UNIVERSITÀ DEGLI STUDI DI MILANO

Facoltà di scienze matematiche, fisiche e naturali

Corso di Laurea in Matematica

CLASSICAL AND RELATIVISTIC DYNAMICS OF EXTRASOLAR PLANETARY SYSTEMS

Relatore:
Prof. Antonio Giorgilli
Correlatore:
Dott. Marco Sansottera

Tesi di Laurea di:
Lorenzo Grassi
Matr. n. 808547

Anno Accademico 2012 - 2013

Contents

Preface **iv**

1 Celestial mechanics **1**

1.1 The two-body problem 2

1.1.1 Central force problem 2

1.2 Orbital elements 4

1.2.1 Determination of the orientation of the ellipse in the space 7

1.3 The problem of N bodies 8

2 Relativistic celestial mechanics **11**

2.1 Special Relativity 12

2.2 General Relativity 13

2.2.1 Distances and time intervals 14

2.2.2 Einstein field equations 15

2.2.3 Motion of a particle in a gravitational field 16

2.2.4 Newton's law 17

2.3 The gravitational field equations 18

2.3.1 Schwarzschild metric 18

2.3.2 Gravitational fields at large distances from bodies 21

2.4 The equation of motion of a system of bodies in the second approximation 22

2.4.1 Relativistic Hamiltonian of N bodies 26

2.4.2 Lense-Thirring effect 28

3 Numerical integration of the planetary problem **32**

3.1 Canonical transformations 33

3.2 Barycentric coordinates 34

3.3 Heliocentric coordinates 35

3.4 Transformation of the Hamiltonian in heliocentric coordinates 36

3.4.1 Classical Hamiltonian 36

3.4.2 Relativistic Hamiltonian 37

3.5 Numerical integration 39

3.6 Application to some extrasolar systems 40

3.7 Simplification of the relativistic Hamiltonian 53

4 Hamiltonian systems **60**
4.1 Integrable Hamiltonian 60
4.1.1 The theorem of Arnold-Jost 61
4.1.2 Periodic and quasi-periodic motion on a torus 63
4.2 Delaunay variables 64
4.3 Quasi-integrable Hamiltonian 68
4.3.1 Lindstedt method 68
4.4 The averaging principle 70
4.5 Introduction to perturbation theory 71
4.5.1 Lie series approach 72
4.6 Birkhoff normal form 74
4.7 The small divisors problem 76
4.8 Beyond the first order 78
4.8.1 Practical calculation 79

5 Development of the Hamiltonian in Poincaré variables **80**
5.1 Hamilton equations in Poincaré variables 80
5.1.1 D'Alembert rules 82
5.2 Conservation of the angular momentum and the reduced three-body problem 83
5.3 Development of the classical Hamiltonian in the reduced problem 86
5.3.1 Development of the Keplerian movement 86
5.3.2 Development of the inverse of the distance 89
5.3.3 Reduction to the plane problem 91
5.3.4 General case 93
5.3.5 Kinetic part 94
5.3.6 Final form of the classical Hamiltonian 95
5.4 Development of the relativistic Hamiltonian in the reduced problem 95
5.4.1 Relativistic precession of pericenter 98
5.4.2 Final form of the relativistic Hamiltonian 99

6 Secular evolution in action-angle coordinates **100**
6.1 The invariance of the semi-major axes according to Lagrange and Poisson 101
6.2 Secular Hamiltonians 103
6.2.1 Expansion of the Hamiltonians 103
6.2.2 Averaging over the mean motions 104
6.3 Secular dynamics of the planets 105
6.3.1 Averaged Hamiltonians in diagonal form 105
6.3.2 Use of action-angle variables 108
6.4 Secular evolution in action-angle coordinates 109
6.4.1 Birkhoff's normal form 110
6.4.2 Analytical integration 111
6.5 Application to some extrasolar systems 111
6.5.1 Comparison between semi-analytical integration and direct numerical integration 112
6.6 A priori criterion to determine the importance of the relativistic correction 118
6.7 Conclusions 124

A Summary of differential geometry **127**

B Hansen coefficients **131**

Bibliography **133**

Preface

In the present thesis, we study the classical and the relativistic (secular) dynamics of some extrasolar planetary systems consisting of two non-resonant coplanar planets. The aim is to evaluate how General Relativity (GR) affects the orbital dynamics of some extrasolar systems, in the limit of point masses, by means of numerical and semi-analytical integrations of the equations of motion.

An *extra-solar planet*, or *exoplanet*, is a planet in an orbit around a star different from the Sun. The first confirmation of an exoplanet orbiting a main-sequence star was made in 1995. Since then the number of extrasolar planets did not cease to grow and today more than a thousand such planets have been discovered.

For the discovery of these exoplanets, many teams have used many techniques like in particular radial velocity measurements, transit method, astrometry and direct imaging. As the discoveries are recent and many of the discovery planets are at the edge of observational capabilities, the uncertainties on their orbital elements and masses are large.

In the current state of art, we are just capable of discovering big planets with not large periods. Therefore, the planets so far discovered have large masses (about three quart have masses in the range $0.15-11\ M_J$) and most of them have orbits close to the central stars. In particular, about half of them are orbiting their host star much closer than Mercury orbits the Sun (0.39 AU). Another characteristic is the large eccentricities of more of them. Indeed, in contrast to the Solar system where the orbits of the planets are almost circular, the exoplanets usually describe true ellipses with high eccentricities. In particular, more than one third have significantly elliptical orbits, with $e > 0.3$, compared with the largest eccentricities in our Solar system, of about 0.2 for Mercury and 0.05 for Jupiter. Thus, the architectures of the extrasolar systems are diverse and, usually, very different from the Solar system configuration.

The recent discoveries of extrasolar planets provide us with a new ensemble of planetary systems to study relativistic effects. Indeed, except for very precise simulations, in the study of the dynamics of Solar system we just deal with Newtonian mechanics and the relativistic effects are in general not taken into account in orbit computations. The effects of relativity are in fact so small as not to worry about it and furthermore the perturbations due to the larger asteroids are almost always much more significant. Moreover, when the relativistic effects are taken into account, only the effects due to the Sun are considered and, also in this case, the secular relativistic effects generated by the Sun are appreciable only for the argument of the perihelion and mean anomaly of the inner Solar system. In particular, the most famous example of a situation where "relativity is important" for orbit computations is the precession of the orbit of Mercury.

Conversely, by virtue of their small semi-major axes and high eccentricities, extrasolar planetary systems with multiple planets allow for General Relativity to exhibit much more pronounced effects than in the case of the Solar system. Due to the characteristics of extrasolar systems, the relativistic corrections due to the star (and in some particular cases the relativistic effects generated by the planets) are in fact important and, in some cases, are indispensable (specially when semi-major axis are of the order of 10^{-1} AU or less) for orbits computation. As a result, these systems provide a new test of General Relativity.

The first aim of this thesis is to perform an extensive comparison between the classical and the relativistic (secular) dynamics of some extrasolar planetary systems consisting of two non-resonant coplanar planets, by a numerical integrations of the equations of motion.

In particular, we are interested in the first corrections to the classical Newtonian equations of motion of a system of point masses, as derived from the GR. They can be found by means of an approximation of this last theory in the case of weak gravitational fields and low velocities, known as the Post Newtonian approximation. This approximation can be accomplished by making an expansion of the GR equations in terms of $\mathbf{v}/c$, where $\mathbf{v}$ are the velocities and c is the speed of light. The PN Hamiltonian of an isolated system of point masses include the Newtonian term and a number of other terms which are suppressed by factors of $1/c$. The equations of motion corresponding to this Hamiltonian are known as the *Einstein-Infeld-Hoffman equations*, and their first version is found in *Einstein et al. (1938)*.

As expected, the results of numerical integration show that relativistic corrections become very important for systems with the innermost planet close to the star, with the other body relatively distant. Moreover, we note that the relativistic corrections seem to provide "stability" to the system, in the cases in which they are important.

The numerical integration of the equations of motion corresponding to the PN Hamiltonian is greatly slow, and it is useful only in situations where the relativistic contribution of every object in the system is to be taken into account. For these reasons, the second aim is to introduce an approximate model of the relativistic Hamiltonian that takes into account the main relativistic effects and that decreases considerably the calculation time. In particular, the PN Hamiltonian can be greatly simplified in the case in which only one object (e.g. the central star) contributes with relevant corrections, skipping the relativistic correction due to the mutual interactions of the two planetary masses.

The simplified relativistic Hamiltonian, although less exact than the PN Hamiltonian, is computationally much more affordable and, as we will show, the dynamic described by the simplified Hamiltonian is very similar to the real one, at least numerically in the systems that we have considered.

On the other hand, the major defect of the numerical integration is that it is CPU consuming and that the time required to integrate a system is very long. The third aim of this thesis is therefore to reconstruct the evolution of the eccentricities (and pericenters) of the planets by using analytical techniques, both in the classical than in the (simplified) relativistic case, extending the Laplace-Lagrange theory.

The classical Laplace-Lagrange theory for the secular motions of the planetary orbits uses the circular approximation as a reference for the orbits and it is based only on a linear approximation of the dynamical equations. Because the orbits of the planets in Solar system are almost circular while the exoplanets usually describe true ellipses with high eccentricities, the applicability of the classical approach, using the circular approximation as a reference, can be doubtful for these systems.

Previous works of *Libert* & *Henrard (2005, 2006)* for coplanar systems have generalized the classical expansion of the perturbation potential to a higher order in the eccentricities, showing that this analytical model gives an accurate description of the behavior of planetary systems which are not close to a mean-motion resonance, up to surprisingly high eccentricities. Moreover, they have shown that an expansion up to order 12 in the eccentricities is usually required for reproducing the secular behavior of extrasolar planetary systems. These results have been obtained considering a secular Hamiltonian at order one in the masses.

In order to study the Hamiltonians with the classical perturbation theory, we rewrite the classical and the (simplified) relativistic Hamiltonian using the *Poincaré variables*. Then we simplify the two Hamiltonians using the averaging principle, which corresponds to fixing the values of the semi-major axes. It is important to remember that, because we consider only non-resonant extrasolar systems, using the averaging principle we can still obtain qualitative information on the long-term changes of the slowly varying orbital elements. Finally, following the works of *Libert and Henrand*, we study the classical and the relativistic secular Hamiltonian using the modern perturbation method based on the Lie series and on the Birkhoff's normal form. This approach allows us to derive a fully-analytical description of the system using semi-automatized computer algebra.

To validate our results, we compare our semi-analytical integration with the direct numerical integration. In particular, we find that the agreement between the results obtained with the two methods is excellent.

Furthermore, evaluating the difference between the quadratic part of the secular classical Hamiltonian and of the secular relativistic Hamiltonian, we have set up a simple (and rough) criterion to discriminate a priori between the cases in which the relativistic corrections are important from those in which they are not (in the case of a coplanar non-resonant three-body system). In particular, our way to quantify the efficacy of GR on the i-th planet is through the dimensionless parameter

$$\frac{c^2 a_i^2 m_1 m_2}{\mathcal{G} a_2 m_0 m_i (m_0 + m_i)} \left[\frac{3}{8} \left(\frac{a_1}{a_2} \right)^2 + \frac{45}{64} \left(\frac{a_1}{a_2} \right)^4 + \frac{525}{512} \left(\frac{a_1}{a_2} \right)^6 \right],$$

where we use the index 1 to indicate the innermost planet and the index 2 to indicate the outer planet (see chapter 6 for more details).

The thesis is organized as follows. Chapter 1 is dedicated to an overview of Celestial mechanics and in particular of the orbital elements, i.e. the quantities that characterize the geometrical property of the orbital ellipse and the position on the ellipse with respect to a fixed reference system. In chapter 2, after a few recalls of General Relativity, we derive the Post Newtonian Hamiltonian, starting from the Einstein equation. To do this, we follow in particular the work of *Einstein et al. (1938)* and *Landau et al. (1971)*. Then, in chapter 3, we study numerically the dynamics described by the classical and relativistic Hamiltonian for some extrasolar systems. The results show how relativistic effects can accumulate over time to induce substantial changes in the dynamics. Then, we derive the simplified relativistic Hamiltonian and we compare numerically the dynamics described by the real and the simplified relativistic Hamiltonian. In chapter 4 are given the tools provided by Hamiltonian's theory and by perturbation theory, that we will use in following chapters. In particular, we present the construction of the Poincaré variables, the averaging principle and the modern perturbation method based on the Lie series and on the Birkhoff's normal form for the study of an isochronous Hamiltonian. In chapter 5 we present a method for the expansion of the classical and relativistic Hamiltonian in planar Poincaré variables, which can be implemented in a straightforward manner on a computer. Following the Lagrange approach, we focus, in chapter 6, on the secular part of the Hamiltonians and we construct a high-order Birkhoff normal form, using the Lie series method, that leads to a very simple form of the equations of motion, being function of the actions only. We apply this method on systems previously studied and we compare our semi-analytical integration with the direct numerical integration. Finally we set up a simple criterion to discriminate between the cases in which the relativistic corrections are important from those in which they are not. An appendix containing a summary of differential geometry and another containing the definition of the Hansen coefficients follow.

Chapter 1

Celestial mechanics

Celestial mechanics is the branch of astronomy that deals with the motions of celestial objects. Johannes Kepler was the first to formulate three scientific laws describing orbital motion, originally formulated to describe the motion of planets around the Sun. He worked as an assistant to the Danish astronomer Tycho Brahe. Brahe took extraordinarily accurate measurements of the motion of the planets of the Solar System. From these measurements, Kepler was able to formulate Kepler's laws, which are:

First law: Each planet moves, relative to the Sun, in an elliptical orbit, with the Sun at one of the two foci of the ellipse.

Second law: The rate of motion in the elliptical orbit is such that the vector pointing to the position of the planet relative to the Sun spans equal areas of the orbital plane in equal times.

Third law: The square of the orbital period T is proportional to the cube of the semi-major axis a of the orbital ellipse.

Kepler published the first two laws in 1609 and the third law in 1619.

Nearly a century later, Isaac Newton had formulated his three laws of motion. They describe the relationship between the forces acting upon a body and its motion in response to said forces. They can be summarized as follows:

First law: When viewed in an inertial reference frame, an object either is at rest or moves at a constant velocity, unless acted upon by an external force.

Second law: The acceleration of a body is directly proportional to, and in the same direction as, the net force acting on the body, and inversely proportional to its mass:

$$\mathbf{F} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = m\mathbf{a}.$$

Third law: *(Principle of action and reaction)* When one body exerts a force on a second body, the second body simultaneously exerts a force equal in magnitude and opposite in direction to that of the first body.

Newton also posed the question of what force produces the elliptical orbits seen by Kepler. He came to formulate his law of universal gravitation, which is the first correct scientific and mathematical formulation of *gravity*. *Newton's law of universal gravitation* states that every point mass in the universe attracts every other point mass with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them. Mathematically

$$\mathbf{F}_{1,0}(\mathbf{x}_0, \mathbf{x}_1) = \mathcal{G}\frac{m_0 m_1}{\parallel \mathbf{x}_1 - \mathbf{x}_0 \parallel^3}(\mathbf{x}_0 - \mathbf{x}_1) \tag{1.1}$$

where $\mathbf{F}$ is the force with which the object 0 is attracted by the object 1, $\mathcal{G}$ is the gravitational constant ($\mathcal{G} = 6.67 \times 10^{-11}\,\mathrm{N\ m^2\ Kg^{-2}}$), m_0 and m_1 are the two masses and $\mathbf{x}_0$ and $\mathbf{x}_1$ denote the position vectors of two bodies in an inertial reference frame.[1] Given this force law and his equations of motion, Newton was able to show that two point masses attracting each other would each follow perfectly elliptical orbits.

1.1 The two-body problem

According to Newton's theory of gravitation, the equations of motion of an isolated system of two bodies having spherical symmetry and mass m_0 and m_1 are

$$\frac{\mathrm{d}^2\mathbf{x}_0}{\mathrm{d}t^2} = \frac{\mathcal{G}m_1}{\parallel \mathbf{x}_1 - \mathbf{x}_0 \parallel^3}(\mathbf{x}_1 - \mathbf{x}_0) \qquad \frac{\mathrm{d}^2\mathbf{x}_1}{\mathrm{d}t^2} = \frac{\mathcal{G}m_0}{\parallel \mathbf{x}_1 - \mathbf{x}_0 \parallel^3}(\mathbf{x}_0 - \mathbf{x}_1) \tag{1.2}$$

where, as before, $\mathbf{x}_0$ and $\mathbf{x}_1$ denote the position vectors of the two bodies in an inertial reference system.

Denoting by $\mathbf{r} = \mathbf{x}_1 - \mathbf{x}_0$ the relative position of the bodies and with $\mathbf{s} = \frac{m_0\mathbf{x}_0 + m_1\mathbf{x}_1}{m_0 + m_1}$ the barycenter of the system, the two vectorial equations above can be reduced to two separate vectorial equation:

$$\begin{aligned} \frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} &= -\mathcal{G}\frac{m_0 + m_1}{\parallel \mathbf{r} \parallel^3}\mathbf{r}, \\ \frac{\mathrm{d}^2\mathbf{s}}{\mathrm{d}t^2} &= 0. \end{aligned} \tag{1.3}$$

The first equation describes a central force problem, while the second shows that the velocity of the center of mass is constant.

1.1.1 Central force problem

We are now reduced to determine the motion of a particle under the influence of a single central force:

$$\mu\ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r}) \tag{1.4}$$

where the force $\mathbf{F}(\mathbf{r})$ is directed as $\mathbf{r}$ (i.e. $\mathbf{r} \times \mathbf{F} = 0$) and μ is the *reduced mass* defined as

$$\mu = \left(\frac{1}{m_0} + \frac{1}{m_1}\right)^{-1} = \frac{m_0 m_1}{m_0 + m_1}. \tag{1.5}$$

It can be proven that the motion of a particle under a central force $\mathbf{F}$ always remains in the plane defined by its initial position and velocity, and that the areal velocity respect to the center

[1] For simplicity, I will adopt the vectorial formalism, denoting by $\mathbf{x}$ the n-uples $x_1, x_2, ...x_n$. Moreover, I will denote by $\dot{\alpha}$ the time derivative of a generic variable α.

is constant. To prove this mathematically, it is sufficient to show that the angular momentum $\mathbf{M} = \mathbf{r} \times \mu\dot{\mathbf{r}}$ of the particle is constant (i.e. $\mathbf{M}$ is a first integral, or a constant of motion). Since the motion is planar and the force radial, it is customary to choose a reference system such that the axis z is oriented as $\mathbf{M}$, so that the motion happen in the plane x, y. Moreover, it's convenient to switch to polar coordinates r, θ with the transformation $x = r\cos\theta$ and $y = r\sin\theta$.

Since $\mathbf{F} = m\mathbf{a}$ by Newton's second law of motion and since $\mathbf{F}$ is a central force, then only the radial component of the acceleration $\mathbf{a}$ can be non-zero. In the orbital plane, the system (1.3) becomes

$$\begin{aligned} \mu(\ddot{r} - r\dot{\theta}^2) &= F \\ \frac{\mathrm{d}}{\mathrm{d}t}(\mu r^2\dot{\theta}) &= 0, \end{aligned} \tag{1.6}$$

where F is the radial component of the force $\mathbf{F}$.

To solve the problem, we introduce the new variable $w = 1/r$. With this new variable the system (1.6) becomes

$$\begin{aligned} \frac{\mathrm{d}^2 w}{\mathrm{d}\theta^2} + w &= -\frac{\mu F}{L^2 w^2}, \\ \dot{\theta} &= \frac{L}{\mu} w^2, \end{aligned} \tag{1.7}$$

where L is the component of the angular momentum $\mathbf{M}$ along axis z. The first equation describe the orbit in the form $w = w(\theta)$, while the second gives the movement on the form $\theta = \theta(t)$.

In the Keplerian case, $F = -\mathcal{G}\mu(m_0 + m_1)w^2$ and the first of equations (1.7) becomes

$$\frac{\mathrm{d}^2 w}{\mathrm{d}\theta^2} + w = \frac{\mathcal{G}\mu^2(m_0 + m_1)}{L^2}. \tag{1.8}$$

The solution is

$$w(\theta) = w_1 + w_2\cos(\theta - \theta_0), \qquad \theta_0 = \theta(t_0), \tag{1.9}$$

where w_1 and w_2 depend on the initial data and where t_0 is the initial time. Finally, remembering that $r = 1/w$, the equation (1.9) can be rewritten as

$$r = \frac{p}{1 + e\cos(\theta - \theta_0)}. \tag{1.10}$$

If $0 \leq e < 1$ the orbit is an ellipse of eccentricity e with the center of the force at the focus of the conic section. In this case, it is also interesting to calculate the ratio $2\pi/T$, where T is the period of revolution of the particle. The period is calculated using the fact that the areal velocity $\dot{A}$ is constant:

$$\dot{A} = \frac{1}{2}r\dot{\theta}^2 = \frac{L}{2\mu} = \frac{ab\pi}{T}, \tag{1.11}$$

where a and b are respectively the length of the semi-major and semi-minor axis of the ellipse ($b = a\sqrt{1 - e^2}$). With elementary calculations one has:

$$\frac{2\pi}{T} = \sqrt{\mathcal{G}(m_0 + m_1)}a^{-3/2}. \tag{1.12}$$

Finally, to obtain $\theta = \theta(t)$ is sufficient to invert the equation

$$t - t_0 = \frac{\mu}{L} \int_{\theta_0}^{\theta} [w_1 + w_2 \cos(\psi - \theta_0)]^{-2} \,\mathrm{d}\psi. \tag{1.13}$$

The solution of the equation (1.3) is

$$\mathbf{r}(t) = \frac{p}{1 + e\cos(\theta(t) - \theta_0)} \left[\cos(\theta(t) - \theta_0)\hat{\mathbf{r}}_0 + \sin(\theta(t) - \theta_0)\hat{\boldsymbol{\theta}}_0 \right], \tag{1.14}$$

where

$$\hat{\mathbf{r}}_0 = \frac{\mathbf{x}_1(t_0) - \mathbf{x}_0(t_0)}{\| \mathbf{x}_1(t_0) - \mathbf{x}_0(t_0) \|}, \qquad \hat{\boldsymbol{\theta}}_0 = \frac{\mathbf{M}}{\| \mathbf{M} \|} \times \hat{\mathbf{r}}_0. \tag{1.15}$$

Concluding, the trajectories of the two bodies are:

$$\mathbf{x}_0(t) = \mathbf{s}(t) - \frac{m_0}{m_0 + m_1}\mathbf{r}(t) \qquad \mathbf{x}_1(t) = \mathbf{s}(t) + \frac{m_1}{m_0 + m_1}\mathbf{r}(t) \tag{1.16}$$

where $\mathbf{s}(t) = \dot{\mathbf{s}}(t_0)(t - t_0) + \mathbf{s}(t_0)$ and $\mathbf{r}(t)$ is defined in (1.14).

It is important to notice that the two bodies describe similar trajectories around the common barycenter and that the ratio between the size of the two orbits is inversely proportional to the mass ratio. As a consequence, when the mass ratio tends to zero, as in the case of small body and a star, the star's orbit shrinks to the barycenter's position: in this case, it is convenient to choose as the central body the more massive of the two bodies, although from the mathematical point of view the choice is arbitrary.

The fact that the orbit in a Keplerian potential is closed is due to the existence of a further first integral, known as the *Runge-Lenz vector*. This vector is

$$\mathbf{A} = \mathbf{v} \times \mathbf{M} - \frac{\mathcal{G} m_0 m_1 \mathbf{r}}{\| \mathbf{r} \|} \tag{1.17}$$

where $\mathbf{v} = \dot{\mathbf{r}}$ and $\mathbf{M}$ is the angular momentum ($\mathbf{M} = \mu\mathbf{r} \times \mathbf{v}$). It is simple to prove that the direction of the vector $\mathbf{A}$ is along the major axis from the focus to the perihelion, and that its magnitude is $\| \mathbf{A} \| = \mathcal{G} m_0 m_1 e$, where e is the eccentricity of the ellipse.

1.2 Orbital elements

It is convenient for astronomers to characterize the relative motion of the two bodies by quantities that describe the geometrical property of the orbital ellipse and the position on the ellipse. These quantities are called *orbital elements*.

The shape of the ellipse can be completely determined by two quantities: the semi-major axis a and the semi-minor axis b, or, equivalently, the semi-major axis a and the eccentricity e. The *eccentricity* e is defined as the ratio between the distance of the focus from the center of the ellipse and the semi-major axis of the ellipse:

$$e = \sqrt{1 - \frac{b^2}{a^2}} \qquad b = a\sqrt{1 - e^2}. \tag{1.18}$$

It is therefore an indicator of how much the orbit differs from a circular one: $e = 0$ means that the orbit is circular while $0 < e < 1$ denote an elliptic orbit. On an elliptic orbit, the closest point to

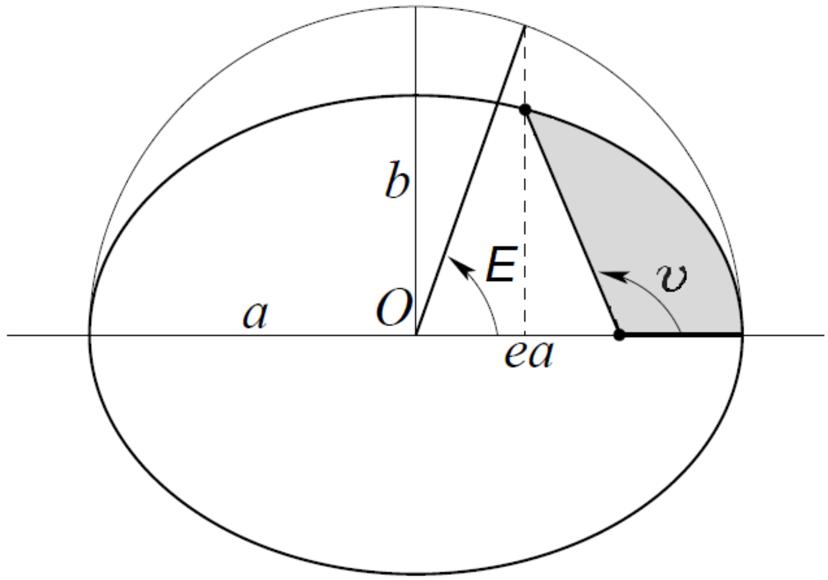


Figure 1.1: Keplerian motion: definition of a, e, ν and E. Reprinted from Fig. 1.3 of Giorgilli, *Meccanica celeste*.

the central body is called *pericenter* (or alternatively *perihelion* if the central body is the Sun) and its distance is equal to $a(1-e)$; the farthest point is called the *apocenter* (respectively *aphelion*) and its distance is equal to $a(1+e)$.

To denote the position of the body on the elliptic orbit it is convenient to use polar coordinate r and ν, where $\nu = 0$ is oriented towards the pericenter of the orbit. The angle ν is called *true anomaly* of the body. In this case, the equation of the ellipse is:

$$r = \frac{a(1-e^2)}{1+e\cos\nu}. \tag{1.19}$$

Alternatively, we can use an orthogonal reference frame $(\mathcal{X}, \mathcal{Y})$ with the origin at the focus of the ellipse occupied by the Sun and $\mathcal{X}$ axis oriented towards the pericenter of the orbit. With elementary geometrical relationships one has

$$\mathcal{X} = r\cos\nu \qquad \mathcal{Y} = r\sin\nu \tag{1.20}$$

where r is given by (1.19).

To describe the motion of the planet along the orbit, it's convenient to introduce a new angle E called *eccentric anomaly*, as Fig. 1.1 shows. E is the angle subtended at the center of the ellipse by the projection of the position of the body on the circle with radius a and tangent to the ellipse at pericenter and apocenter. The following relationships between E and ν are all equivalent

$$\begin{aligned} &(1-e\cos E)(1-e\cos\nu) = 1-e^2, \\ &\tan\frac{E}{2} = \sqrt{\frac{1-e}{1+e}}\tan\frac{\nu}{2}, \\ &\sin E = \frac{\sqrt{1-e^2}\sin\nu}{1+e\cos\nu}. \end{aligned} \tag{1.21}$$

The equation of the orbit relative to the angle E becomes:

$$r = a(1 - e\cos E) \tag{1.22}$$

or using the orthogonal reference frame $(\mathcal{X}, \mathcal{Y})$

$$\begin{aligned} \mathcal{X} &= a(\cos E - e), \\ \mathcal{Y} &= a\sqrt{1-e^2}\sin E. \end{aligned} \tag{1.23}$$

The position of a body in its orbit can be expressed in terms of a, e and E only. Using the fact the areal velocity $\dot{A}$, defined in (1.11), is constant, it is possible to derive the evolution law of E with respect to the time, usually called *Kepler equation*

$$E - e\sin E = n(t - t_0), \tag{1.24}$$

where

$$n = \frac{2\pi}{T} = \sqrt{\mathcal{G}(m_0 + m_1)}a^{-3/2} \tag{1.25}$$

is the orbital frequency or *mean motion* of the body, T is the period of revolution of the planet, t is the time and t_0 is the time of passage of pericenter.

It is also useful to introduce a new angle

$$M = n(t - t_0) \tag{1.26}$$

called the *mean anomaly* as an orbital element that changes linearly with time and still denotes the position of the body in its orbit.

Kepler equation is a transcendental equation, thus it cannot be solved with simple mathematical method. Using the Fourier series, it can be proven that the solution of Kepler equation is

$$E(M) = M + 2\sum_{k=1}^{\infty}\frac{1}{k}J_k(ke)\sin(kM), \tag{1.27}$$

where $J_k(x)$ is the *Bessel function* of order k, defined as[2]

$$J_k(x) = \frac{1}{\pi}\int_0^{\pi}\cos[k(z - x\sin z)]\mathrm{d}z = \left(\frac{x}{2}\right)^k\sum_{h=0}^{\infty}(-1)^h\frac{(x^2/4)^h}{h!(h+k)!}. \tag{1.28}$$

Finally, to find the components of velocity of the body on the elliptic orbit, it is sufficient to derive $\mathcal{X}$ and $\mathcal{Y}$ respect to time:

$$\begin{aligned} \dot{\mathcal{X}} &= -a\dot{E}\sin E, \\ \dot{\mathcal{Y}} &= a\dot{E}\sqrt{1-e^2}\cos E. \end{aligned} \tag{1.29}$$

[2] It is simple to prove the following useful properties (see *Whittaker* for more details):

$$J_k(-x) = (-1)^k J_k(x) = J_{-k}(x), \qquad kJ_k(x) = \frac{x}{2}(J_{k-1}(x) + J_{k+1}(x)).$$

These properties show that it is sufficient to know J_0 and J_1 to calculate the other J_k recursively.

To apply these formulæ, we need to know $\dot{E}$. To do this, we derive the Kepler equation respect to time:

$$\dot{E}(1 - e\cos E) = n \tag{1.30}$$

and using the equation (1.22) of the ellipse, we obtain

$$\dot{E} = \frac{na}{r} = \sqrt{\frac{\mathcal{G}(m_0 + m_1)}{r^2 a}}, \tag{1.31}$$

where $r = \sqrt{\mathcal{X}^2 + \mathcal{Y}^2}$. The result is

$$\begin{aligned} \dot{\mathcal{X}} &= -\sqrt{\frac{\mathcal{G}(m_0 + m_1)}{a}} \frac{2\pi \sin E}{1 - e\cos E} = -\sqrt{\frac{\mathcal{G}(m_0 + m_1)}{a(1 - e^2)}} \sin\nu, \\ \dot{\mathcal{Y}} &= \sqrt{\frac{\mathcal{G}(1 - e^2)(m_0 + m_1)}{a}} \frac{2\pi \cos E}{1 - e\cos E} = \sqrt{\frac{\mathcal{G}(m_0 + m_1)}{a(1 - e^2)}} (e + \cos\nu), \end{aligned} \tag{1.32}$$

where it has been used the relation between ν and E.

1.2.1 Determination of the orientation of the ellipse in the space

To characterize the orientation of the ellipse in space, with respect to an arbitrary orthogonal reference frame (x, y, z) centered on the position of the central body, we have to introduce three additional angles.

The first one is the *inclination* i of the orbital plane (the plane which contains the ellipse) with respect to the (x, y) plane. If the orbit has a nonzero inclination, it intersects the (x, y) plane in two points, called the *nodes* of the orbit. We define the *ascending node* as the node where the body passes from negative to positive z. The orientation of the orbital plane in space is then completely determined when the angular position of the *ascending node* from the x axis is given. This angle is called the *longitude of node* Ω. The last angle that needs to be introduced is the *argument of pericenter* ω. It characterizes the orientation of the ellipse in its plane and it is defined as the angular position of the pericenter, measured in the orbital plane relative to the line connecting the central body to the ascending node.

The orbital elements a, e, i, ω, Ω and M completely define the position $\mathbf{r}$ and the velocity $\mathrm{d}\mathbf{r}/\mathrm{d}t$ of the secondary body with respect to the central one. In particular, there is a one-to-one correspondence between $\mathbf{r} = [r_x, r_y, r_z]$, $\mathrm{d}\mathbf{r}/\mathrm{d}t = [\mathrm{d}r_x/\mathrm{d}t, \mathrm{d}r_y/\mathrm{d}t, \mathrm{d}r_z/\mathrm{d}t]$ and the orbital elements which is given by the relationship:

$$\mathbf{r} = R\mathbf{u} \qquad \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = R\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}, \tag{1.33}$$

where, as we have seen, the vector $\mathbf{u}$ and $\mathrm{d}\mathbf{u}/\mathrm{d}t$ have components

$$\mathbf{u} = [\mathcal{X}, \mathcal{Y}, 0] = [a(\cos E - e), a\sqrt{1 - e^2}\sin E, 0] \tag{1.34}$$

and

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \left[\frac{\mathrm{d}\mathcal{X}}{\mathrm{d}t}, \frac{\mathrm{d}\mathcal{Y}}{\mathrm{d}t}, 0\right] = \left[-\frac{na\sin E}{1 - e\cos E}, \frac{na\sqrt{1 - e^2}\cos E}{1 - e\cos E}, 0\right] \tag{1.35}$$

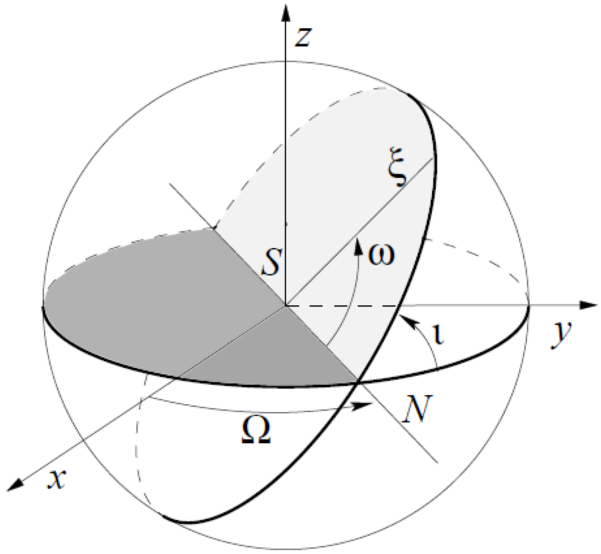


Figure 1.2: Keplerian motion: definition of i, ω and Ω. Reprinted from Fig. 1.9 of Giorgilli, *Meccanica celeste*.

respectively, and the rotation matrix R has entries

$$R = \begin{bmatrix} \cos\Omega\cos\omega - \sin\Omega\cos i\sin\omega & -\cos\Omega\sin\omega - \sin\Omega\cos i\cos\omega & \sin\Omega\sin i \\ \sin\Omega\cos\omega + \cos\Omega\cos i\sin\omega & -\sin\Omega\sin\omega + \cos\Omega\cos i\cos\omega & -\cos\Omega\sin i \\ \sin i\sin\omega & \sin i\cos\omega & \cos i \end{bmatrix}. \tag{1.36}$$

It is simple to prove that $R^{-1} = R^T$.

Note that, in the definition of the orbital elements above, ω and Ω are not defined when the inclination is zero (because the position of the ascending node is not determined), and, moreover, M is not defined when the eccentricity is zero (because the position of the pericenter is not determined). It is convenient, therefore, to introduce the *longitude of pericenter* $\varpi = \omega + \Omega$ and the *mean longitude* $\lambda = M + \omega + \Omega$. The first angle is well defined when $i = 0$, while the second one is well defined when $i = 0$ and/or $e = 0$. It is evident that also the set of orbital elements $a, e, i, \varpi, \Omega, \lambda$ unequivocally defines the position and the velocity of the body.

1.3 The problem of N bodies

In an inertial reference system, the equations of motion of an isolated system made of a star of mass m_0 and of N planets of masses $m_1, m_2, ...m_N$ are:

$$m_i \frac{\mathrm{d}^2\mathbf{x}_i}{\mathrm{d}t^2} = \sum_{j\neq i} \mathbf{F}_{ij} \qquad 0 \le i \le N \tag{1.37}$$

where $\mathbf{x}_0$ is the position of the star, $\mathbf{x}_i$ is the position of the i-th planet and

$$\mathbf{F}_{ij} = -\mathcal{G} m_i m_j \frac{\mathbf{x}_i - \mathbf{x}_j}{\| \mathbf{x}_i - \mathbf{x}_j \|^3}. \tag{1.38}$$

To study these equations, it's useful the knowledge of *constants of motion*, which are quantities that are conserved throughout the motion and that impose, in effect, constraints on the motion. In this case, the constants of motion known are the *energy* E, the *total linear momentum* $\mathbf{P}$ and the *angular momentum* $\mathbf{M}$:

$$\begin{aligned}
E &= \frac{1}{2}\sum_j m_j \dot{\mathbf{x}}_j^2 - \frac{\mathcal{G}}{2}\sum_{j\neq i} m_i m_j \frac{\mathbf{x}_i - \mathbf{x}_j}{\| \mathbf{x}_i - \mathbf{x}_j \|^2}, \\
\mathbf{P} &= \sum_j m_j \dot{\mathbf{x}}_j, \\
\mathbf{M} &= \sum_j \mathbf{x}_j \times m_j \dot{\mathbf{x}}_j.
\end{aligned} \tag{1.39}$$

Introducing the heliocentric position of the planets $\mathbf{r}_i = \mathbf{x}_i - \mathbf{x}_0$, the equation (1.37) can be rewritten as

$$\frac{\mathrm{d}^2\mathbf{r}_i}{\mathrm{d}t^2} = -\frac{\mathcal{G}(m_0 + m_i)}{\| \mathbf{r}_i \|^3}\mathbf{r}_i + \sum_{\substack{j=1\\ j\neq i}}^{N} \mathcal{G} m_j \left(\frac{\mathbf{r}_j - \mathbf{r}_i}{\| \mathbf{r}_j - \mathbf{r}_i \|^3} - \frac{\mathbf{r}_j}{\| \mathbf{r}_j \|^3} \right), \qquad 1 \leq i \leq N, \tag{1.40}$$

while the motion of the star is given by

$$\frac{\mathrm{d}^2\mathbf{x}_0}{\mathrm{d}t^2} = \sum_{i\neq 0} -\mathcal{G} m_i \frac{\mathbf{r}_i}{\| \mathbf{r}_i \|^3}. \tag{1.41}$$

The equations (1.40) are evidently very close to the equations of motion of the two-body problem (1.2). Indeed, if the masses of the planets are small compared to that of the star (i.e. $m_i \ll m_0$ for $1 \leq i \leq N$), and none of their mutual distance $\| \mathbf{r}_i - \mathbf{r}_j \|$ becomes small, equation (1.40) can be well approximated for time not too long as

$$\frac{\mathrm{d}^2\mathbf{r}_i}{\mathrm{d}t^2} \simeq -\frac{\mathcal{G}(m_0 + m_i)}{\| \mathbf{r}_i \|^3}\mathbf{r}_i, \qquad 1 \leq i \leq N \tag{1.42}$$

which are the equations of Keplerian motion.

As a consequence, because the motion resulting by the equations (1.40) will be close to the Keplerian motion, rewriting equations (1.40) in terms of the orbital elements, the equations for a, e, i, ω, Ω of each planet have the form $\mathrm{d}\alpha/\mathrm{d}t = O(m_j/m_0)$, where α denotes any of these elements and $O(m_j/m_0)$ denotes a function which is small as the mass of the planet in solar mass units; the equation for M has the form $\mathrm{d}M/\mathrm{d}t = n + O(m_j/m_0)$ where n is the unperturbed mean motion resulting from the two-body problem. These equations show that the orbital elements a, e, i, ω, Ω change slowly with time, while M deviates slowly from its linear unperturbed motion. Because they are difficult to handle, the idea is to rewrite them in a slightly different form, a Hamiltonian form, in order to study the motion in detail.

In particular, it is very easy to write the Hamilton's equations for systems that conserve a quantity (called "energy") which in variables $\mathbf{x}$ and $\dot{\mathbf{x}}$ can be written as the sum of a "kinetic energy" $T(\mathbf{x}, \dot{\mathbf{x}})$ and of a "potential energy" $U(\mathbf{x})$. Indeed, if we can decompose T as $T = T_2 + T_1 + T_0$, where T_2, T_1 and T_0 are respectively the terms of motion of order 2, 1 and 0 in $\dot{\mathbf{x}}$, the equations of motion can be written as

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\partial H(\mathbf{x}, \mathbf{p})}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H(\mathbf{x}, \mathbf{p})}{\partial x_i} \tag{1.43}$$

for $0 \leq i \leq N$, where p_i are the *momenta* conjugates to the coordinates x_i

$$p_i = \frac{\partial T}{\partial \dot{x}_i}, \qquad 0 \leq i \leq N \tag{1.44}$$

and where the function H is the *Hamiltonian* of the system:

$$H(\mathbf{x}, \mathbf{p}) = T_2(\mathbf{x}, \dot{\mathbf{x}})\Big|_{\dot{\mathbf{x}}=\dot{\mathbf{x}}(\mathbf{x},\mathbf{p})} - T_0(\mathbf{x}) + U(\mathbf{x}). \tag{1.45}$$

The Hamiltonian of $N+1$ of "point-like" bodies, which interact with one other but with no others bodies, coincides with the total energy of the system, i.e.

$$H = \sum_{i=0}^{N} \frac{\mathbf{p}_i \cdot \mathbf{p}_i}{2m_i} - \mathcal{G} \sum_{i=0}^{N} \sum_{\substack{j=0 \\ j \neq i}}^{N} \frac{m_i m_j}{2 \parallel \mathbf{x}_i - \mathbf{x}_j \parallel}, \tag{1.46}$$

where $\mathbf{p}_i = m_i \dot{\mathbf{x}}_i$, for $0 \leq i \leq N$. It is simple to prove that the Hamilton's equations

$$\begin{aligned} \dot{\mathbf{x}}_i &= \frac{\mathbf{p}_i}{m_i}, \\ \dot{\mathbf{p}}_i &= -\sum_{j \neq i} \mathcal{G} m_i m_j \frac{\mathbf{x}_i - \mathbf{x}_j}{\parallel \mathbf{x}_i - \mathbf{x}_j \parallel^3}, \end{aligned} \tag{1.47}$$

for $0 \leq i \leq N$, are equivalent to the equations of motion (1.37).

Chapter 2

Relativistic celestial mechanics

Newton's law of gravitation soon became accepted because it gave very accurate predictions of the orbits of all the planets. In particular, the highest apotheosis of the Newtonian theory of gravitation came when the planet Neptune was mathematically predicted before it was directly observed. By 1846 the planet Uranus had completed nearly one full orbit since its discovery in 1781, and astronomers had detected a series of irregularities in its path which could not be entirely explained by Newton's law of gravitation. These irregularities could, however, be resolved if the gravity of a farther, unknown planet were disturbing its path around the Sun. The astronomers began calculations to determine the nature and position of such a planet, which was finally discovered in 1846 by Urbain Le Verrier.

In 1859 Le Verrier discovered that the orbital precession of the planet Mercury (roughly 574.4" of rotation per century[1]) was not quite what it should be; the ellipse of its orbit was rotating (precessing) slightly faster than predicted by the traditional theory of Newtonian gravity, even after all the effects of the other planets had been accounted for. The effect is small (roughly 43" of rotation per century, approximately 7.5% of the total), but well above of the measurement error. Several classical explanations were proposed, such as interplanetary dust, unobserved oblateness of the Sun, an undetected moon of Mercury, or a new planet named Vulcan. After these explanations were discounted, some physicists were driven to the more radical hypothesis that Newton's inverse-square law of gravitation was incorrect. For example, some physicists proposed a power law with an exponent that was slightly different from 2. A number of ad hoc and ultimately unsuccessful solutions were proposed, but they tended to introduce more problems.

Another possible explanation was given by Laplace in his treatise on celestial mechanics (see *Laplace, 1805*). He had shown that if the gravitational influence does not act instantaneously but it does propagate at a finite speed, then a planet is attracted to a point where the Sun was some time before, and not towards the instantaneous position of the Sun. On the assumption of the classical fundamentals, Laplace had shown that if gravity would propagate at a velocity on the order of the speed of light then the Solar system would be unstable, and would not exist for a long time. The observation that the Solar system is old allows one to put a lower limit on the speed of gravity that is many orders of magnitude faster than the speed of light. Around 1904 – 1905, the works of H. Lorentz, H. Poincaré and A. Einstein's special theory of relativity, exclude the possibility of propagation of any effects faster than the speed of light. It followed that Newton's

[1]The arcsecond is defined as 1/3 600 of an degree, i.e. as $\pi/648\,000$ radians. The standard symbol for marking the arcsecond is the double prime ".

law of gravitation would have to be replaced with another law, which reduced to the Newtonian one when relativistic effects are negligible. This new law is precisely the Einstein's General Theory of Relativity, which explains, among other things, the remaining precession of Mercury or the change of orientation of the orbital ellipse within its orbital plane[2].

The *General Theory of Relativity* (GR) has been one of the greatest achievements of the XX century. Its formulation has revolutionized our way to understand the physical Universe and has led, for example, to a consistent theory describing the global behavior of the Universe, which was impossible in the framework of Newtonian mechanics.

In Einstein's theory the space-time is thought as a four dimensional pseudo-Riemannian manifold whose geometry interacts non-linearly with matter. Because of this feature, general calculations in GR require much more work than the corresponding ones in Newtonian gravity. Another consequence of this difficulty concerns the validation of GR: our inability to solve in general the gravitational field equation limits the possibility of devising full test of GR. Indeed, most of the investigation on the validity of this theory was performed in a regime of weak field leaving the strong field regime almost completely unexplored.

Einstein himself was well aware of the fact that the understanding of his theory in the perturbation regime was crucial for its testing and its application to the problems in which Newtonian gravity was most successful. For this reason in the 1930s he developed a perturbation approach able to describe the subtle changes induced on the Newtonian evolution of the bodies in the Solar system by General Relativity. The so-called *Einstein-Infeld-Hoffman (EIH) equations* are the results as first *Post-Newtonian* level of this attempt. The derivation of these equations was the birth of a new field called *Relativistic Celestial Mechanics.*

2.1 Special Relativity

Special Relativity is a fundamental theory concerning space and time, developed by Albert Einstein in 1905 as a modification of Galilean relativity.

It is based on two postulates, which are the *Principle of Relativity* and the *Principle of Invariant Light Speed.*

The *Principle of Relativity* states that all the laws of nature are identical in all inertial system of reference, i.e the equations expressing the laws of nature are invariant with respect to transformations of coordinate and time from one inertial system to other.

The *Principle of Invariant Light Speed* states that light is always propagated in empty space with a definite velocity c which is independent of the state of motion of the emitting body. Moreover, it states that the velocity of propagation of interactions coincides with the velocity of light in empty space, whose numerical value is

$$c = 2.997\,93 \times 10^{8}\,\mathrm{m\,s^{-1}}. \tag{2.1}$$

From the principle of relativity, it follows in particular that the velocity of light and the velocity of propagation of interactions are the same in all inertial system of reference.

Thus the classical mechanics based on the assumption of instantaneous propagation of interaction contain intrinsic errors. On the other hand, the large value of this velocity explains the fact that, in practice, classical mechanics appears to be sufficiently accurate in most case. In fact, the velocities with which we have occasion to deal are usually so small compared with c and so

[2]It is important to remember that the other planets perihelion shifts as well, but, since they are farther from the Sun and have longer periods, their shifts are lower, and could not be observed accurately until long after Mercury's.

the assumption that the velocity of light is infinity does not materially affects the accuracy of the results.

The mechanics based on the Einsteinian Principle of Relativity (i.e. the combination of the principle of relativity with the finiteness of the velocities of propagation of interactions) is called *relativistic*. In the limiting case when the velocity of the moving bodies is small compared with c, the relativistic mechanics goes over into the usual (Newtonian or classical) mechanics, based on the assumption of instantaneous propagation of interactions.

The Einsteinian principle of relativity implies a wide range of consequences, which are for example *length contraction*, *time dilatation* and *relativity of simultaneity*, which have been experimentally verified. In particular, it has replaced the conventional notion of an absolute universal time with the notion of a time that is dependent on reference frame and spatial position. For these reasons, each *event* (i.e. a single moment in space and time) is characterized uniquely by a four-vector (ct, x, y, z), which depends on the chosen reference system.

This theory is called "special" because it applies the principle of relativity only to the special case of *inertial reference frames*. Einstein later published a paper on General Relativity in 1915 to apply the principle in the general case, that is to any reference frame, so as to handle general coordinate transformations, and gravitational effects.

2.2 General Relativity

Gravitational fields, or fields of gravity, have the basic property that all bodies move in them in the same manner, independently of mass, provided the initial conditions are the same. This property of gravitational fields provides the possibility of establishing an analogy between the motion of a body in a gravitational field and the motion of a body not located in any external field, but which is considered from the point of view of a non-inertial system of reference. Thus the property of the motion in a non-inertial system are the same as those in an inertial system in the presence of a gravitational field, i.e. a non-inertial reference is equivalent to a certain gravitational field. This is called the *Principle of Equivalence*.

The theory of gravitational fields, constructed on the basis of the theory of relativity, is called *General Theory of Relativity*. The General Theory of Relativity generalizes the special relativity and Newton's law of universal gravitation, providing a unified description of gravity as a geometric property of space and time, or space-time.

One of the basic concepts of General Relativity is that the presence of a mass changes the geometrical properties of the space-time, in the sense that it tends to twist. Vice-versa, all the curvature of space-time indicate the presence of a field whose source is a mass. Gravity becomes so a geometric property of space-time, rather than a mysterious force that propagates with infinite speed between two material objects as in the classical theory.

In General Relativity, the fundamental object of study is the *metric tensor* (or simply, the *metric*). The metric captures indeed all the geometric structure of space-time, and it is used to define notions such as distance, volume, curvature, angle, future and past. Mathematically, the space-time is represented by a 4-dimensional differentiable manifold M and the metric is given as a covariant, second-rank, symmetric tensor on M, conventionally denoted by g. Moreover the metric is required to be non-degenerate with signature $(-, +, +, +)$. A manifold M equipped with such a metric is called a *Lorentzian manifold*. In local coordinates x^i (where i is an index which runs from 0 to 3)[3] the metric can be written in the form

$$g = g_{ik} \mathrm{d}x^i \mathrm{d}x^k. \tag{2.2}$$

With the quantity $\mathrm{d}x^i$ being an infinitesimal coordinate displacement, the metric acts as an infinitesimal invariant interval squared. For this reason one often sees the notation $\mathrm{d}s^2$ for the metric:

$$\mathrm{d}s^2 = g_{ik} \mathrm{d}x^i \mathrm{d}x^k. \tag{2.3}$$

This is the invariant interval, i.e. the measure of separation between two arbitrarily close events in space-time.

The metric (and the associated geometry of space-time) is determined by the matter and energy content of space-time. *Einstein's field equations* relate the metric (and the associated curvature tensors) to the *stress-energy tensor* T_{ik}, which is a tensor that describes the density and the flux of energy and momentum in space-time. The Einstein field equations are a set of 10 equations which describe the gravitational effects produced by a given mass in general relativity. Exact solutions of Einstein's field equations are very difficult to find.

When a body enters inside a gravitational field, represented by a curvature of space-time, it moves along a path as short as possible, which depends by the curvature of space-time and so by the metric tensor and which is called *geodesic line.*

The set of concepts related to General Relativity have been masterfully summarized by the famous phrase of J. A. Wheeler:

"Mass tells space-time how to curve, and space-time tells mass how to move".

Some of the most important experimental verifications of General Relativity are:

- *Perihelion precession of Mercury;*
- *Deflection of light by the Sun;*
- *Gravitational redshift of light.*

2.2.1 Distances and time intervals

In General Relativity, the choice of a coordinate system is not limited in any way: the triple of space coordinate x^1, x^2, x^3, can be any quantities defining the position of the bodies in space and the coordinate x^0 can be defined by an arbitrary running clock. The question that arise is how to determine spatial and time intervals in term of the values of the quantities x^0, x^1, x^2, x^3.

We define the *proper time* for a given object as the time read by a clock moving with this object. It can be proven that the relation between the proper time τ and the coordinate x_0 is given by

$$\mathrm{d}\tau = \frac{1}{c}\sqrt{(g_{00})}\mathrm{d}x^0 \qquad \tau = \frac{1}{c}\int \sqrt{(g_{00})}\mathrm{d}x^0. \tag{2.4}$$

This relation determines the actual time interval for a change of coordinate x^0. We remember that in special relativity, proper time elapses differently for clocks moving relative to one other. In the

[3] The indices go from zero to three: $x^0 \equiv ct, x^1 \equiv x, x^2 \equiv y, x^3 \equiv z$. space-time indices are always in Latin; occasionally we will use Greek indices if we mean only the spatial components, e.g. $\mu = 1, 2, 3$. Moreover, we use the summation convention, i.e. indices which appear both as superscripts and subscripts are summed over.

General Theory of Relativity, proper time elapses differently even at different points of space in the same reference system. This means that the interval of proper time between two events occurring at same point in the space, and the interval of time of two events simultaneous with these at another point is space, are in general different from one another.

It can be proven that the element dl of spatial distance is:

$$\mathrm{d}l^2 = \gamma_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta, \qquad \gamma_{\alpha\beta} = -g_{\alpha\beta} + \frac{g_{0\alpha}g_{0\beta}}{g_{00}}, \tag{2.5}$$

where $\gamma_{\alpha\beta}$ is the three dimensional metric tensor determining the metric, i.e. the geometrical properties of the space. Because the space metric $\gamma_{\alpha\beta}$ generally changes with time (g_{ik} generally depend on x^0), it is meaningless to integrate dl, such an integral would depend on the world line chosen between the two given space points. Thus in General Theory of Relativity, the concept of distance between bodies loses its meaning, remaining valid only for infinitesimally distance or in the case in which g_{ik} don't depend on the time.

2.2.2 Einstein field equations

As already mentioned, the *Einstein Field Equations* are used to determine the space-time geometry resulting from the presence of mass-energy and linear momentum, i.e. they determine the metric tensor of space-time for a given arrangement of stress-energy in the space-time.

The Einstein field equations may be written as

$$R_{ik} - \frac{1}{2}g_{ik}R = \frac{8\pi\mathcal{G}}{c^4}T_{ik} \tag{2.6}$$

where R_{ik} is the Ricci curvature tensor, R the scalar curvature, g_{ik} the metric tensor, $\mathcal{G}$ is Newton's gravitational constant, c the speed of light in vacuum, and T_{ik} the stress-energy tensor (for a brief introduction of these concepts, see appendix A). The solutions of the Einstein field equations are the components of the metric tensor.

We can notice that the expression on the left represents the curvature of space-time as determined by the metric; the expression on the right represents the matter/energy content of space-time.

One can write the Einstein field equations in a more compact form, using the Einstein tensor

$$G_{ik} = \frac{8\pi\mathcal{G}}{c^4}T_{ik}. \tag{2.7}$$

If the energy-momentum tensor T_{ik} is zero in the region under consideration (an empty space), the equation of the gravitational field reduce to the equations $R_{ik} = 0$. Flat Minkowski space is the simplest example of solution of these equations; in this case the metric is simply

$$\eta = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2.8}$$

It is important to note that the equations of gravitational field are non linear equations. Therefore for gravitational fields the principle of superposition is not valid. However, in the approximation of weak gravitational fields, for which the equations of the fields in first approximation are linear, the principle of superposition is valid.

Finally, it can be proven that the Einstein field equations reduce to Newton's law of gravitation when the gravitational field is weak and velocities are much less than the speed of light.

2.2.3 Motion of a particle in a gravitational field

The motion of a particle in a gravitational field is determined by the principle of least action. The *Principle of least action* is defined by the statement that for each mechanical system there exist a certain integral S, called action, which has a stationary value for the actual motion, so that its variation δS is zero.

The *action* between two particular events a and b is defined as

$$S = -mc \int_a^b \mathrm{d}s = -mc \int_a^b \sqrt{-g_{\alpha\beta} \frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}} \mathrm{d}\lambda \tag{2.9}$$

where the integral is along the world line of the particle and λ is an invariant parameter which described the curve. In particular, for material particles the parameter λ is usually chosen to be the *proper time* τ or, equivalently, the space-time interval s ($\mathrm{d}s = c\mathrm{d}\tau$).

Using the principle of least action, it can be proven that in a gravitational field the particle moves along a *geodesic line* in the four space x^0, x^1, x^2, x^3. It can be shown that the geodesic equation is

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}s^2} + \Gamma^i{}_{kl} \frac{\mathrm{d}x^k}{\mathrm{d}s} \frac{\mathrm{d}x^l}{\mathrm{d}s} = 0, \qquad 0 \leq i \leq 3, \tag{2.10}$$

where $\Gamma^i{}_{kl}$ are the Christoffel symbols, which depend on the metric g_{ik}. This explains the second part of the phrase of Wheeler, i.e. the fact that the motion of a body depends on the distribution of matter and so on the curvature of space-time.

It is important to remember that the equation of geodesic in the form (2.10) is not applicable to the propagation of a light signal, since along the world line of the propagation of a light ray the interval ds is zero, so that all terms in equation (2.10) become infinity.

In the limiting case of small velocities and weak gravitational field, the relativistic equations of motion of a particle in a gravitational field go into the corresponding non-relativistic equation.

The action integral can also be represented as an integral of a function L

$$S = \int_a^b L \mathrm{d}\lambda, \tag{2.11}$$

where L represents the *Lagrangian function* of the mechanical system. In Lagrangian mechanics, the equations of motion for a material particle (2.10) can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}s} \frac{\partial L}{\partial(\mathrm{d}x^i/\mathrm{d}s)} - \frac{\partial L}{\partial x^i} = 0, \qquad 0 \leq i \leq 3 \tag{2.12}$$

which are called the *Euler-Lagrange equations.*

Because the action is invariant under re-parameterizations, we can choose t as the parameter λ in (2.9)-(2.11) and we obtain the following definition of the Lagrangian

$$L = -mc\sqrt{-g_{\alpha\beta} u^\alpha u^\beta} \tag{2.13}$$

where $u = \mathrm{d}x^i/\mathrm{d}t = (c, \mathbf{v})$ is the four-vector velocity of the test mass. The corresponding equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}^\alpha} - \frac{\partial L}{\partial x^\alpha} = 0, \qquad 1 \leq \alpha \leq 3. \tag{2.14}$$

2.2.4 Newton's law

We carry now the transition of the Einstein equation (2.6) to the non-relativistic limit (i.e. weak field and small velocities, compared with light velocity, of all bodies).

If the gravitational field is weak enough, then space-time will be only slightly deformed from the gravity-free Minkowski space (2.8), and we can consider the space-time metric as a small perturbation of the Minkowski metric:

$$g_{ij} = \eta_{ij} + h_{ij}, \qquad h << 1. \tag{2.15}$$

Using the fact that in non-relativistic mechanics the motion of a particle in a gravitational field is determined by the Lagrangian $L = -mc^2 + \frac{1}{2}mv^2 - m\phi$ (ϕ is the non-relativistic potential of the gravitational field), it can be proven that, in the limiting case which are considering, the expression for the component g_{00} of the metric tensor is $g_{00} = 1 + \frac{2\phi}{c^2}$. In the non-relativistic limit, the stress-energy tensor is given by

$$T^0{}_0 = \mu c^2, \qquad T^\mu{}_0 = T^0{}_\mu = 0, \qquad T^\mu{}_\nu = 0, \qquad 1 \leq \mu, \nu \leq 3 \tag{2.16}$$

Indeed, the stress-energy tensor is given by $T^k{}_i = \mu c^2 u^k u_i$ and in the non-relativistic limit we must set $u^\alpha = 0$ and $u^0 = u_0 = 1$. Of course the scalar $T = T^i{}_i$ will be equal to μc^2.

Using the field equation, we obtain that $R^0{}_0 = \frac{4\pi\mathcal{G}}{c^4}\mu$ and that all other equations are vanishing. Because $\Gamma^\alpha{}_{00} \simeq -\frac{1}{2}\eta^{\alpha\beta}\partial_\beta h_{00}$ and $R^0{}_0 = \partial_\alpha \Gamma^\alpha{}_{00}$, we obtain $R^0{}_0 = \frac{1}{c^2}\Delta\phi$. Thus the Einstein field equations give

$$\Delta\phi = 4\pi\mathcal{G}\mu. \tag{2.17}$$

This is the equation of gravitational field in non-relativistic mechanics. In particular we have for potential of the field of a single particle of mass m

$$\phi = -\frac{\mathcal{G}m}{R} \tag{2.18}$$

and, consequently, the force $F = -m'\frac{\partial\phi}{\partial R'}$ acting in this field on another particle of mass m' is

$$F = -\frac{\mathcal{G}mm'}{R^2}, \tag{2.19}$$

which is the well known law of attraction of Newton.

Now we show that the geodesic equation reduces to Newton's second law. Remember the geodesic equation (2.10), using proper time as the parameter of the worldline, every term containing one or two spatial four-velocity components will be dwarfed by the term containing two time components, because the particle in question is moving slowly. We can therefore take the approximation

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu{}_{00}\frac{\mathrm{d}x^0}{\mathrm{d}\tau}\frac{\mathrm{d}x^0}{\mathrm{d}\tau} = 0. \tag{2.20}$$

Using the definition of g_{00} and the fact that $\tau \approx t$ (in this limit), we obtain

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}t^2} = -\partial_\mu\phi \tag{2.21}$$

which is Newton's second law of motion.

2.3 The gravitational field equations

A system of bodies in motion loses some of its energy by radiating gravitational waves. However, this loss of energy is revealed only in the approximation of the fifth order of $1/c$, i.e. up to the fourth approximation the energy of the system remains constant. This means that a system of gravitating bodies can be described by a Lagrangian correctly less than fifth order terms. The aim of these sections is to derive the Lagrangian of a system of bodies to terms of second order, i.e. in the next approximation after the Newtonian. To do this, we shall neglect the dimensions and the internal structure of the bodies, regarding them as "point-like". Moreover, we do not consider the rotation of these bodies around their axes. To derive the required Lagrangian, we follow the works of *Einstein, Infeld and Hoffmann (1938)*, *Landau and Lifshitz (1971)* and *Brumberg (2008)*.

As first thing, we must determine the gravitation field produced by any centrally symmetric non-rotating distribution of matter and the weak gravitational field produced by the bodies at large distances compared to their dimensions, but, at the same time, at small distances compared to the wavelength of the gravitational waves radiated by the system.

2.3.1 Schwarzschild metric

Let us consider a static gravitational field possessing central symmetry. Such a field can be produced by any centrally symmetric non-rotating distribution of matter. The central symmetry of the field means that the space-time metric, that is the expression for the interval ds, must be the same for all points located at the same distance from the center. It is important to note that, while in a euclidean space this distance is equal to the radius vector, in a non-euclidean space, such as we have in the presence of a gravitational field, there is no quantity which has all the property of the euclidean radius vector.

If we use "spherical" space coordinates r, θ, ϕ, it can been proven that the metric is

$$\mathrm{d}s^2 = \left(1 - \frac{r_g}{r}\right) c^2\mathrm{d}t^2 - r^2(\sin^2\theta \mathrm{d}\phi^2 + \mathrm{d}\theta^2) - \frac{\mathrm{d}r^2}{1 - \frac{r_g}{r}}, \tag{2.22}$$

where M is the total mass of the bodies producing the field and $r_g = \frac{2\mathcal{G}M}{c^2}$ is called gravitational radius of the body. The metric can be written in matrix form:

$$g_{ik} = \begin{bmatrix} 1 - \frac{r_g}{r} & 0 & 0 & 0 \\ 0 & -\left(1 - \frac{r_g}{r}\right)^{-1} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{bmatrix}. \tag{2.23}$$

An interesting thing that can be noticed is that at finite distance from the masses there is a "slowing down" of the time compared with the time at infinity. In fact, combining $g_{00} \leq 1$ with the formula (2.4) $\mathrm{d}\tau = \sqrt{g_{00}}\mathrm{d}t$ defining the proper time, it follows that $\mathrm{d}\tau \leq \mathrm{d}t$ and the equality sign holds only at infinity, where t coincides with the proper time.

The Schwarzschild solution is a useful approximation for describing slowly rotating astronomical objects such as many stars, planets and black holes, including Earth and the Sun. In particular, the Schwarzschild geodesics are a good approximation to the relative motion of two bodies of arbitrary mass, provided that the Schwarzschild mass M is set equal to the sum of the two individual masses m_1 and m_2.

An approximate expression for $\mathrm{d}s^2$ at large distances from the origin of coordinates is

$$\mathrm{d}s^2 = \mathrm{d}s_0^2 - \frac{2GM}{c^2 r}(\mathrm{d}r^2 + c^2\mathrm{d}t^2). \tag{2.24}$$

The second term represents a small correction to the Galilean metric $\mathrm{d}s_0^2$. At large distances from the masses producing it, every field appears centrally symmetric. Therefore (2.24) determines the metric at large distances from any system of bodies.

Motion in a centrally symmetric gravitational field

Let us consider the motion of a test body in a centrally symmetric gravitational field. As in every centrally symmetric field, the motion occurs in a single plane passing through the origin (e.g. $\theta = \pi/2$).

The geodesic equation for θ is

$$0 = \frac{\mathrm{d}^2\theta}{\mathrm{d}s^2} + \frac{2}{r}\frac{\mathrm{d}\theta}{\mathrm{d}s}\frac{\mathrm{d}r}{\mathrm{d}s} - \sin\theta\cos\theta\left(\frac{\mathrm{d}\phi}{\mathrm{d}s}\right)^2 \tag{2.25}$$

so $\theta = \pi/2$ is a valid solution. Fixed $\theta = \pi/2$ ($\mathrm{d}\theta/\mathrm{d}t = 0$), the geodesic equations for ϕ and t become

$$0 = \frac{\mathrm{d}^2\phi}{\mathrm{d}s^2} + \frac{2}{r}\frac{\mathrm{d}\phi}{\mathrm{d}s}\frac{\mathrm{d}r}{\mathrm{d}s} \qquad 0 = \frac{\mathrm{d}^2 t}{\mathrm{d}s^2} + \frac{1}{w}\frac{\mathrm{d}w}{\mathrm{d}r}\frac{\mathrm{d}t}{\mathrm{d}s}\frac{\mathrm{d}r}{\mathrm{d}s} \tag{2.26}$$

where $w(r) = 1 - \frac{r_g}{r}$ and $v(r) = 1/w(r)$. Solving (2.26), we obtain two constants of the motion:

$$L = cr^2\frac{\mathrm{d}\phi}{\mathrm{d}s} \qquad E = c^3\left(1 - \frac{r_g}{r}\right)\frac{\mathrm{d}t}{\mathrm{d}s}. \tag{2.27}$$

If we rewrite the metric as

$$c^2 = \left(1 - \frac{r_g}{r}\right)c^2\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \frac{1}{1 - \frac{r_g}{r}}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 - r^2\left(\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\right)^2, \tag{2.28}$$

remembering that $\mathrm{d}s^2 = c^2\mathrm{d}\tau^2$ and $\theta = \pi/2$, we obtain

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 = \frac{E^2 r^4}{L^2 c^2} - \left(1 - \frac{r_g}{r}\right)\left(\frac{c^2 r^4}{L^2} + r^2\right). \tag{2.29}$$

Using the inverse radius $u = 1/r$, the orbital equation (2.29) can be rewritten in the form

$$\left(\frac{\mathrm{d}u}{\mathrm{d}\phi}\right)^2 = \frac{E^2}{c^2 L^2} - \frac{c^2}{L^2} + \frac{c^2}{L^2}r_g u - u^2 + r_g u^3. \tag{2.30}$$

Deriving both sides of (2.30) respect to ϕ and dividing for $2\mathrm{d}u/\mathrm{d}\phi$, we obtain

$$\frac{\mathrm{d}^2 u}{\mathrm{d}\phi^2} + u = \frac{\mathcal{G}M}{L^2} + \frac{3\mathcal{G}M}{c^2}u^2, \tag{2.31}$$

where we have used $r_g = \frac{2\mathcal{G}M}{c^2}$.

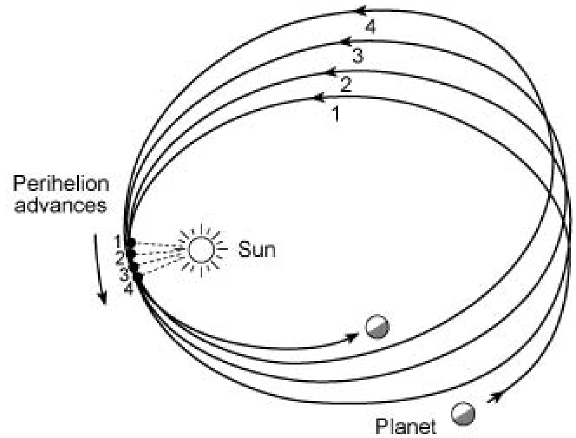


Figure 2.1: Relativistic shift in the perihelion of the orbit. Courtesy of Kenneth R. Lang, from his book *The Cambridge Guide to the Solar System: Second Edition*, Cambridge University Press, 2011, p. 217.

The term $\frac{3\mathcal{G}M}{c^2}u^2$ is absent in the Newtonian theory (see (1.8)), where we have

$$\begin{aligned} \frac{\mathrm{d}^2 u_0}{\mathrm{d}\phi^2} + u_0 &= \frac{\mathcal{G}M}{L^2}, \\ u_0 &= \frac{\mathcal{G}M}{L^2}(1 + e\cos\phi). \end{aligned} \tag{2.32}$$

This term is important because it leads to a systematic (secular) shift in the perihelion of the orbit.

To solve (2.31), we look for a solution of the form

$$u(\phi, \epsilon) = u_0(\phi) + \epsilon u_1(\phi) + \epsilon^2 u_2(\phi) + ..., \tag{2.33}$$

where $\epsilon = \frac{1}{c^2}$, ϕ is not periodic and u_0 is given by (2.32). Replacing u in (2.31) and collecting together all coefficients of the same power of ε, we get the infinite system of equations of the form

$$\frac{\mathrm{d}^2 u_k}{\mathrm{d}\phi^2} + u_k = 3\mathcal{G}M u_{k-1}^2(\phi) \tag{2.34}$$

for $k \geq 1$. We can notice that the r.h.s. is a known function, because it depends only on $u_{k-1}(\phi)$ which is known or determined by the equations of preceding steps. Hence we may attempt a recursive solution of the system.

In particular, the equation for u_1 is:

$$\frac{\mathrm{d}^2 u_1}{\mathrm{d}\phi^2} + u_1 = 3\mathcal{G}M u_0^2(\phi) = \frac{3\mathcal{G}^3 M^3}{L^4}(1 + 2e\cos\phi + e^2\cos^2\phi). \tag{2.35}$$

It can be proven that the solution of this equation is

$$u_1 = \frac{3\mathcal{G}^3 M^3}{L^4}\left\{1 + e\phi\sin(\phi) + e^2\left[\frac{1}{2} - \frac{1}{6}\cos(2\phi)\right]\right\}. \tag{2.36}$$

The solution of (2.31) is so

$$u = \frac{\mathcal{G}M}{L^2}\left\{1 + e\cos\phi + \frac{3\mathcal{G}^2M^2}{c^2L^2}\left[1 + e\phi\sin(\phi) + e^2\left(\frac{1}{2} - \frac{1}{6}\cos(2\phi)\right)\right]\right\} + O\left(\frac{1}{c^4}\right), \tag{2.37}$$

which can be simplified in the following way

$$u \simeq \frac{\mathcal{G}M}{L^2}\left[1 + e\cos\left(\left(1 - \frac{3\mathcal{G}^2M^2}{c^2L^2}\right)\phi\right)\right]. \tag{2.38}$$

The approximated solution is yet periodic, but with period equal to

$$T \simeq \frac{2\pi}{1 - \frac{3\mathcal{G}^2M^2}{c^2L^2}} \simeq 2\pi\left(1 + \frac{3\mathcal{G}^2M^2}{c^2L^2}\right). \tag{2.39}$$

Concluding, we have found that

$$\Delta T \simeq \frac{6\pi\mathcal{G}^2M^2}{c^2L^2} \tag{2.40}$$

which is the required angular displacement of the Newtonian ellipse during one revolution, i.e. the shift in the perihelion of the orbit. Expressing it in term of the length a of the semi-major axis and the eccentricity e of the ellipse ($a = L^2/(\mathcal{G}M(1-e^2))$), we obtain

$$\Delta T \simeq \frac{6\pi\mathcal{G}M}{c^2a\,(1-e^2)}. \tag{2.41}$$

Numerical values of the shifts determined from the formula (2.41) for Mercury and Earth are equal, respectively, to 43.0" and 3.8" per century. Astronomical measurements give 43.1" $\pm$ 0.4" and 5.0" $\pm$ 1.2", in excellent agreement with theory.

2.3.2 Gravitational fields at large distances from bodies

Using the considerations made in the previous section, it's easy to derive the stationary gravitational field at large distance r from the body (placed in the origin of the coordinates) which produces it, and determine the first term of its expansion in power of $1/r$. Because away from the body the field is weak, the space-time metric is almost Galilean, i.e

$$g_{ij} = \eta_{ij} + h_{ij}, \qquad h << 1. \tag{2.42}$$

where η_{ij} is given in (2.8) and where h_{ij} are small corrections that determine the gravitational field. Less than zero of third-order, the determinant of the metric tensor is

$$g = \eta\left(1 + h + \frac{1}{2}h^2 - \frac{1}{2}h^i{}_k h^k{}_i\right). \tag{2.43}$$

In the first approximation, to term of order $1/r$, the small corrections h_{ik} are given by the corresponding terms in the expansion of the Schwarzschild solution (2.24), i.e.

$$h_{00} = -\frac{r_g}{r} \qquad h_{\alpha\beta} = -\frac{r_g}{r}\delta_{\alpha\beta} \qquad h_{0\alpha} = 0 \tag{2.44}$$

where $r_g = 2\mathcal{G}M/c^2$.

It is important to remember that the condition for which h_{ik} must be infinitesimal quantities, does not determine a unique choice of the reference system. In fact, if this condition is met in some system, it will be satisfied by any transformation $x'^i = x^i + \xi^i$, where ξ^i are infinitesimal quantities. In this case, it is simple to prove that the new metric g'_{ik} is given by

$$g'_{ik} = \eta_{ij} + h'_{ij} \tag{2.45}$$

where

$$h'_{ij} = h_{ik} - \frac{\partial \xi_i}{\partial x^k} - \frac{\partial \xi_k}{\partial x^i}. \tag{2.46}$$

2.4 The equation of motion of a system of bodies in the second approximation

The field h, found in the previous section, is the field at large distance from a “point-like” non-rotating body placed in the origin of coordinate. Since it is a solution of the linearized Einstein equations, for it the principle of superposition applies. Consequently, a field at large distances from a system of bodies is calculated by summing the fields of each of these bodies:

$$h^0{}_0 = \frac{2}{c^2}\phi, \qquad h^\alpha{}_0 = 0, \qquad h^\alpha{}_\beta = -\frac{2}{c^2}\phi\delta^\alpha{}_\beta, \tag{2.47}$$

where

$$\phi(\mathbf{r}) = -\mathcal{G}\sum_a \frac{m_a}{\parallel \mathbf{r} - \mathbf{r}_a \parallel} \tag{2.48}$$

is the Newtonian gravitational potential of a system of point-like bodies ($\mathbf{r}_a$ is the radius vector of the body of mass m_a). The expression of the interval of the metric tensor is

$$\mathrm{d}s^2 = \left(1 + \frac{2}{c^2}\phi\right) c^2 \mathrm{d}t^2 - \left(1 - \frac{2}{c^2}\phi\right)(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2). \tag{2.49}$$

As will be seen from the sequel, to obtain the required equations of motion it’s sufficient to know the spatial components $h_{\alpha\beta}$ to the accuracy $\sim 1/c^2$, with which are given in (2.44); the mixed components (which are absent in the $1/c^2$ approximation) are needed up to terms of order $1/c^3$, and the time component h_{00} to terms in $1/c^4$. To calculate them we turn once again to the general equations of gravitation, and consider the terms of corresponding order in these equations.

Disregarding the fact that the bodies are macroscopic, we must write the energy-momentum tensor of of the system of non-interacting particles. In curvilinear coordinates, it can be proven that it is

$$T^{ik} = \sum_a \frac{m_a c}{\sqrt{-g}} \frac{\mathrm{d}x^i}{\mathrm{d}s} \frac{\mathrm{d}x^k}{\mathrm{d}t} \delta(\mathbf{r} - \mathbf{r}_a), \tag{2.50}$$

where the summation extends over all the bodies in the system, $\mathbf{r}_a$ is the radius-vector of the particles, m_a is the mass of the particle a and δ is the Dirac delta function. The component

$$T_{00} = \sum_a \frac{m_a c^3}{\sqrt{-g}} g_{00}^2 \frac{\mathrm{d}t}{\mathrm{d}s} \delta(\mathbf{r} - \mathbf{r}_a) \tag{2.51}$$

is equal to $\sum_a m_a c^2 \delta(\mathbf{r}-\mathbf{r}_a)$ in first approximation (i.e. for Galilean g_{ik}); in the next approximation, we substitute for g_{ik} from (2.49) and find, after a simple computation:

$$T_{00} = \sum_a m_a c^2 \left(1 + \frac{5\phi_a}{c^2} + \frac{v_a^2}{2c^2}\right) \delta(\mathbf{r}-\mathbf{r}_a), \tag{2.52}$$

where v_a is the ordinary three-dimensional velocity ($v^\alpha = \mathrm{d}x^\alpha/\mathrm{d}t$) and ϕ_a is the potential of the field at the point $\mathbf{r}_a$ (it is simple to note that $\phi(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}_a) = \phi_a\delta(\mathbf{r}-\mathbf{r}_a)$)[4].

As regards the components $T_{\alpha\beta}$, $T_{0\alpha}$ of the energy momentum tensor, in this approximation it's sufficient to keep for them only the first terms in the expansion of the expression (2.50)

$$T_{\alpha\beta} = \sum_a m_a v_{a\alpha} v_{a\beta} \delta(\mathbf{r}-\mathbf{r}_a), \qquad T_{0\alpha} = -\sum_a m_a c v_{a\alpha} \delta(\mathbf{r}-\mathbf{r}_a). \tag{2.53}$$

Next we compute the components of the Ricci tensor R_{ik}, using the formula $R_{ik} = g^{lm} R_{limk}$ with R_{limk} given by $(A.12)$. A simple computation gives for R_{00} the result

$$\begin{aligned} R_{00} =& \frac{1}{c}\frac{\partial}{\partial t}\left(\frac{\partial h^\alpha{}_0}{\partial x^\alpha} - \frac{1}{2c}\frac{\partial h^\alpha{}_\alpha}{\partial t}\right) + \frac{1}{2}\Delta h_{00} + \frac{1}{2}h^{\alpha\beta}\frac{\partial^2 h_{00}}{\partial x^\alpha \partial x^\beta} + \\ & - \frac{1}{4}\left(\frac{\partial h_{00}}{\partial x^\alpha}\right)^2 - \frac{1}{4}\frac{\partial h_{00}}{\partial x^\beta}\left(2\frac{\partial h^\alpha{}_\beta}{\partial x^\alpha} - \frac{\partial h^\alpha{}_\alpha}{\partial x^\beta}\right). \end{aligned} \tag{2.54}$$

In this computation we have still not used any auxiliary condition for the quantities h_{ik}. Making use of this freedom (see (2.45) – (2.46)), we now impose the condition

$$\frac{\partial h^\alpha{}_0}{\partial x^\alpha} - \frac{1}{2c}\frac{\partial h^\alpha{}_\alpha}{\partial t} = 0 \tag{2.55}$$

as a result of which all term containing the components h drop out of R_{00}. In the remaining terms, we substitute

$$h^0{}_0 = \frac{2}{c^2}\phi + O\left(\frac{1}{c^4}\right), \qquad h^\beta{}_\beta = -\frac{2}{c^2}\phi\delta^\beta{}_\beta \tag{2.56}$$

and obtain, to the required accuracy,

$$R_{00} = \frac{1}{2}\Delta h_{00} + \frac{2}{c^4}\phi\Delta\phi - \frac{2}{c^4}(\nabla\phi)^2, \tag{2.57}$$

where we have gone over to three-dimensional notation.

In a similar way, we find that the components $R_{0\alpha}$ is

$$R_{0\alpha} = \frac{1}{2c}\frac{\partial^2 h^\beta{}_\alpha}{\partial t \partial x^\beta} + \frac{1}{2}\frac{\partial^2 h^\beta{}_0}{\partial x^\alpha \partial x^\beta} - \frac{1}{2c}\frac{\partial^2 h^\beta{}_\beta}{\partial t \partial x^\alpha} + \frac{1}{2}\Delta h_{0\alpha} \tag{2.58}$$

and then, using the condition (2.55):

$$R_{0\alpha} = \frac{1}{2}\Delta h_{0\alpha} + \frac{1}{2c^3}\frac{\partial^2 \phi}{\partial t \partial x^\alpha}. \tag{2.59}$$

[4] As yet, we pay no attention to the fact that ϕ_a contains an infinite part, i.e. the potential of the self-field of the particle m_a: concerning this, see below.

Using the expressions from (2.52) to (2.59), we now write the Einstein equations

$$R_{ik} = \frac{8\pi\mathcal{G}}{c^4}\left(T_{ik} - \frac{1}{2}g_{ik}T\right). \tag{2.60}$$

The time component of equation (2.60) gives

$$\Delta h_{00} + \frac{4}{c^4}\phi\Delta\phi - \frac{4}{c^4}(\nabla\phi)^2 = \frac{8\pi\mathcal{G}}{c^4}\sum_a m_a c^2\left(1 + \frac{5\phi_a}{c^2} + \frac{3v_a^2}{2c^2}\right)\delta(\mathbf{r} - \mathbf{r}_a), \tag{2.61}$$

which can be rewritten in the form

$$\Delta\left(h_{00} - \frac{2}{c^4}\phi^2\right) = \frac{8\pi\mathcal{G}}{c^4}\sum_a m_a c^2\left(1 + \frac{5\phi'_a}{c^2} + \frac{3v_a^2}{2c^2}\right)\delta(\mathbf{r} - \mathbf{r}_a), \tag{2.62}$$

using the identity

$$4(\nabla\phi)^2 = 2\Delta(\phi^2) - 4\phi\Delta\phi$$

and the equation of the Newtonian potential

$$\Delta\phi = 4\pi\mathcal{G}\sum_a m_a\delta(\mathbf{r} - \mathbf{r}_a).$$

In the right side of equation (2.62), after completing all the computations, we have replaced ϕ_a by

$$\phi'_a = -\mathcal{G}\sum_{b\neq a}\frac{m_b}{\parallel \mathbf{r}_a - \mathbf{r}_b \parallel}, \tag{2.63}$$

i.e. by the potential at the point $\mathbf{r}_a$ of the field produced by all particles except for the particle m_a[5].

Using the relation

$$\Delta\frac{1}{\parallel \mathbf{r} \parallel} = -4\pi\delta(\mathbf{r}),$$

it can been proven that the solution of (2.62) is

$$h_{00} = \frac{2\phi}{c^2} + \frac{2\phi^2}{c^4} - \frac{2\mathcal{G}}{c^4}\sum_a\frac{m_a\phi'_a}{\parallel \mathbf{r} - \mathbf{r}_a \parallel} - \frac{3\mathcal{G}}{c^4}\sum_a\frac{m_a v_a^2}{\parallel \mathbf{r} - \mathbf{r}_a \parallel}. \tag{2.64}$$

The mixed component of equation (2.60) gives

$$\Delta h_{0\alpha} = -\frac{16\pi\mathcal{G}}{c^3}\sum_a m_a v_{a\alpha}\delta(\mathbf{r} - \mathbf{r}_a) - \frac{1}{c^3}\frac{\partial^2\phi}{\partial t\partial x^\alpha}. \tag{2.65}$$

The solution of this linear equation is

$$h_{0\alpha} = \frac{4\mathcal{G}}{c^3}\sum_a\frac{m_a v_{a\alpha}}{\parallel \mathbf{r} - \mathbf{r}_a \parallel} - \frac{1}{c^3}\frac{\partial^2 f}{\partial t\partial x^\alpha}, \tag{2.66}$$

[5] The exclusion of the infinite self potential of the (point-like) particles corresponds to a "renormalization" of their masses, as a result of which they take on their true values, which take into account the field produced by the particles themselves.

where f is the solution of the auxiliary equation

$$\Delta f = \phi = -\sum \frac{\mathcal{G} m_a}{\| \mathbf{r} - \mathbf{r}_a \|}. \tag{2.67}$$

Using the relation $\Delta r = 2/r$, we find

$$f = -\frac{\mathcal{G}}{2} \sum_a m_a \| \mathbf{r} - \mathbf{r}_a \|, \tag{2.68}$$

and then, after a simple computation, we finally obtain:

$$h_{0\alpha} = \frac{\mathcal{G}}{2c^3} \sum_a \frac{m_a}{\| \mathbf{r} - \mathbf{r}_a \|} [7 v_{a\alpha} + (\mathbf{v}_a \cdot \mathbf{n}_a) n_{a\alpha}] \tag{2.69}$$

where $\mathbf{n}_a$ is a unit vector along the direction of the vector $\mathbf{r} - \mathbf{r}_a$.

Using the expressions (2.47) – (2.64) – (2.69), the required Lagrangian for a single particle to terms of second order, in a gravitational field produced by other particles and assumed to be given, is

$$L_a = -m_a c \frac{\mathrm{d}s}{\mathrm{d}t} = -m_a c^2 \left(1 + h_{00} + 2h_{0\alpha} \frac{v_a^\alpha}{c} - \frac{v_a^2}{c^2} + h_{\alpha\beta} \frac{v_a^\alpha v_a^\beta}{c^2} \right)^{1/2}, \tag{2.70}$$

where $v_a^2 = \mathbf{v}_a \cdot \mathbf{v}_a$.

Expanding the square root and dropping the unrelevant constant $-m_a c^2$, we rewrite this expression, to the required accuracy, as

$$L_a = \frac{m_a v_a^2}{2} + \frac{m_a v_a^4}{8c^2} - m_a c^2 \left(\frac{h_{00}}{2} + h_{0\alpha} \frac{v_a^\alpha}{c} + \frac{1}{2c^2} + h_{\alpha\beta} v_a^\alpha v_a^\beta - \frac{h_{00}^2}{8} + \frac{h_{00}}{4c^2} v_a^2 \right). \tag{2.71}$$

The total Lagrangian of the system is, of course, not equal to the sum of the Lagrangians L_a for the individual bodies, but must be constructed so that it leads to the correct values of the force $\mathbf{f}_a$ acting on each of the bodies for a given motion of the others. For this purpose we compute the forces $\mathbf{f}_a$ by differentiating the Lagrangian L_a:

$$\mathbf{f}_a = \left(\frac{\partial L_a}{\partial \mathbf{r}} \right)_{\mathbf{r} = \mathbf{r}_a}. \tag{2.72}$$

It is then easy to form the total Lagrangian L, from which all of the forces $\mathbf{f}_a$ are obtained by taking the partial derivatives $\partial L / \partial \mathbf{r}_a$. The final results for Lagrangian is:

$$\begin{aligned} L = & \sum_a \frac{m_a v_a^2}{2} + \sum_a \sum_{b \neq a} \frac{\mathcal{G} m_a m_b}{r_{ab}} + \\ & + \sum_a \frac{m_a v_a^4}{8c^2} + \sum_a \sum_{b \neq a} \frac{3 \mathcal{G} m_a m_b v_a^2}{2c^2 r_{ab}} - \\ & - \sum_a \sum_{b \neq a} \frac{\mathcal{G} m_a m_b}{4c^2 r_{ab}} [7(\mathbf{v}_a \cdot \mathbf{v}_b) + (\mathbf{v}_a \cdot \mathbf{n}_{ab})(\mathbf{v}_b \cdot \mathbf{n}_{ab})] - \\ & - \sum_a \sum_{b \neq a} \sum_{c \neq a} \frac{\mathcal{G}^2 m_a m_b m_c}{2c^2 r_{ab} r_{ac}}, \end{aligned} \tag{2.73}$$

where $r_{ab} = \| \mathbf{r}_a - \mathbf{r}_b \|$ and $\mathbf{n}_{ab}$ is a unit vector along the direction $\mathbf{r}_a - \mathbf{r}_b$.

The equation of motion corresponding to this Lagrangian were first obtained by A. Einstein, L. Infeld and B. Hoffmann in 1938. We can notice that, in the limit $c \to \infty$, one recovers the Newton's Lagrangian of N bodies.

The relativistic center of inertia

Finally, we want to find the coordinates of the center of inertia of a system of gravitating bodies in the second approximation.

The *relativistic center of inertia* $\mathbf{R}$ of the system is given by the formula

$$\mathbf{R} = \frac{\sum_a E_a \mathbf{r}_a + \int W \mathbf{r} \mathrm{d}V}{\sum_a E_a + \int W \mathrm{d}V}, \tag{2.74}$$

where E_a is the kinetic energy of the particle (including its rest energy) and W is the energy density of the gravitational field. Since the E_a contains the large quantities $m_a c^2$, it is sufficient to consider only those terms of E_a and of W which do not contain c, i.e. we consider only the non relativistic kinetic energy of the particles and the energy of gravitational field. It can be proven that the energy density of the gravitational field in the Newtonian theory is $W = -\frac{1}{8\pi\mathcal{G}}(\nabla\phi)^2$.

The coordinates of the center of inertia are so given by the formula

$$\begin{aligned} \mathbf{R} &= \frac{1}{E} \sum_a \mathbf{r}_a \left(m_a c^2 + \frac{p_a^2}{2m_a} - \frac{\mathcal{G} m_a}{2} \sum_{b \neq a} \frac{m_b}{r_{ab}} \right), \\ E &= \sum_a \left(m_a c^2 + \frac{p_a^2}{2m_a} - \frac{\mathcal{G} m_a}{2} \sum_{b \neq a} \frac{m_b}{r_{ab}} \right), \end{aligned} \tag{2.75}$$

where E is the total energy of the system.

It is easy to show that E is conserved, i.e. $\dot{E} = O(c^{-4})$, and that the barycentre's acceleration vanishes, i.e. $\ddot{\mathbf{R}} = O(c^{-4})$. Therefore the barycenter moves uniformly with constant velocity, and $\mathbf{R}$ can be set equal to zero by placing the origin of the coordinate system at the relativistic center of mass.

As expected, if the velocities of all particles are small compared to c and the gravitational field is weak, we can approximately set $E \approx mc^2$ (where E is the total energy $E = \sum_a E_a + \int W \mathrm{d}V$) so that (2.75) goes over into the usual classical expression

$$\mathbf{R} = \frac{\sum_a m_a \mathbf{r}_a}{\sum_a m_a}. \tag{2.76}$$

2.4.1 Relativistic Hamiltonian of N bodies

Because the Lagrange equations are difficult to handle, we rewrite them in a Hamiltonian form before to study them in detail.

Consider a n-dimensional differentiable manifold (the *configuration space*) endowed with (local) coordinates $q_1, q_2, ..., q_n$, and its tangent space described by the generalized components $\dot{q}_1, \dot{q}_2, ..., \dot{q}_n$ of the velocity. The dynamical state of the system at a given time t is completely determined by the knowledge of $\mathbf{q}(t), \dot{\mathbf{q}}(t)$. In particular, the dynamics is determined by the *Lagrangian function* $L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)$ through the n differential equations of second order

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0, \qquad 0 \leq j \leq n. \tag{2.77}$$

To rewrite the Lagrange equations in a Hamiltonian form, we first introduce the momenta $p_1, ..., p_n$ conjugated to $q_1, ..., q_n$ defined as

$$p_j = \frac{\partial L}{\partial \dot{q}_j}(\mathbf{q}, \dot{\mathbf{q}}, t), \qquad 1 \leq j \leq N. \tag{2.78}$$

which are given as function of $\mathbf{q}, \dot{\mathbf{q}}, t$. If the condition $\det\left(\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j}\right) \neq 0$ is fulfilled, then (2.78) can be solved with respect to $\dot{q}_1, ..., \dot{q}_n$, thus giving $\dot{\mathbf{q}}_i = \dot{\mathbf{q}}_i(\mathbf{q}, \mathbf{p}, t)$, and the momenta can be used in place of the velocities $\dot{\mathbf{q}}(t)$ in order to determine the dynamical state.

The Hamiltonian function of the system is defined as the *Legendre transform* of the Lagrangian:

$$H(\mathbf{q}, \mathbf{p}, t) = \left[\sum_{j=1}^{n} p_j(\mathbf{q}, \dot{\mathbf{q}}, t)\dot{q}_j - L(\mathbf{q}, \dot{\mathbf{q}}, t)\right]_{\dot{\mathbf{q}} = \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t)}, \tag{2.79}$$

where $\dot{\mathbf{q}}$ must be replaced everywhere with its expression as a function of $\mathbf{q}, \mathbf{p}, t$. It is simple to prove that *Hamilton's equations*

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \tag{2.80}$$

correspond to the Euler-Lagrange equations (2.77).

In the case of relativistic Lagrangian (2.73), the momenta conjugates to the coordinates are

$$\mathbf{p}_i = m_i \mathbf{v}_i + \frac{1}{c^2}\left\{\frac{1}{2} m_i v_i^2 \mathbf{v}_i + \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{3\mathcal{G} m_i m_j}{r_{ij}} \mathbf{v}_i - \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{\mathcal{G} m_i m_j}{2 r_{ij}} [7\mathbf{v}_j + (\mathbf{v}_i \cdot \mathbf{n}_{ij})\mathbf{n}_{ij}]\right\} \tag{2.81}$$

where $\mathbf{v}_i = \dot{\mathbf{q}}_i$. It can be proven that, to the accuracy $\sim 1/c^2$, we have

$$\dot{\mathbf{q}}_i = \frac{\mathbf{p}_i}{m_i} - \frac{1}{m_i c^2}\left\{\frac{1}{2m_i^2} p_i^2 \mathbf{p}_i + \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{3\mathcal{G} m_j}{r_{ij}} \mathbf{p}_i - \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{\mathcal{G} m_i m_j}{2 r_{ij}} \left[\frac{7}{m_j}\mathbf{p}_j + \frac{1}{m_i}(\mathbf{p}_i \cdot \mathbf{n}_{ij})\mathbf{n}_{ij}\right]\right\}. \tag{2.82}$$

In fact, consider the following problem. Let $n \geq 1$ and let $\mathbf{f}$ a function defined as

$$\begin{aligned} \mathbf{f} \colon \mathbb{R}^n &\to \mathbb{R}^n \\ \mathbf{x} \mapsto \mathbf{y} &= \mathbf{f}(\mathbf{x}) = A\mathbf{x} + \epsilon \boldsymbol{\varphi}(\mathbf{x}) \end{aligned} \tag{2.83}$$

where A is a $n \times n$ real matrix such that $\det A \neq 0$ (i.e. A is invertible), ϵ is a parameter which satisfies the condition $|\epsilon| \ll 1$ and $\boldsymbol{\varphi}$ is a $C^1(\mathbb{R}^n, \mathbb{R}^n)$ function such that for each $\mathbf{x} \in \mathbb{R}^n$

$$\det\left[A + \epsilon D_{\mathbf{x}}(\boldsymbol{\varphi}(\mathbf{x}))\right] \neq 0,$$

where $D_{\mathbf{x}}(\boldsymbol{\varphi}(\mathbf{x}))$ is the Jacobian matrix of $\boldsymbol{\varphi}(\mathbf{x})$. Because $\mathbf{f}$ satisfies the hypotheses of the *inverse function theorem*, for every $\mathbf{x} \in \mathbb{R}^n$ there exists a neighborhood $U \subset \mathbb{R}^n$ of $\mathbf{x}$ such that the function $\mathbf{f}$ is invertible in U.

To find a function that approximates the inverse of $\mathbf{f}$, we proceed as follows. Using the fact that $\mathbf{x} = A^{-1}\mathbf{y} - \epsilon A^{-1}\boldsymbol{\varphi}(\mathbf{x})$ and that $\mathbf{x} \simeq A^{-1}\mathbf{y}$ (because $\epsilon \ll 1$), we have

$$\mathbf{x} = A^{-1}\mathbf{y} - \epsilon A^{-1}\boldsymbol{\varphi}(\mathbf{x})|_{\mathbf{x}=A^{-1}\mathbf{y}-\epsilon A^{-1}\boldsymbol{\varphi}(\mathbf{x})} = A^{-1}\mathbf{y} - \epsilon A^{-1}\boldsymbol{\varphi}(A^{-1}\mathbf{y}) + \mathbf{R}(\mathbf{y}) \tag{2.84}$$

where $\| \mathbf{R}(\mathbf{y}) \| = O(\epsilon^2)$ (because $\mathbf{R}(\mathbf{y}) \cong \epsilon A^{-1}[\boldsymbol{\varphi}(A^{-1}\mathbf{y}) - \boldsymbol{\varphi}(A^{-1}\mathbf{y} - \epsilon A^{-1}\boldsymbol{\varphi}(A^{-1}\mathbf{y}))]$). Thus, less than errors of order ϵ^2, the inverse of the function (2.83) is

$$\mathbf{x} = \mathbf{f}^{-1}(\mathbf{y}) \cong A^{-1}\mathbf{y} - \epsilon A^{-1}\boldsymbol{\varphi}(A^{-1}\mathbf{y}). \tag{2.85}$$

In our case, the problem to invert the function $\mathbf{p} = \mathbf{p}(\mathbf{x}, \mathbf{v})$ given in (2.81) (where $\mathbf{x}$ are considered fixed but arbitrary) can be solved using the solution of the more generic problem studied in (2.83). Thus, using (2.85), the inverse of $\mathbf{p} = \mathbf{p}(\mathbf{x}, \mathbf{v})$ is given by $\mathbf{v} = \mathbf{v}(\mathbf{x}, \mathbf{p})$ in (2.82), less than errors of order c^{-4}.

Using (2.73)-(2.79)-(2.82), it simple to prove that, less than terms of fourth-order, the relativistic Hamiltonian is

$$\begin{aligned} H = & \sum_{i=1}^{N} \frac{1}{2m_i} p_i^2 - \mathcal{G} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{m_i m_j}{2 r_{ij}} - \\ & - \sum_{i=1}^{N} \frac{p_i^4}{8c^2 m_i^3} - \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{3\mathcal{G} p_j^2 m_i}{2c^2 m_j r_{ij}} + \\ & + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{\mathcal{G}}{4c^2 r_{ij}} [7\mathbf{p}_i \cdot \mathbf{p}_j + (\mathbf{p}_i \cdot \mathbf{n}_{ij})(\mathbf{p}_j \cdot \mathbf{n}_{ij})] + \\ & + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \sum_{\substack{k=1 \\ k \neq i}}^{N} \frac{\mathcal{G}^2 m_i m_j m_k}{2c^2 r_{ij} r_{jk}}. \end{aligned} \tag{2.86}$$

where, as before, $p_i^2 = \mathbf{p}_i \cdot \mathbf{p}_i$ and $p_i^4 = (\mathbf{p}_i \cdot \mathbf{p}_i)^2$.

2.4.2 Lense-Thirring effect

In the Newtonian case, the rotation of the bodies around their axes has no effect on the orbit of each body, because, in Newtonian mechanics, the gravitational field of a body depends only on its mass and not on its rotation. Instead, Einstein's General Theory of Relativity predicts that, given a system of massive bodies that rotate around their axes, part of the distortion of space-time due to the presence of bodies is given precisely by the rotation of the bodies on themselves. In particular, General Relativity predicts that the rotation of massive objects would distort the space-time metric, so that the space-time surrounding the objects is "dragged" along with them as they rotate (for this reason the effect was called "frame-dragging"). One way to visualize this effect is to place a small ball in a thick fluid such as honey. As the ball spins, it pulls the honey around itself. Anything stuck in the honey will also move around the ball. Similarly, as a massive object rotates, it pulls space-time in its vicinity around itself, making the orbit of a nearby test particle precess. On the other hand, this effect is very small, and in particular it is irrelevant if the distances between the bodies are very large and the speed of rotation of the bodies around their axes is very small.

For simplicity, let us consider the stationary weak gravitational field possessing central symmetry produced by a symmetric rotating central body of mass m'. This case was studied for the first time in 1918 by the physicists J. Lense and H. Thirring and this simple model allows us to determine the effects due to the rotation of the central Star on the orbit of the (non-rotating) planets.

The problem is to determine the systematic ("secular") shift of the orbit of a particle of mass m moving in the field of a central body, associated with the rotation of the latter. It can be proven that, in Cartesian coordinates, the space-time metric at large distance r from the rotating central body is given by

$$g_{ik} = \eta_{ik} + h^{(1)}_{ik} + h^{(2)}_{ik}, \tag{2.87}$$

where η_{ik} is the Minkowski metric and

$$\begin{aligned} &h^{(1)}_{00} = -\frac{r_g}{r}, \qquad h^{(1)}_{\alpha\beta} = -\frac{r_g}{r}\mathbf{n}_\alpha\mathbf{n}_\beta, \qquad h^{(1)}_{0\alpha} = 0, \\ &h^{(2)}_{00} = 0, \qquad h^{(2)}_{\alpha\beta} = -\left(\frac{r_g}{r}\right)^2\mathbf{n}_\alpha\mathbf{n}_\beta, \qquad h^{(2)}_{0\alpha} = -\frac{2\mathcal{G}}{c^3}M'_{\alpha\beta}\frac{\mathbf{n}_\beta}{r^2} \end{aligned} \tag{2.88}$$

where $r_g = 2\mathcal{G}m'/c^2$ and where $M'_{\alpha\beta}$ is the angular momentum of the rotating body.
In particular, if the central rotating body is a homogeneous sphere of radius R and mass m', then

$$M'_{\alpha\beta} = \begin{bmatrix} 0 & M'_z & -M'_y \\ -M'_z & 0 & M'_x \\ M'_y & -M'_x & 0 \end{bmatrix} \tag{2.89}$$

and

$$\mathbf{M}' = [M'_x, M'_y, M'_z] = I\boldsymbol{\omega}, \qquad I = \frac{2}{5}m'R^2 \tag{2.90}$$

where I is the moment of inertia and $\boldsymbol{\omega}$ is the angular velocity vector.

It can be proven that, in a stationary gravitational field, there acts on the particle a "Coriolis force" equal to that which would act on the particle if it were on a body rotating with angular velocity

$$\boldsymbol{\Omega} = \frac{c}{2}\sqrt{g_{00}}\nabla \times g, \qquad \mathbf{g} = \frac{2\mathcal{G}}{c^3 r^2}\mathbf{n} \times \mathbf{M}'. \tag{2.91}$$

Therefore we may say that in the field produced by a rotating body (with total angular momentum $\mathbf{M}'$) there acts on a particle distant from the body a force which is equivalent to the Coriolis force which would appear for a rotation with angular velocity

$$\boldsymbol{\Omega} \simeq \frac{c}{2}\nabla \times \mathbf{g} = \frac{\mathcal{G}}{c^2 r^2}[\mathbf{M}' - 3\mathbf{n}(\mathbf{n}\cdot\mathbf{M}')]. \tag{2.92}$$

The Lagrangian for a particle moving in the field (2.87) – (2.88) of a central rotating body is $L = -mc\frac{\mathrm{d}s}{\mathrm{d}t}$. Because all the relativistic effects are small, they superpose linearly with one another, so in calculating the effects resulting from the rotation of the central body we can neglect the influence of the non-Newtonian centrally symmetric force field which we considered in section

2.3.2; in other words, we can make the computations assuming that of all the h_{ik} only the $h_{0\alpha}$ are different from zero.

Thus, the required Lagrangian is (*J. Lense and H. Thirring, 1918*)

$$L = L_0 + \delta L = \frac{1}{2}\mu\mathbf{v}\cdot\mathbf{v} + \frac{\mathcal{G}mm'}{\parallel \mathbf{r} \parallel} + \frac{2\mathcal{G}\mu}{c^2 r^3}\mathbf{M}'\cdot(\mathbf{v}\times\mathbf{r}), \tag{2.93}$$

where $\mathbf{r}$ is the astrocentric position of the particle ($r = \parallel \mathbf{r} \parallel$), $\mu = mm'/(m+m')$, $\mathbf{v} = \dot{\mathbf{r}}$ and $\mathbf{M}'$ is the angular momentum of the central body. Then the Hamiltonian is[6]:

$$H = H_0 + \delta H = \frac{1}{2\mu}\mathbf{p}\cdot\mathbf{p} - \frac{\mathcal{G}mm'}{\parallel \mathbf{r} \parallel} + \frac{2\mathcal{G}}{c^2 r^3}\mathbf{M}'\cdot(\mathbf{r}\times\mathbf{p}), \tag{2.94}$$

where $\mathbf{p} = \mu\mathbf{v}$.

As we have seen, the orientation of the classical orbit of the particle is determined by two conserved quantities, which are the angular momentum of the particle $\mathbf{M}$ and the Runge-Lenz vector $\mathbf{A}$:

$$\mathbf{M} = \mathbf{r}\times\mathbf{p}, \qquad \mathbf{A} = \frac{\mathbf{p}}{\mu}\times\mathbf{M} - \frac{\mathcal{G}mm'\mathbf{r}}{\parallel \mathbf{r} \parallel}. \tag{2.95}$$

In particular, the vector $\mathbf{M}$ is perpendicular to the plane of the orbit, while the vector $\mathbf{A}$ is directed along the major axis of the ellipse toward the perihelion (and $\parallel \mathbf{A} \parallel = \mathcal{G}mm'e$, where e is the eccentricity of the orbit). Thus, the required shift of the orbit can be described in terms of the change in direction of these vectors.

Using Hamilton's equations, it can be proven that the secular change of $\mathbf{M}$ is given by the formula

$$\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = \frac{2\mathcal{G}}{c^2 a^3 (1-e^2)^{\frac{3}{2}}}\mathbf{M}'\times\mathbf{M}, \tag{2.96}$$

where a and e are the semi-major axis and the eccentricity of the ellipse, i.e. the vector $\mathbf{M}$ rotates around the axis of rotation of the central body, remaining fixed in magnitude.

Moreover, it can be shown that the secular change of the vector $\mathbf{A}$ is given by the formula

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \mathbf{\Omega}\times\mathbf{A}, \tag{2.97}$$

where

$$\mathbf{\Omega} = \frac{2\mathcal{G}M'}{c^2 a^3 (1-e^2)^{\frac{3}{2}}}\{\mathbf{n}' - 3\mathbf{n}(\mathbf{n}\cdot\mathbf{n}')\}, \tag{2.98}$$

and where $M' = \parallel \mathbf{M}' \parallel$ and $\mathbf{n}$ and $\mathbf{n}'$ are unit vector along the direction of $\mathbf{M}$ and $\mathbf{M}'$. Formula (2.97) shows that the vector $\mathbf{A}$ rotates with angular velocities $\mathbf{\Omega}$, remaining fixed in magnitude, i.e. the eccentricity of the orbit does not undergo any secular change.

Formula (2.96) can be re-written as

$$\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = \mathbf{\Omega}\times\mathbf{M} \tag{2.99}$$

[6]Let the Lagrangian be of the form $L = L_0 + L'$, where L' is a small correction of $L = L_0$. It can be proven that the corresponding addition H' in the Hamiltonian $H = H_0 + H'$ is related to L' by

$$(H')_{q,p} = -(L')_{q,\dot{q}}.$$

with the same $\mathbf{\Omega}$ as in (2.98). Thus, $\mathbf{\Omega}$ is the angular velocity of rotation of the ellipse "as a whole": this rotation includes both the additional (compared to that considered in section 2.3.1) shift of the perihelion of the orbit, and the secular rotation of its plane about the direction of the axis of the two body (where the latter effect is absent if the plane of the orbit coincides with the equatorial plane of the central body, i.e. $i = \pi/2$).

For comparison we note that the effect considered in section 2.3.1 corresponds to

$$\mathbf{\Omega} = \frac{6\pi\mathcal{G}(m_0 + m_1)}{c^2 a(1 - e^2)T}\mathbf{n}, \tag{2.100}$$

where T is the period of revolution of the particle around the central body.

Using (2.96) and (2.97)-(2.98), we find for the longitude of the ascending node Ω and for the argument of the pericenter ω, respectively, a pro-/retrograde precession of

$$\Delta\Omega = \frac{2\mathcal{G}M'}{c^2 a^3 (1 - e^2)^{\frac{3}{2}}}, \qquad \Delta\omega = -\frac{6\mathcal{G}M' \cos i}{c^2 a^3 (1 - e^2)^{\frac{3}{2}}}. \tag{2.101}$$

Now we can apply the above formulæ in the case of the Solar system to determine the effects due to the rotation of the Sun. By assuming an homogeneous, spherically and uniformly rotating Sun, if we adopt $M' = 1.9 \times 10^{41}$ Kg m^2 s^{-1} for the solar proper angular momentum, we found that the numerical values of the shifts of the argument of the pericenter for Mercury and Earth are equal, respectively, to -0.003" and -0.00015" per century. As expected, in the case of the Solar system, the Lense-Thirring effect is quite small (in fact the angular momentum of the Sun is relatively small and the Lense-Thirring precession fall off with the inverse of the third power of the planet's semi-major axis).

For these reasons, because in the following we consider only systems in which the angular momentum of the central star is small, we decide to assume that all bodies are not rotating, making a negligible error.

Chapter 3

Numerical integration of the planetary problem

We have seen that, in the configuration space $\mathbb{R}^{6(N+1)}$, the non-relativistic Hamiltonian of a system made of a star of mass m_0 and of N planets of masses $m_1, m_2, ...m_N$ which interact with one other but with no others bodies, is

$$H = \sum_{i=0}^{N} \frac{1}{2m_i} y_i^2 - \mathcal{G} \sum_{i=0}^{N} \sum_{\substack{j=0 \\ j\neq i}}^{N} \frac{m_i m_j}{2r_{ij}} \tag{3.1}$$

where $\mathbf{x}_0, ..., \mathbf{x}_N$ are the positions of the $N+1$ bodies, $r_{ij} = \parallel \mathbf{x}_j - \mathbf{x}_i \parallel$ for $0 \leq i, j \leq N$, $\mathbf{y}_i = m_i \dot{\mathbf{x}}_i$ and $y_i^2 = \mathbf{y}_i \cdot \mathbf{y}_i$ for $i = 0, ..., N$.

In the relativistic case, less than errors of order c^{-4}, the Hamiltonian is

$$\begin{aligned} H =& \sum_{i=0}^{N} \frac{1}{2m_i} y_i^2 - \mathcal{G} \sum_{i=0}^{N} \sum_{\substack{j=0 \\ j\neq i}}^{N} \frac{m_i m_j}{2r_{ij}} + \frac{1}{c^2} \Bigg\{ - \sum_{i=0}^{N} \frac{y_i^4}{8m_i^3} - \sum_{i=0}^{N} \sum_{\substack{j=0 \\ j\neq i}}^{N} \frac{3\mathcal{G} y_j^2 m_i}{2m_j r_{ij}} + \\ &+ \sum_{i=0}^{N} \sum_{\substack{j=0 \\ j\neq i}}^{N} \frac{\mathcal{G}}{4r_{ij}} [7\mathbf{y}_i \cdot \mathbf{y}_j + (\mathbf{y}_i \cdot \mathbf{n}_{ij})(\mathbf{y}_j \cdot \mathbf{n}_{ij})] + \sum_{i=0}^{N} \sum_{\substack{j=0 \\ j\neq i}}^{N} \sum_{\substack{k=0 \\ k\neq i}}^{N} \frac{\mathcal{G}^2 m_i m_j m_k}{2r_{ij} r_{jk}} \Bigg\}. \end{aligned} \tag{3.2}$$

where $\mathbf{n}_{ij} = (\mathbf{r}_i - \mathbf{r}_j)/r_{ij}$, $y_i^2 = \mathbf{y}_i \cdot \mathbf{y}_i$, $y_i^4 = (\mathbf{y}_i \cdot \mathbf{y}_i)^2$ and where $\mathbf{y}_i$ is defined as:

$$\mathbf{y}_i = m_i \dot{\mathbf{x}}_i + \frac{1}{c^2} \left\{ \frac{1}{2} m_i (\dot{\mathbf{x}}_i \cdot \dot{\mathbf{x}}_i) \dot{\mathbf{x}}_i + \sum_{\substack{j=0 \\ j\neq i}}^{N} \frac{3\mathcal{G} m_i m_j}{r_{ij}} \dot{\mathbf{x}}_i - \sum_{\substack{j=0 \\ j\neq i}}^{N} \frac{\mathcal{G} m_i m_j}{2r_{ij}} [7\dot{\mathbf{x}}_j + (\dot{\mathbf{x}}_i \cdot \mathbf{n}_{ij}) \mathbf{n}_{ij}] \right\}. \tag{3.3}$$

The aim of this chapter is to study numerically the dynamics described by the Hamiltonian (3.1) and (3.2) for some extrasolar systems to see if there are significant differences between the Newtonian and the relativistic case. In particular, we study only the case $N+1 = 3$ and we consider only systems far enough from collisions and such that no strong mean motion resonances are present.

The numerical integration of the equations of motion corresponding to the Hamiltonian (3.2) is greatly slow, and it is useful only in situations where the relativistic contribution of every object in the system is to be taken into account. Thus we look for a simplification of the relativistic Hamiltonian. In particular, we skip the relativistic corrections due to the mutual interactions between the two planets, i.e. we consider the mutual interaction between the two planets as a Newtonian one. This choice leads to a strong simplification of the relativistic Hamiltonian and it is justified by the fact that the relativistic corrections due to the mutual interactions between the two planets is a negligible correction compared with the relativistic corrections due to the mutual interactions between the star and the planets.

The simplified Hamiltonian, although less exact than Hamiltonian (3.2), is computationally much more affordable and, as we will show, the dynamic described by the simplified Hamiltonian is very similar to the real one, at least numerically in the systems that we have considered.

3.1 Canonical transformations

As we have seen, a system of ordinary differential equations of the type

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{F}(\mathbf{r}) \tag{3.4}$$

is said to be *Hamiltonian form* if $\mathbf{r}$ is a $2n$-uple and, denoting by $q_1, q_2, ..., q_n$ and $p_1, p_2, ..., p_n$ its $2n$ components, there exist a function $H(q_1, ..., q_n, p_1, ..., p_n)$ (called *Hamiltonian*) such that equations (3.4) can be rewritten as

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i} \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q_i} \tag{3.5}$$

for $i = 1, ..., n$. The variables $\mathbf{q}$ and $\mathbf{p}$ are called the *conjugate variables* and, in particular, $q_1, q_2, ..., q_n$ and $p_1, p_2, ..., p_n$ are respectively called *coordinates* and *momenta*. The $(\mathbf{q}, \mathbf{p})$ space is usually a $2n-$dimensional differentiable manifold called the *phase space* of the system and the dimension n of the vector $\mathbf{q}$ and $\mathbf{p}$ is called the *number of degrees of freedom*.

As for a generic system of differential equations, coordinate transformations can be used in order to bring the system to a simpler form. In particular, in the Hamiltonian world, the idea is to use the link between the Hamilton's equations and the Hamiltonian to make the change of coordinates in each of the $2n$ differential equations, i.e. we put the Hamiltonian in the new variables and from it we derive the equations of Hamilton in the new variables. However, not all changes of coordinate preserve the form of the canonical equations of motion, in the sense that the second members are not expressed as derivatives of a unique function. This leads to introduce the concept of *canonical transformation*.

Definition 1. *A time independent transformation of coordinate and momenta in phase space is said to be canonical if it preserve the Hamiltonian form of the equations of motion, whatever the Hamiltonian function. More precisely, a transformation* $(\mathbf{q}, \mathbf{p}) \to (\mathbf{q}', \mathbf{p}')$ *is canonical if to every Hamiltonian* $H(\mathbf{q}, \mathbf{p})$ *one can associate another function* $K(\mathbf{q}', \mathbf{p}')$ *such that the equation of motion for* $\mathbf{q}'$ *and* $\mathbf{p}'$ *become:*

$$\dot{q}'_i = \frac{\partial K}{\partial p'_i} \qquad \dot{p}'_i = -\frac{\partial K}{\partial q'_i}, \qquad 1 \le i \le n. \tag{3.6}$$

In some cases, the new Hamiltonian $K(\mathbf{q}', \mathbf{p}')$ can be constructed by a straightforward substitution of the transformation in the old one:

$$K(\mathbf{q}', \mathbf{p}') = H(\mathbf{q}(\mathbf{q}', \mathbf{p}'), \mathbf{p}(\mathbf{q}', \mathbf{p}')). \tag{3.7}$$

The class of canonical transformations is a very restrictive class among all possible transformations on the phase space, but if we don't want to lose the Hamiltonian form of the equations we are compelled to restrict to canonical transformations.

The most useful criterion, to test whether a given transformation is canonical, is that of Poisson bracket[1]. It can be proven that a transformation $(\mathbf{q}, \mathbf{p}) \to (\mathbf{q}', \mathbf{p}')$ is canonical if and only if it preserves the fundamental Poisson bracket, i.e., considering the components $q'_1, q'_2, ..., q'_n$ of $\mathbf{q}'$ and $p'_1, p'_2, ..., p'_n$ of $\mathbf{p}'$ as a function of $(\mathbf{q}, \mathbf{p})$, one has

$$\{p'_i, p'_j\} = 0, \qquad \{q'_i, q'_j\} = 0, \qquad \{p'_i, q'_j\} = \delta_i, j \tag{3.9}$$

where $\delta_{i,j}$ is 1 if $i = j$ and 0 otherwise.

3.2 Barycentric coordinates

In astronomy, *barycentric coordinates* are non-rotating coordinates with origin at the center of mass of the $N + 1$ bodies.

In classical mechanics, using the conservation of the *total linear momentum* $\mathbf{P} = \sum_{i=0}^{N} \mathbf{y}_i$ (where $\mathbf{y}_i = m_i \dot{\mathbf{x}}_i$), it is simple to prove that the center of mass $\mathbf{R}$ moves of uniform rectilinear motion:

$$\mathbf{R}(t) = \frac{\sum_{i=0}^{N} m_i \mathbf{x}_i(t)}{\sum_{i=0}^{N} m_i} = \mathbf{R}(0) + \frac{\mathbf{P}}{\sum_{i=0}^{N} m_i} t. \tag{3.10}$$

In relativistic mechanics, the coordinates of the center of inertia are given by the formula

$$\begin{aligned} \mathbf{R}(t) &= \frac{1}{E} \sum_{i=0}^{N} \mathbf{x}_i(t) \left(m_i c^2 + \frac{y_i^2(t)}{2m_i} - \mathcal{G} \sum_{\substack{j=0 \\ j \neq i}}^{N} \frac{m_i m_j}{2 r_{ij}(t)} \right) \\ E &= \sum_{i=0}^{N} \left(m_i c^2 + \frac{p_i^2}{2m_i} - \frac{\mathcal{G} m_i}{2} \sum_{j \neq i} \frac{m_i}{r_{ij}} \right) \end{aligned} \tag{3.11}$$

where E is the total energy of the system ($\dot{E} = O(c^{-4})$). Also in this case, we have seen that the barycenter moves uniformly with constant velocity $\dot{\mathbf{R}} = \mathbf{P}/E$, i.e. $\ddot{\mathbf{R}} = O(c^{-4})$, where $\mathbf{P} = \sum_{i=0}^{N} \mathbf{y}_i$ with $\mathbf{y}_i$ defined in (3.3).

Thus, using the conservation of the total linear momentum, a first possible reduction of the equations of motion is obtained imposing in both cases the conditions

$$\mathbf{R}(t) = 0, \qquad \dot{\mathbf{R}}(t) = 0, \qquad \forall t, \tag{3.12}$$

[1] In general, let $f(\mathbf{q}, \mathbf{p})$ and $g(\mathbf{q}, \mathbf{p})$ be differentiable dynamical variables; then the *Poisson bracket* is defined as

$$\{f, g\} = \sum_{j=1}^{n} \left(\frac{\partial f}{\partial q_j} \frac{\partial g}{\partial p_j} - \frac{\partial f}{\partial p_j} \frac{\partial g}{\partial q_j} \right). \tag{3.8}$$

in order to reduce of 3 the number of degrees of freedom of the system.

A reference system which satisfies the condition (3.12) is called a barycentric inertial reference system.

3.3 Heliocentric coordinates

In a barycentric inertial reference system, we introduce the canonical *heliocentric coordinates* which describe the position of each body with respect to the position of the central star.

The heliocentric coordinates $\mathbf{r}_0, ..., \mathbf{r}_N$ are given by the transformation

$$\mathbf{r}_0 = \mathbf{x}_0, \qquad \mathbf{r}_k = \mathbf{x}_k - \mathbf{x}_0, \qquad k = 1, ..., N \tag{3.13}$$

and the inverse transformation is

$$\mathbf{x}_0 = \mathbf{r}_0, \qquad \mathbf{x}_k = \mathbf{r}_0 + \mathbf{r}_k, \qquad k = 1, ..., N. \tag{3.14}$$

In matrix notation, the change of coordinate is given by

$$\begin{bmatrix} \mathbf{r}_0 \\ \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_{N-1} \\ \mathbf{r}_N \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 \\ -1 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & 0 & 0 & \dots & 1 & 0 \\ -1 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_{N-1} \\ \mathbf{x}_N \end{bmatrix} = L \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_{N-1} \\ \mathbf{x}_N \end{bmatrix}. \tag{3.15}$$

Obviously, the linear transformation of coordinates $\mathbf{x}_0, ..., \mathbf{x}_N$ to the coordinates $\mathbf{r}_0, ..., \mathbf{r}_N$ given by

$$\mathbf{r}_k = \sum_{j=0}^{N} A_{kj} \mathbf{x}_j, \qquad k = 0, ..., N \tag{3.16}$$

where A is a $3(N+1) \times 3(N+1)$ non-singular matrix, is a point transformation. It can be extended to a canonical transformation in the following way: if $\mathbf{y}_0, ..., \mathbf{y}_N$ are the momenta conjugate to the old coordinates $\mathbf{x}_0, ..., \mathbf{x}_N$ and if $\mathbf{p}_0, ..., \mathbf{p}_N$ are the momenta conjugate to the new coordinates $\mathbf{r}_0, ..., \mathbf{r}_N$, it can be proven that the canonical transformation of the momenta is

$$\mathbf{p}_k = \sum_{j=0}^{N} \left[(A^{-1})^T \right]_{kj} \mathbf{y}_j, \qquad k = 0, ..., N. \tag{3.17}$$

Moreover, it is easy to show that the angular momentum maintains the same form under a linear transformation of the type (3.16)-(3.17), i.e.

$$\mathbf{M} = \sum_{i=0}^{N} \mathbf{x}_i \times \mathbf{y}_i = \sum_{i=0}^{N} \mathbf{r}_i \times \mathbf{p}_i. \tag{3.18}$$

In our cases, the canonical transformation on the momenta is given by

$$\begin{bmatrix} \mathbf{p}_0 \\ \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_{N-1} \\ \mathbf{p}_N \end{bmatrix} = \left(L^{-1} \right)^T \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_{N-1} \\ \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_{N-1} \\ \mathbf{y}_N \end{bmatrix}. \tag{3.19}$$

In a more compact form, the change of momenta is given by

$$\mathbf{p}_0 = \sum_{i=0}^{N} \mathbf{y}_i, \qquad \mathbf{p}_k = \mathbf{y}_k, \qquad k = 1, ..., N \tag{3.20}$$

and the inverse transformation is

$$\mathbf{y}_0 = \mathbf{p}_0 - \sum_{i=1}^{N} \mathbf{p}_i, \qquad \mathbf{y}_k = \mathbf{p}_k, \qquad k = 1, ..., N. \tag{3.21}$$

where $\mathbf{p}_0$ is the total linear momentum. As we have seen, $\mathbf{p}_0$ is constant, and, in particular, because we are working in the barycentric reference system, it is equal to zero, i.e. $\mathbf{p}_0 = \mathbf{0}$.

Using the heliocentric coordinates, the Hamiltonians can be written as the sum of two parts: a *main part*, which describes the Keplerian motion of the (individual) planets under the action of the central star, and a *perturbation part*, which is small with respect to the main part and which takes into account the interactions of the planets and the relativistic corrections.

3.4 Transformation of the Hamiltonian in heliocentric coordinates

As already said, in the following we limit ourselves to study the problem of three bodies. In this case, we assign a privileged role to the central star taking into account that its mass m_0 is appreciably higher than that of the planets (i.e. $m_1, m_2 \ll m_0$).

3.4.1 Classical Hamiltonian

With simple calculations, we find that, in the barycentric reference system, the classical Hamiltonian in the heliocentric variables $(\mathbf{r}, \mathbf{p})$ becomes

$$H = H_0 + \varepsilon H_1, \tag{3.22}$$

where

$$\begin{aligned} H_0 &= \sum_{i=1}^{2} \left(\frac{1}{2\mu_i} p_i^2 - \mathcal{G} \frac{(m_0 + m_i)\mu_i}{\parallel \mathbf{r}_i \parallel} \right), \\ \varepsilon H_1 &= \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{m_0} - \mathcal{G} \frac{m_1 m_2}{\parallel \mathbf{r}_1 - \mathbf{r}_2 \parallel}, \end{aligned} \tag{3.23}$$

and where μ_i is the reduced masses of the planet i-th

$$\mu_i = \frac{m_0 m_i}{m_0 + m_i}, \qquad i = 1, 2. \tag{3.24}$$

The Hamiltonian (3.22) depends only by the 12 canonical coordinate $\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2$, which means that, as expected, eliminating the motion of the center of gravity, we have reduced of 3 the number of degrees of freedom of the system.

We can notice that the Hamiltonian H_0 can be read as the sum of 2 Hamiltonians

$$H_0^i = \frac{1}{2\mu_i} p_i^2 - \mathcal{G} \frac{(m_0 + m_i)\mu_i}{\parallel \mathbf{r}_i \parallel}, \qquad i = 1, 2 \tag{3.25}$$

of two-body problem with reduced mass μ_i. Thus H_0^i describes the Keplerian motion of the $i-$th planet around the star.

The terms H_1 contains instead the interaction between the planets. If the mutual distance $\| \mathbf{r}_1 - \mathbf{r}_2 \|$ does not become small, its size relative to the two-body problem Hamiltonian is proportional to a small factor ε, which can be approximated by

$$\varepsilon \simeq \max_{i=1,2} \frac{\mu_i}{m_0}, \tag{3.26}$$

where $\mu_j/m_0 \simeq m_j/m_0$. For example, in the case of the Solar system, ε is well approximated by the ratio of the mass of Jupiter and of the Sun, which is approximately $\varepsilon \simeq 10^{-3}$.

The canonical equations for the Hamiltonian (3.22) are

$$\begin{aligned} \dot{\mathbf{r}}_k &= \frac{\mathbf{p}_k}{\mu_k} + \frac{\mathbf{p}_{3-k}}{m_0}, \\ \dot{\mathbf{p}}_k &= -\frac{\mathcal{G}(m_0 + m_k)\mu_k \mathbf{r}_k}{\| \mathbf{r}_k \|^3} - \mathcal{G} m_k m_{3-k} \frac{\mathbf{r}_k - \mathbf{r}_{3-k}}{\| \mathbf{r}_k - \mathbf{r}_{3-k} \|^3}, \end{aligned} \tag{3.27}$$

for $k = 1, 2$. We can notice that $\mathbf{p}_k$ are not proportional to the velocity $\dot{\mathbf{r}}_k$. This is a consequence of the fact that the kinetic energy has not a diagonal shape, a property which instead was true for the coordinates in the absolute reference system.

Finally, being the positions of the planets known, the position of the star $\mathbf{x}_0$ is given by

$$\mathbf{x}_0 = -\frac{\sum_{i=1}^{2} m_i \mathbf{r}_i}{\sum_{i=0}^{2} m_i}. \tag{3.28}$$

3.4.2 Relativistic Hamiltonian

In the barycentric reference system, we find that the relativistic Hamiltonian in the heliocentric variables $(\mathbf{r}, \mathbf{p})$ [2] becomes

$$H = H_0 + \varepsilon H_1 + \frac{1}{c^2} H_2, \tag{3.29}$$

where

$$\begin{aligned} H_0 =& \sum_{i=1}^{2} \left(\frac{1}{2\mu_i} p_i^2 - \mathcal{G} \frac{(m_0 + m_i)\mu_i}{\| \mathbf{r}_i \|} \right), \\ \varepsilon H_1 =& \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{m_0} - \mathcal{G} \frac{m_1 m_2}{\| \mathbf{r}_1 - \mathbf{r}_2 \|}, \end{aligned} \tag{3.30}$$

[2] Here we use the same notation $(\mathbf{r}, \mathbf{p})$ that we have introduced for the classical Hamiltonian. Of course it is immediate to observe that they are not the same coordinates. On the other hand, in the following we never use simultaneously both cases, so there is no risk of confusion.

$$
\begin{aligned}
H_2 = &-\sum_{i=1}^{2}\frac{p_i^4}{8}\left(\frac{1}{m_0^3}+\frac{1}{m_i^3}\right)-\frac{1}{4m_0^3}\left(p_1^2p_2^2+2(\mathbf{p}_1\cdot\mathbf{p}_2)^2+2\sum_{i=1}^{2}p_i^2(\mathbf{p}_1\cdot\mathbf{p}_2)\right)+\\
&-\frac{3\mathcal{G}}{2m_0}\sum_{i=1}^{2}\frac{m_i}{\parallel \mathbf{r}_i \parallel}\left(p_1^2+2\mathbf{p}_1\cdot\mathbf{p}_2+p_2^2\right)-\sum_{i=1}^{2}\frac{3\mathcal{G}p_i^2}{2m_i}\left(\frac{m_0}{\parallel \mathbf{r}_i \parallel}+\frac{m_{3-i}}{\parallel \mathbf{r}_1-\mathbf{r}_2 \parallel}\right)+\\
&-\sum_{i=1}^{2}\frac{7\mathcal{G}}{2\parallel \mathbf{r}_i \parallel}\mathbf{p}_i\cdot(\mathbf{p}_1+\mathbf{p}_2)+\frac{7\mathcal{G}}{2\parallel \mathbf{r}_1-\mathbf{r}_2 \parallel}\mathbf{p}_1\cdot\mathbf{p}_2-\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\mathcal{G}}{2\parallel \mathbf{r}_i \parallel}(\mathbf{p}_i\cdot\mathbf{n}_i)(\mathbf{p}_j\cdot\mathbf{n}_i)+\\
&+\frac{\mathcal{G}}{2\parallel \mathbf{r}_2-\mathbf{r}_1 \parallel}(\mathbf{p}_1\cdot\mathbf{n}_{12})(\mathbf{p}_2\cdot\mathbf{n}_{12})+\frac{\mathcal{G}^2m_0m_1m_2}{\parallel \mathbf{r}_1 \parallel\parallel \mathbf{r}_2 \parallel}+\sum_{i=1}^{2}\frac{\mathcal{G}^2m_0m_1m_2}{\parallel \mathbf{r}_i \parallel\parallel \mathbf{r}_1-\mathbf{r}_2 \parallel}+\\
&+\frac{\mathcal{G}^2}{2}\left(\sum_{i=1}^{2}\frac{m_0m_i}{\parallel \mathbf{r}_i \parallel^2}(m_0+m_i)+\frac{m_1m_2}{\parallel \mathbf{r}_2-\mathbf{r}_1 \parallel^2}(m_1+m_2)\right),
\end{aligned}
$$

where $\mathbf{n}_j = \mathbf{r}_j/\parallel \mathbf{r}_j \parallel$ and $\mathbf{n}_{12} = (\mathbf{r}_2-\mathbf{r}_1)/\parallel \mathbf{r}_2-\mathbf{r}_1 \parallel$. As before, the terms H_1 contains the interaction between the planets and, if none of their mutual distance $\parallel \mathbf{r}_i-\mathbf{r}_j \parallel$ becomes small, its size relative to H_0 is proportional to the small factor ε given in (3.26). The terms H_2 contains the relativistic corrections and its size relative to H_0 is proportional to c^{-2}.

Also in this case, we can notice that the Hamiltonian (3.29) does not depend on $\mathbf{r}_0$ and so $\mathbf{p}_0$ is constant (in particular $\mathbf{p}_0 = \mathbf{0}$ because we are working in the barycentric reference system).

The canonical equations for the Hamiltonian (3.29) are

$$
\begin{aligned}
\dot{\mathbf{r}}_k = &\frac{\mathbf{p}_k}{\mu_k}+\frac{\mathbf{p}_{3-k}}{m_0}+\frac{1}{c^2}\Bigg\{\mathbf{p}_{3-k}\left[-\frac{p_1^2+p_2^2}{2m_0^3}-\sum_{i=1}^{2}\frac{7\mathcal{G}}{2\parallel \mathbf{r}_i \parallel}+\frac{7\mathcal{G}}{2\parallel \mathbf{r}_1-\mathbf{r}_2 \parallel}\right]+\\
&+\mathbf{p}_k\left[-\frac{p_k^2}{2}\left(\frac{1}{m_0^3}+\frac{1}{m_k^3}\right)-\frac{p_{3-k}^2}{2m_0^3}-\frac{7\mathcal{G}}{\parallel \mathbf{r}_k \parallel}-\frac{3\mathcal{G}}{m_k}\left(\frac{m_0}{\parallel \mathbf{r}_k \parallel}-\frac{m_{3-k}}{\parallel \mathbf{r}_1-\mathbf{r}_2 \parallel}\right)\right]+\\
&+(\mathbf{p}_1+\mathbf{p}_2)\left[-\frac{\mathbf{p}_1\cdot\mathbf{p}_2}{m_0^3}-\frac{3\mathcal{G}}{m_0}\sum_{i=1}^{2}\frac{m_i}{\parallel \mathbf{r}_i \parallel}\right]+\frac{\mathcal{G}(\mathbf{p}_{3-k}\cdot\mathbf{n}_{12})}{2\parallel \mathbf{r}_1-\mathbf{r}_2 \parallel}\mathbf{n}_{12}+\\
&-\frac{\mathcal{G}}{2\parallel \mathbf{r}_k \parallel}\left(\sum_{i=1}^{2}\mathbf{p}_i\cdot\mathbf{n}_k\right)\mathbf{n}_k-\frac{\mathcal{G}(\mathbf{p}_{3-k}\cdot\mathbf{n}_{3-k})}{2\parallel \mathbf{r}_{3-k} \parallel}\mathbf{n}_{3-k}\Bigg\}
\end{aligned}
\tag{3.31}
$$

$$
\begin{aligned}
\dot{\mathbf{p}}_k = &-\frac{\mathcal{G}(m_0+m_k)\mu_k\mathbf{r}_k}{\parallel \mathbf{r}_k \parallel^3}-\mathcal{G}m_km_{3-k}\frac{\mathbf{r}_k-\mathbf{r}_{3-k}}{\parallel \mathbf{r}_k-\mathbf{r}_{3-k} \parallel^3}-\frac{1}{c^2}\Bigg\{\frac{\mathbf{r}_k}{\parallel \mathbf{r}_k \parallel^3}\Bigg[\frac{3\mathcal{G}m_k}{2m_0}(p_1^2+p_2^2+2\mathbf{p}_1\cdot\mathbf{p}_2)+\\
&+\frac{3\mathcal{G}p_k^2m_0}{2m_k}+\frac{7\mathcal{G}}{2}\mathbf{p}_k\cdot(\mathbf{p}_1+\mathbf{p}_2)+\frac{3\mathcal{G}(\mathbf{p}_k\cdot\mathbf{r}_k)}{2\parallel \mathbf{r}_k \parallel^2}((\mathbf{p}_1+\mathbf{p}_2)\cdot\mathbf{r}_k)+\\
&-\mathcal{G}^2m_0m_1m_2\left(\frac{1}{\parallel \mathbf{r}_{3-k} \parallel}+\frac{1}{\parallel \mathbf{r}_1-\mathbf{r}_2 \parallel}\right)-\frac{\mathcal{G}^2m_0m_k(m_0+m_k)}{\parallel \mathbf{r}_k \parallel}\Bigg]+\frac{\mathbf{r}_k-\mathbf{r}_{3-k}}{\parallel \mathbf{r}_1-\mathbf{r}_2 \parallel^3}\Bigg[\sum_{i=1}^{2}\frac{3\mathcal{G}p_i^2m_{3-i}}{2m_i}+\\
&-\frac{7\mathcal{G}\mathbf{p}_1\cdot\mathbf{p}_2}{2}-\frac{\mathcal{G}^2m_1m_2(m_1+m_2)}{\parallel \mathbf{r}_1-\mathbf{r}_2 \parallel}-\mathcal{G}^2m_0m_1m_2\left(\frac{1}{\parallel \mathbf{r}_1 \parallel}+\frac{1}{\parallel \mathbf{r}_2 \parallel}\right)-\frac{3\mathcal{G}}{2}(\mathbf{p}_1\cdot\mathbf{n}_{12})(\mathbf{p}_2\cdot\mathbf{n}_{12})\Bigg]+\\
&-\sum_{i=1}^{2}\left(\frac{\mathcal{G}(\mathbf{n}_k\cdot\mathbf{p}_i)}{2\parallel \mathbf{r}_k \parallel^2}-\frac{\mathcal{G}(\mathbf{p}_i\cdot\mathbf{n}_{12})}{2\parallel \mathbf{r}_1-\mathbf{r}_2 \parallel^2}\right)\mathbf{p}_{3-i}-\frac{\mathcal{G}(\mathbf{p}_k\cdot\mathbf{n}_k)}{\parallel \mathbf{r}_k \parallel^2}\mathbf{p}_k\Bigg\},
\end{aligned}
\tag{3.32}
$$

for $k=1,2$. Note the positions of the planets, the position of the star $\mathbf{x}_0$ is given by

$$\mathbf{x}_0 = -\frac{\sum_{i=1}^{2} m_i \mathbf{r}_i \left[1 + \frac{1}{c^2}\left(\frac{p_i^2}{2m_i^2} - \mathcal{G}\left(\frac{m_0}{2\|\mathbf{r}_i\|} + \frac{m_{3-i}}{2\|\mathbf{r}_i - \mathbf{r}_{3-i}\|}\right)\right)\right]}{\sum_{i=0}^{2} m_i + \frac{1}{c^2}\left[\sum_{i=0}^{2}\left(\frac{p_i^2}{2\mu_i} + \frac{\mathbf{p}_i \cdot \mathbf{p}_{3-i}}{2m_0} - \mathcal{G}\frac{m_i m_0}{\|\mathbf{r}_i\|} - \mathcal{G}\frac{m_i m_{3-i}}{2\|\mathbf{r}_i - \mathbf{r}_{3-i}\|}\right)\right]}. \tag{3.33}$$

3.5 Numerical integration

Numerical solution of ordinary differential equations is the most important technique in continuous time dynamics. Since most ordinary differential equations can not be solved analytically, numerical integration is the only way to obtain information about the trajectory. Many different methods have been proposed and used in an attempt to solve accurately various types of ordinary differential equations. All these discretize the differential system to produce a difference equation or map. The methods obtain different maps from the same differential equation, but they have the same aim, i.e. that the dynamics of the maps should correspond closely to the dynamics of the differential equations.

Let the phase space Ω be a domain in $\mathbb{R}^{2n}$. The smooth Hamiltonian function $H \in C^1(\Omega)$ gives rise to the Hamiltonian system of ODE's

$$q_i = \frac{\partial H}{\partial p_i}, \qquad p_i = -\frac{\partial H}{\partial q_i} \tag{3.34}$$

for $i = 1, ..., n$, or, in a more compact form,

$$\dot{\mathbf{x}} = J H_{\mathbf{x}}(\mathbf{x}), \qquad J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \tag{3.35}$$

where $\mathbf{x} = [\mathbf{q}^T, \mathbf{p}^T]^T$, I is the identity matrix and where the subscript $\mathbf{x}$ denotes differentiation.

Let $h > 0$ be fixed. A (single-step) numerical scheme to solve such a system consists of a function $\boldsymbol{\psi}_{H,h}\colon \Omega \to \Omega$ depending smoothly on the step-size h and on the Hamiltonian H. Given an initial condition $(\mathbf{p}_0, \mathbf{q}_0)$, the approximate solution at time mh defined as $(\mathbf{q}_m, \mathbf{p}_m)$ can be computed iteratively by

$$(\mathbf{q}_m, \mathbf{p}_m) = \boldsymbol{\psi}_{H,h}(\mathbf{q}_{m-1}, \mathbf{p}_{m-1}). \tag{3.36}$$

Now, let $\boldsymbol{\phi}_t$ be the flow of (3.35). To approximate $\boldsymbol{\phi}_t(\mathbf{x})$ in the interval $[0, t]$, at least heuristically, the idea is to divide $[0, t]$ into N parts and to iterate $\boldsymbol{\psi}_{H,h}$ with a step size $h = t/N$ for N times.

The method $\boldsymbol{\psi}_{H,h}$ is said to be of order $r \in \mathbb{N}$ if, as $h \to 0$,

$$\| \boldsymbol{\phi}_h(\mathbf{x}) - \boldsymbol{\psi}_{H,h}(\mathbf{x}) \| = O(h^{r+1}) \tag{3.37}$$

for any $\mathbf{x} \in \Omega$. It is simple to prove that the total accumulated error is order $O(h^r)$. Indeed, we would perform $O(h^{-1})$ computations, so

$$\| \boldsymbol{\phi}_h(\mathbf{x}) - \boldsymbol{\psi}_{H,h}^N(\mathbf{x}) \| = O(h^{r+1}) \cdot O(h^{-1}) = O(h^r). \tag{3.38}$$

One of the most used methods is the *classical Runge-Kutta method*, which I present briefly. For generality, we consider the non-autonomous initial value problem

$$\dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y}), \qquad \mathbf{y}(t_0) = \mathbf{y}_0, \tag{3.39}$$

where $\mathbf{y}$ is an unknown function (scalar or vector) of time t which we would like to approximate. At the initial time t_0 the corresponding $\mathbf{y}$-value is $\mathbf{y}_0$. The function $\mathbf{f}$ and the data t_0, $\mathbf{y}_0$ are given.

Pick a step-size $h > 0$. The classical Runge-Kutta method of order 4 (*RK4*) is defined as

$$\mathbf{y}_{m+1} = \mathbf{y}_m + \frac{1}{6}h\left(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4\right) \tag{3.40}$$

$$t_{m+1} = t_m + h \tag{3.41}$$

for $m = 0, 1, 2, 3, ...$, where

$$\begin{aligned}
\mathbf{k}_1 &= \mathbf{f}(t_m, \mathbf{y}_m), \\
\mathbf{k}_2 &= \mathbf{f}\left(t_m + \frac{1}{2}h, \mathbf{y}_m + \frac{h}{2}\mathbf{k}_1\right), \\
\mathbf{k}_3 &= \mathbf{f}\left(t_m + \frac{1}{2}h, \mathbf{y}_m + \frac{h}{2}\mathbf{k}_2\right), \\
\mathbf{k}_4 &= \mathbf{f}(t_m + h, \mathbf{y}_m + h\mathbf{k}_3).
\end{aligned} \tag{3.42}$$

Here $\mathbf{y}_{m+1}$ is the Runge-Kutta approximation of $\mathbf{y}(t_{m+1})$, and the value $\mathbf{y}_{m+1}$ is determined by the previous value $\mathbf{y}_m$ plus the weighted average of four increments, where each increment is the product of the size of the interval h and an estimated slope specified by function $\mathbf{f}$ on the right-hand side of the differential equation. It is simple to prove that the classical Runge-Kutta method is a fourth-order method, meaning, as we have seen, that the local truncation error is on the order of $O(h^5)$, while the total accumulated error is order $O(h^4)$.

In simulations of conservative systems, the energy H is usually monitored as a check on the calculation. As we have seen, the energy should be a constant of motion and should not change. However, in numerical simulations the energy might fluctuate on a short time scale and increase or decrease on a very long time scale due to numerical integration artifacts that arise with the use of a finite time step h. For this reason, in the following we use the conservation of the total energy H as a useful checks of the accuracy of the numerical simulations. The gradual change in the total energy of a closed system over time is called *energy drift*. Moreover, we can also use the conservation of the total angular momentum to have another check of the accuracy of the numerical simulations.

3.6 Application to some extrasolar systems

We consider some extrasolar systems consisting of a central star and two planets rotating around it, such that the two planets are not in resonance, in order to use in the following the principle of the average. We are interested in the evolution of the semi-major axes and of the eccentricities of the two planets, and in particular we are interested to see if there are significant differences between the Newtonian and the relativistic case.

The initial orbital elements that characterize the selected systems, and the parameters of their parents star, come from the Jean Schneider Encyclopedia of Extrasolar Planets[3].

We proceed as follows. First, from the initial orbital elements which describe the extrasolar system, we calculate the canonical coordinates in a heliocentric reference system, i.e. the position $\mathbf{r}_i = (r_{x,i}, r_{y,i}, r_{z,i})$ and the momentum $\mathbf{p}_i = (p_{x,i}, p_{y,i}, p_{z,i})$ of the i-planet, for $i = 1, 2$. Then we integrate the equation of motions (3.27) and (3.31)-(3.32) using the classical Runge-Kutta method

[3] http://exoplanet.eu

(3.40) and finally, at each step or after a fixed period, we calculate the new orbital elements in order to obtain the evolution of the semi-major axes and of the eccentricities of the two planets.

To integrate the equations of motion, I implemented a program in the C language with a 15 digit precisions.

To have a check on the quality of the numerical integration, in both cases we calculate at each step also the drift of the total energy of the system, given by the formula

$$\Delta H(t) = \frac{H(0) - H(t)}{H(0)}. \tag{3.43}$$

Obviously, the integration is to be considered good if this quantity remains small. To get good results, we observe that the integration step must be very small if the period of revolution of the inner planet around the star is small.

To switch from canonical coordinates to the orbital elements, and vice versa, we have to introduce the concept of the *osculating orbit.* The osculating orbit of an object in space at a given moment in time is the gravitational Kepler orbit (i.e. the ellipse in heliocentric coordinate) that it would have about its central body if perturbations were not present, i.e. the orbit that coincides with the current orbital state vectors (position and velocity). An osculating orbit and the object's position upon it can be fully described by the six standard Keplerian orbital elements (osculating elements), which are easy to calculate as long as one knows the object's position and velocity relative to the central body.

The osculating elements would remain constant in the absence of perturbations, but the real astronomical orbits experience perturbations that cause the osculating elements to evolve. Thus, if we assume that the variations of the osculating orbits are slow and regular, we can describe the motion of the planets thinking of osculating ellipses whose parameters vary slowly over time. Indeed, suppose that at a given instant we can "turn off" the interactions between the planets and the relativistic corrections, so that the Hamiltonian of our system is reduced to H_0. From this moment, each planet moves on a Keplerian orbit with orbital elements $a_j(t_0), e_j(t_0), ...$: the ellipse covered by the planet from that moment is called the osculating orbits at time t_0. Obviously, the real motion of the planet will be very different from being an ellipse, but if we repeat the procedure at the time $t_1 = t_0 + \tau$, where τ is small time increment, we obtain new orbital elements $a_j(t_1), e_j(t_1), ...$ which describe a new osculating orbit.

The idea is to use the osculating orbits to make these changes of coordinates, i.e. we suppose that the orbit of each planets is equal to its osculating orbit. With this assumption, to make these changes of coordinates we can simple use the relationships between the orbital elements and the canonical coordinates given in the case of the two-body problem. In particular, the formulæ to change from orbital elements to canonical coordinates $(\mathbf{r}, \mathbf{p})$ are given in chapter 1. Vice-versa, to change from canonical coordinates to orbital elements, we consider for example the motion of one of the two planets, having mass m, around the star of mass m_0. We introduce the angular momentum $\mathbf{M} = (M_1, M_2, M_3)$ and the energy $\mathcal{E}$ of the system:

$$\begin{aligned} &\mathbf{M} = \mathbf{r} \times \mu\dot{\mathbf{r}}, \qquad \mathcal{M} = \sqrt{M_1^2 + M_2^2 + M_3^2}, \\ &\mathcal{E} = \frac{1}{2}\mu\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} - \frac{\mathcal{G}(m_0 + m)}{\| \mathbf{r} \|}, \end{aligned} \tag{3.44}$$

where μ is the reduced mass (3.24).

Using the expressions from (1.33) to (1.36), it is simple to prove that the semi-major axis a and the eccentricity e are given by the following formulæ

$$\begin{aligned} a &= -\frac{\mathcal{G}(m_0+m)}{2\mathcal{E}}, \\ e &= \sqrt{1-\frac{\mathcal{M}^2}{a\mathcal{G}(m_0+m)}}, \end{aligned} \tag{3.45}$$

and that the inclination i is given by

$$\cos i = \frac{M_3}{\mathcal{M}}, \qquad \sin i = \frac{\sqrt{M_1^2+M_2^2}}{\mathcal{M}}. \tag{3.46}$$

If $i \neq 0$ or π, the longitude of the ascending node Ω is given by

$$\cos\Omega = -\frac{M_2}{\mathcal{M}\sin i}, \qquad \sin\Omega = \frac{M_1}{\mathcal{M}\sin i}. \tag{3.47}$$

If $i = 0$ or π, we assume for the following that $\Omega = 0$.

If $e \neq 0$, the eccentric anomaly E is given by

$$\cos E = \frac{a-r}{e\,a}, \qquad \sin E = \frac{\mathbf{r}\cdot\dot{\mathbf{r}}}{e\sqrt{a\mathcal{G}(m+m_0)}}, \tag{3.48}$$

where the second relation is obtained by calculating $\mathbf{r}\cdot\dot{\mathbf{r}}$ and where r is given by

$$r = \sqrt{r_x^2+r_y^2+r_z^2}. \tag{3.49}$$

The true anomaly ν is given by

$$\cos\nu = \frac{e-\cos E}{1-e\cos E}, \qquad \sin\nu = \frac{\sqrt{1-e^2}\sin E}{1-e\cos E}. \tag{3.50}$$

If $e \neq 0$, to define the argument of pericenter ω we need to pass to the orbital plane

$$\begin{bmatrix} q_1 \\ q_2 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\Omega & \sin\Omega & 0 \\ -\sin\Omega\cos i & \cos\Omega\cos i & \sin i \\ \sin\Omega\sin i & -\cos\Omega\sin i & \cos i \end{bmatrix} \begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix}. \tag{3.51}$$

It is simple to prove that

$$\omega = u - \nu \tag{3.52}$$

where u is

$$\cos u = \frac{q_1}{\sqrt{q_1^2+q_2^2}}, \qquad \sin u = \frac{q_2}{\sqrt{q_1^2+q_2^2}}. \tag{3.53}$$

Finally, the relationship between the momentum and the velocity is given in the Newtonian case by

$$\mathbf{p} = \frac{m_0 m}{m_0+m}\dot{\mathbf{r}}, \tag{3.54}$$

and in the relativistic case (see (3.63)) by

$$\begin{aligned}\mathbf{p} =&\frac{m_0 m}{m_0+m}\dot{\mathbf{r}} + \frac{1}{c^2}\Bigg\{\frac{m_0 m}{m_0+m}\Bigg[\left(\frac{1}{2m_0^3}+\frac{1}{2m^3}\right)(\dot{\mathbf{r}}\cdot\dot{\mathbf{r}})\dot{\mathbf{r}}+\\ &+\frac{\mathcal{G}}{\|\mathbf{r}\|}\left(\frac{3m}{m_0}+\frac{3m_0}{m}+7\right)\dot{\mathbf{r}}+\frac{\mathcal{G}}{\|\mathbf{r}\|^2}(\mathbf{r}\cdot\dot{\mathbf{r}})\dot{\mathbf{r}}\Bigg]\Bigg\}.\end{aligned} \tag{3.55}$$

In most cases of extrasolar systems, the inclinations and the initial mean anomaly of the two planets aren't known. In these cases, as regards the inclination we decide to put $i = 0$ for both the planets. With this choice, it is simple to prove that $r_{z,1} = r_{z,2} = 0$ and $p_{z,1} = p_{z,2} = 0$ for any t, i.e. the system is coplanar for any t, so the dynamic is greatly simplified. The choice of the initial mean anomaly M_0 does not affect much on the dynamics of the system over a long time, so it can be chosen in an arbitrary manner or, when it is possible, with the formula:

$$M_0 = \sqrt{\mathcal{G}(m_0+m)}a^{-3/2}\,t. \tag{3.56}$$

The relation between the mean anomaly and the eccentric anomaly is given by the Kepler equation (1.24). To solve the Kepler equation, it can be also useful to use numerical method, as the *bisection method* or the *Newton's method*, which I show briefly. Suppose we have to find an approximate solution of the equation

$$f(x) = 0, \tag{3.57}$$

where f is a real-valued function. If f is a differentiable function such that $f(a)\cdot f(b) \le 0$ and such that $f'(x) \ne 0$ for any $x \in [a,b]$, then an approximate solution of equation (3.57) can be constructed using the following algorithm

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad \forall n \ge 0 \tag{3.58}$$

where x_0 is an arbitrary initial valued in $[a,b]$[4].

For the following, we adopt the astronomical unit AU (i.e. the average Earth-Sun distance) as unit of length, the year (yr) as time unit and the Jupiter mass (M_J) as mass unit:

$$\begin{aligned}1\,AU &= 149\,597\,871\,000\,\mathrm{m},\\ 1\,yr &= 31\,560\,000\,\mathrm{s},\\ 1\,M_J &= 1.8986\times 10^{27}\,\mathrm{Kg}.\end{aligned} \tag{3.60}$$

With these units, the gravitational constant $\mathcal{G}$ and the speed of light in empty space c are

$$\begin{aligned}\mathcal{G} =&3.76242\times 10^{-2}\,\frac{AU^3}{M_J\cdot yr^2},\\ c =&63\,197.79\,\frac{AU}{yr}.\end{aligned} \tag{3.61}$$

[4]**Theorem** *Let $f \in C^2([a,b])$ a real-valued function such that $f(a)\cdot f(b) \le 0$. If f satisfies the conditions*

$$0 < d \le f'(x), \qquad 0 \le f''(x) \le M \tag{3.59}$$

for any $x \in [a,b]$, then there exist exactly one fixed point $\tilde{x}$ such that $f(\tilde{x}) = 0$ and the sequence $\{x_n\}_{n=0}^{\infty}$ defined by the process (3.58) with an arbitrary starting point $x_0 \in [a,b]$, converges to $\tilde{x}$.

The extrasolar systems that we have studied are the following:

- **HD 190360**

  Stellar mass= 1089.2544 M_J
  Integration step size=5×10^{-5} (yr)

	HD 190360 b	HD 190360 c
mass	$1.502 \pm 0.13\, M_J$	$0.057 \pm 0.0015\, M_J$
a	$3.92 \pm 0.02\, AU$	$0.128 \pm 0.002\, AU$
e	0.36 ± 0.03	0.01 ± 0.1
ω	12.4 ± 9.3 deg	153.7 ± 32 deg
(Chosen) M_0	5 deg	20 deg
Orbital period	2891 ± 85 day	17.1 ± 0.0015 day

- **HD 11964**

  Stellar mass= 1178.28 M_J
  Integration step size=1×10^{-4} (yr)

	HD 11964 b	HD 11964 c
mass	$0.622 \pm 0.056\, M_J$	$0.079 \pm 0.01\, M_J$
a	$3.16 \pm 0.19\, AU$	$0.229 \pm 0.013\, AU$
e	0.041 ± 0.017	0.3 ± 0.17
ω	155 deg	102 deg
(Chosen) M_0	5 deg	180 deg
Orbital period	1945 ± 26 day	37.91 day

- **HD 169830**

  Stellar mass= 1466.304 M_J
  Integration step size=2.5×10^{-4} (yr)

	HD 169830 b	HD 169830 c
mass	$2.88\, M_J$	$4.04\, M_J$
a	$0.81\, AU$	$3.6\, AU$
e	0.31 ± 0.01	0.33 ± 0.02
ω	148 ± 2 deg	252 ± 8 deg
(Chosen) M_0	5 deg	5 deg
Orbital period	225.62 ± 0.22 day	2102 ± 264 day

- **HD 12661**

Stellar mass= $1120.6752\, M_J$
Integration step size=2.5×10^{-4} (yr)

	HD 12661 b	HD 12661 c
mass	$2.3\, M_J$	$1.57\, M_J$
a	$0.83\, AU$	$2.56\, AU$
e	0.377 ± 0.008	0.031 ± 0.022
ω	296 ± 1.5 deg	165 deg
(Chosen) M_0	5 deg	5 deg
Orbital period	263.2 ± 1.2 day	1708 ± 14 day

- **BD 082823**

Stellar mass= $775.0464 \pm 73.3152\, M_J$
Integration step size=1×10^{-5} (yr)

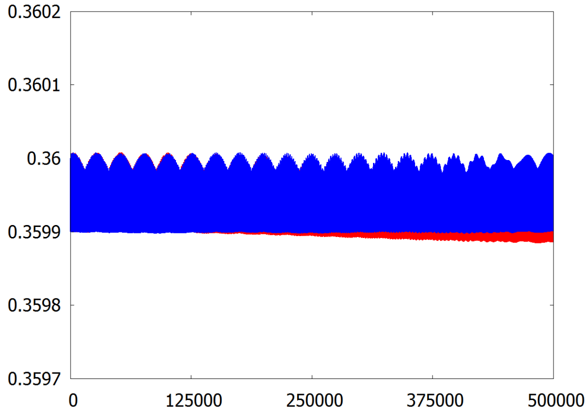
	BD 082823 b	BD 082823 c
mass	$0.045 \pm 0.007\, M_J$	$0.33 \pm 0.03\, M_J$
a	$0.056 \pm 0.0002\, AU$	$0.68 \pm 0.02\, AU$
e	0.15 ± 0.15	0.19 ± 0.09
ω	30 ± 100 deg	127 ± 21 deg
(Chosen) M_0	5 deg	15 deg
Orbital period	5.6 ± 0.02 day	237.6 ± 1.5 day

- **HIP 5158**

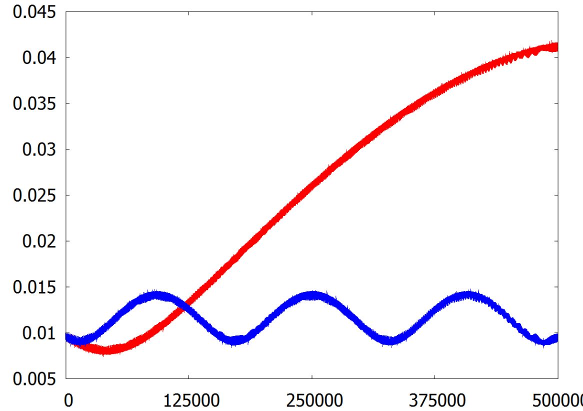
Stellar mass= $816.9408 \pm 21.995\, M_J$
Integration step size=2.5×10^{-4} (yr)

	HIP 5158 b	HIP 5158 c
mass	$1.44 \pm 0.14\, M_J$	$15.04 \pm 10.55\, M_J$
a	$0.89 \pm 0.14\, AU$	$7.7 \pm 1.88\, AU$
e	0.54 ± 0.04	0.14 ± 0.1
ω	70 ± 4 deg	142 ± 75 deg
(Chosen) M_0	5 deg	50 deg
Orbital period	345.63 ± 1.99 day	9081 ± 318 day

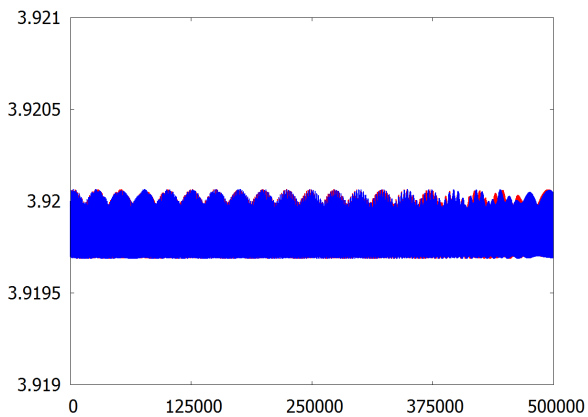
In the following figures, the red curves are for the point-mass planets interacting mutually through Newtonian forces, while the blue curves are for generalized model of motion, including the general relativistic corrections. The execution times of the programs used for the numerical integration of the equations of motion are given in Table 6.1.



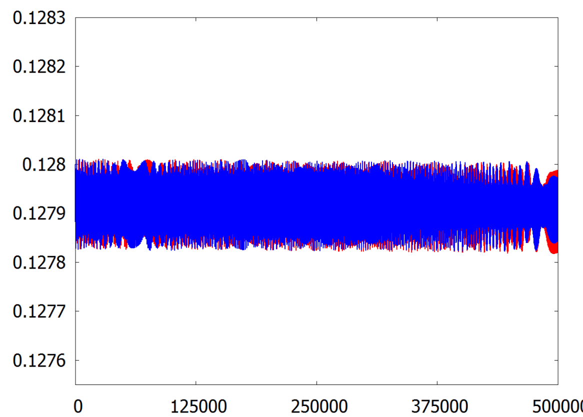
(a) Eccentricity planet HD 190360 b



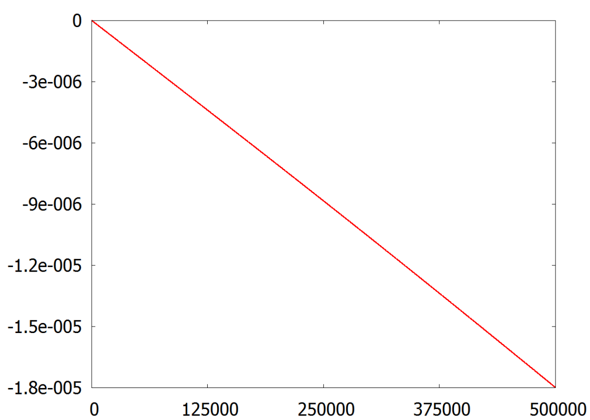
(b) Eccentricity planet HD 190360 c



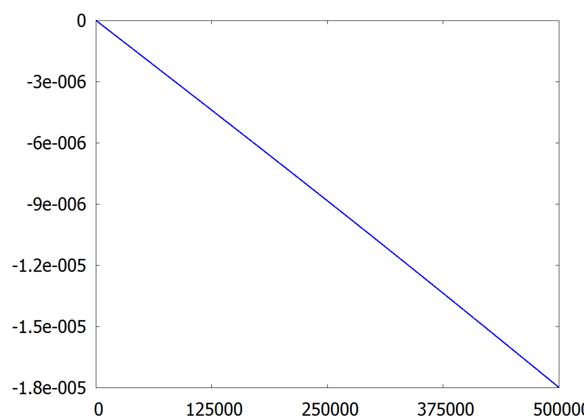
(c) Semi-major axis planet HD 190360 b



(d) Semi-major axis planet HD 190360 c

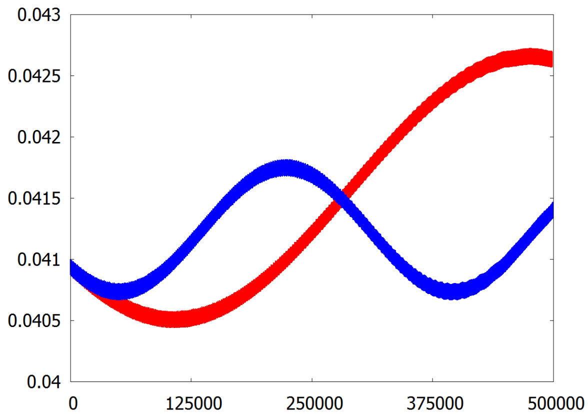


(e) The relative error of the Hamiltonian in the Newtonian case - order: -1.8×10^{-5}

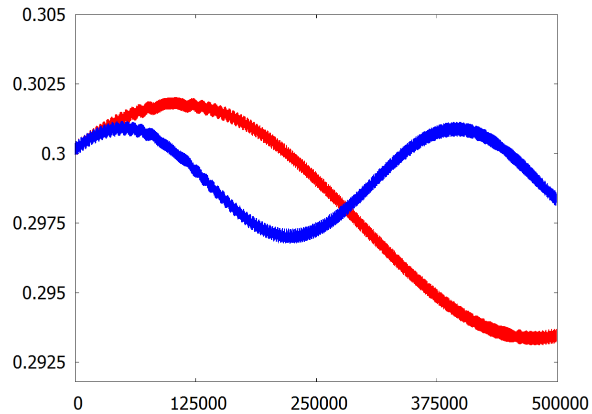


(f) The relative error of the Hamiltonian in the relativistic case - order: -1.8×10^{-5}

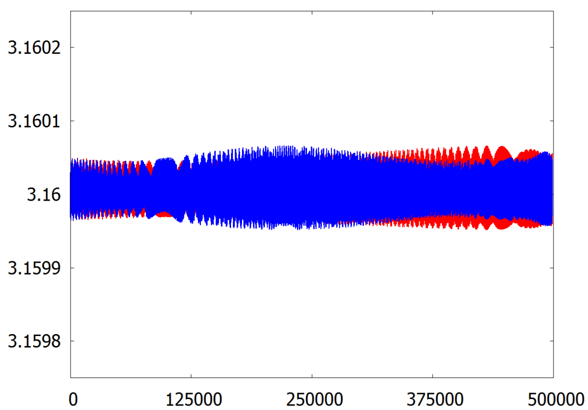
Figure 3.1: The long-term secular evolution of eccentricities and semi-major axis of the planets of the extrasolar system **HD 190360**.



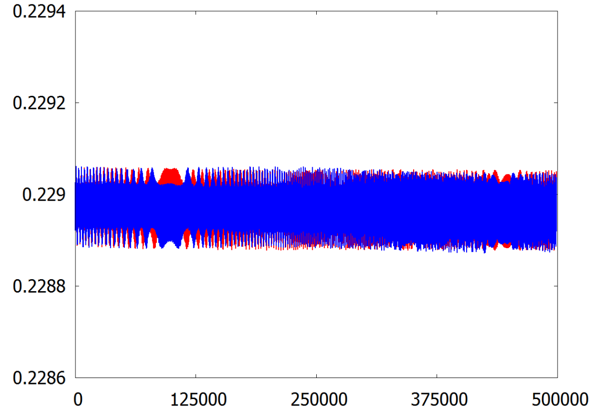
(a) Eccentricity planet HD 11964 b



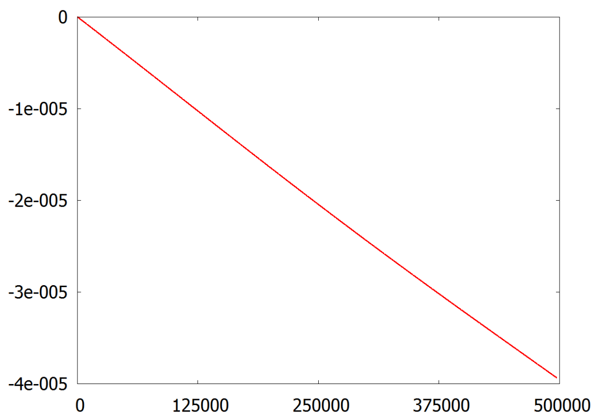
(b) Eccentricity planet HD 11964 c



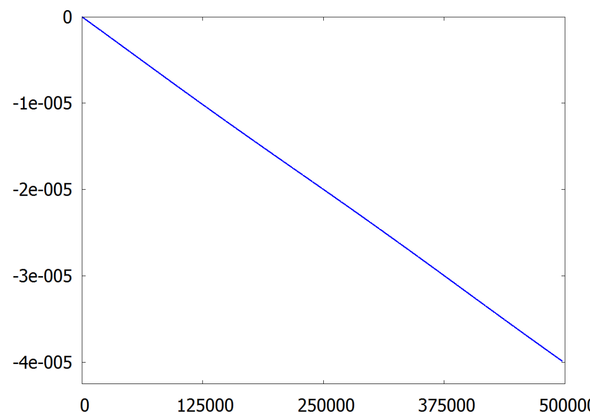
(c) Semi-major axis planet HD 11964 b



(d) Semi-major axis planet HD 11964 c

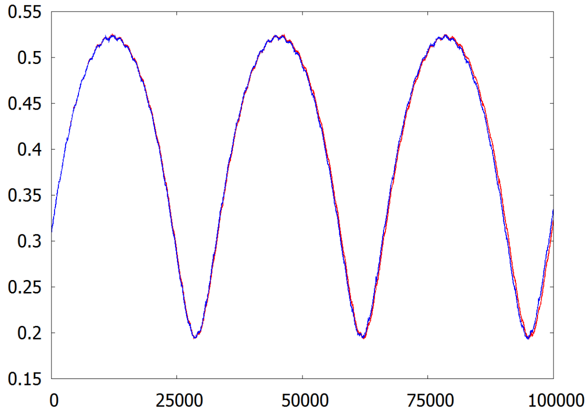


(e) The relative error of the Hamiltonian in the Newtonian case - order: -4×10^{-5}

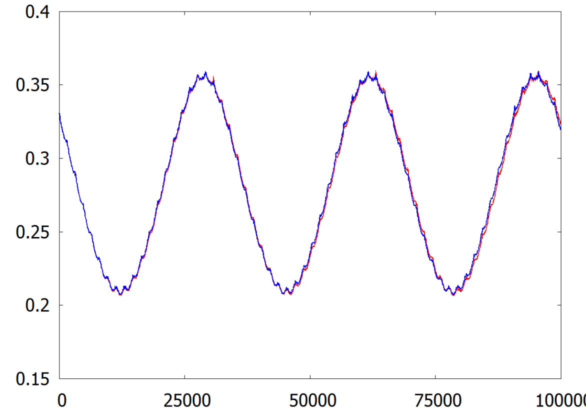


(f) The relative error of the Hamiltonian in the relativistic case - order: -4×10^{-5}

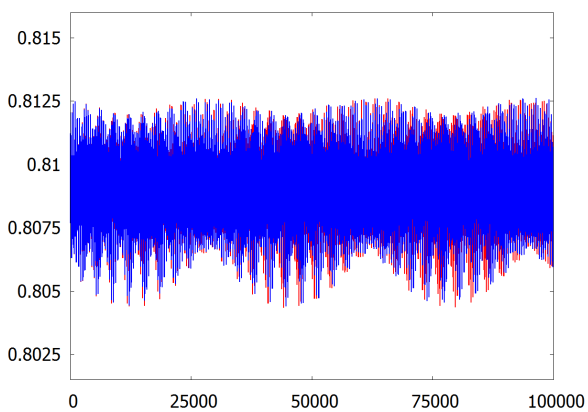
Figure 3.2: The long-term secular evolution of eccentricities and semi-major axis of the planets of the extrasolar system **HD 11964**.



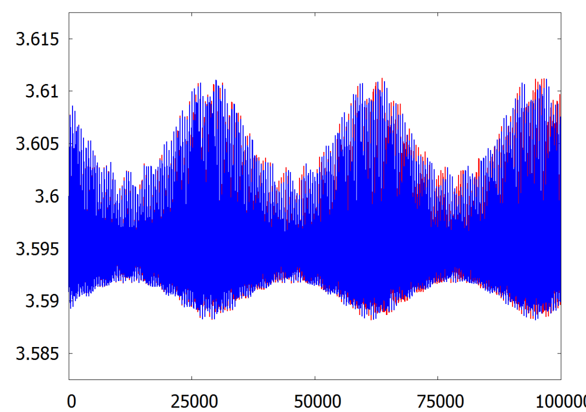
(a) Eccentricity planet HD 169830 b



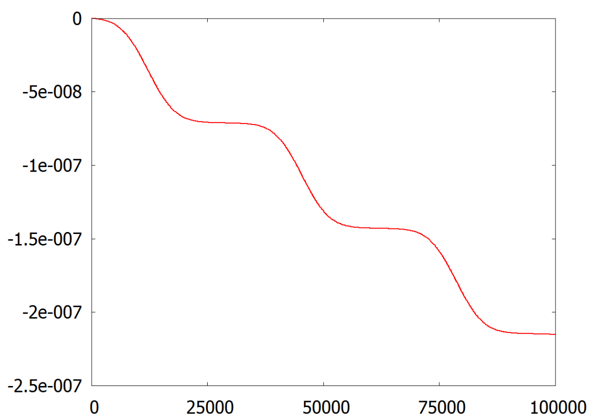
(b) Eccentricity planet HD 169830 c



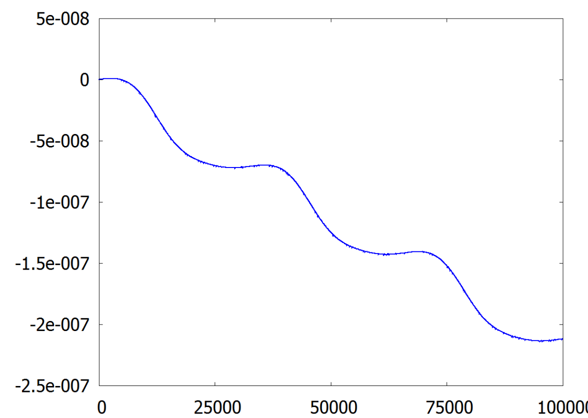
(c) Semi-major axis planet HD 169830 b



(d) Semi-major axis planet HD 169830 c

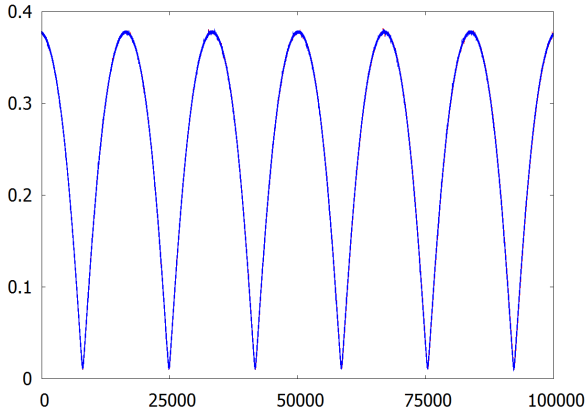


(e) The relative error of the Hamiltonian in the Newtonian case - order: -2.5×10^{-7}

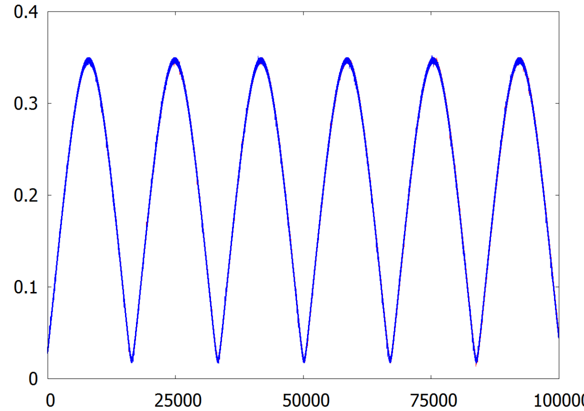


(f) The relative error of the Hamiltonian in the relativistic case - order: -2.5×10^{-7}

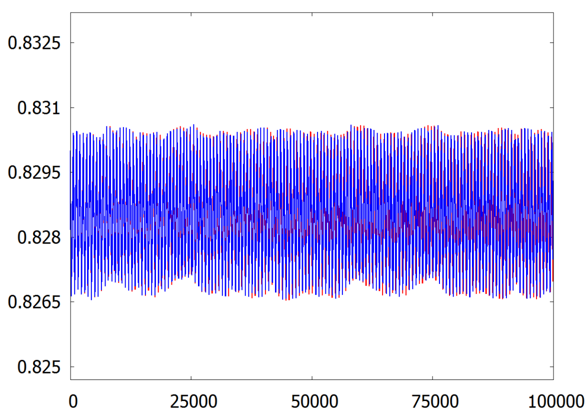
Figure 3.3: The long-term secular evolution of eccentricities and semi-major axis of the planets of the extrasolar system **HD 169830**.



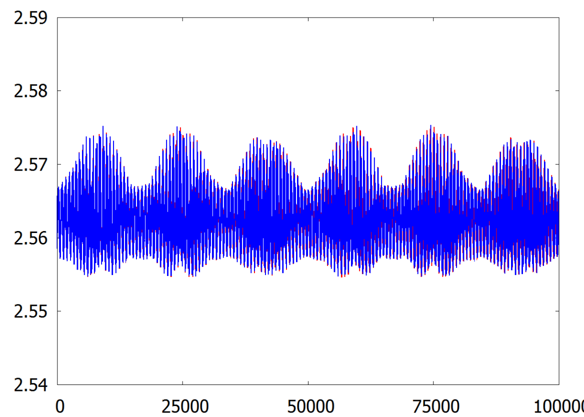
(a) Eccentricity planet HD 12661 b



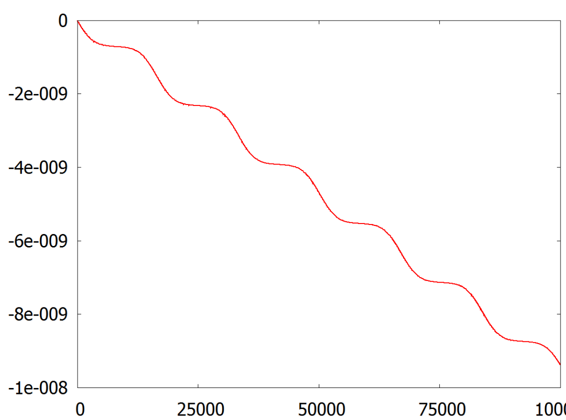
(b) Eccentricity planet HD 12661 c



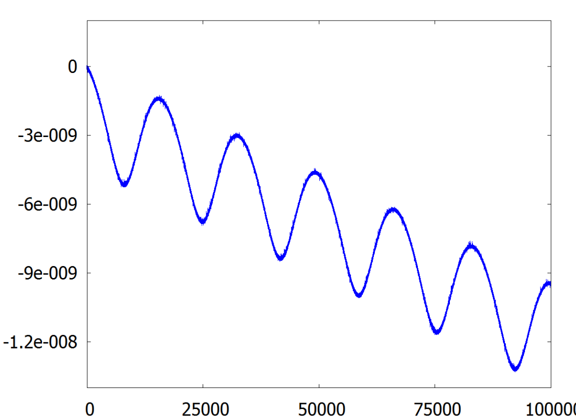
(c) Semi-major axis planet HD 12661 b



(d) Semi-major axis planet HD 12661 c

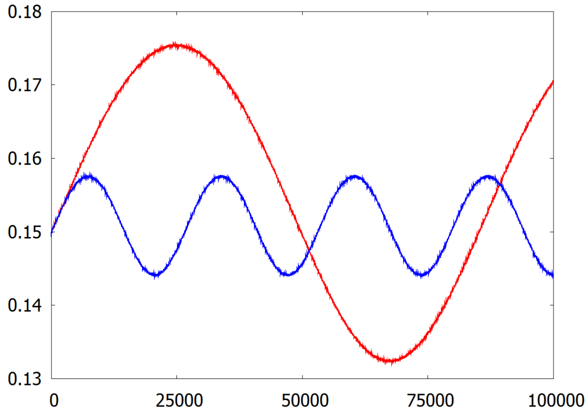


(e) The relative error of the Hamiltonian in the Newtonian case - order: -1×10^{-8}

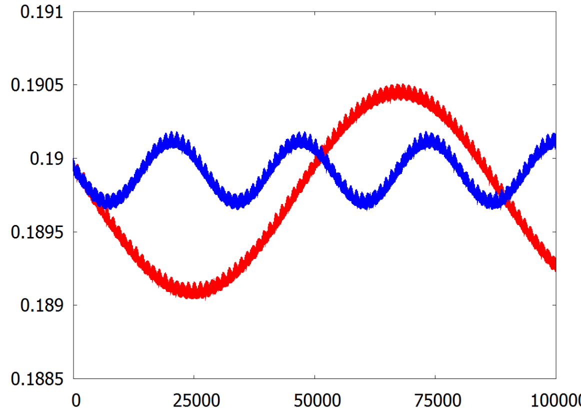


(f) The relative error of the Hamiltonian in the relativistic case - order: -1.4×10^{-8}

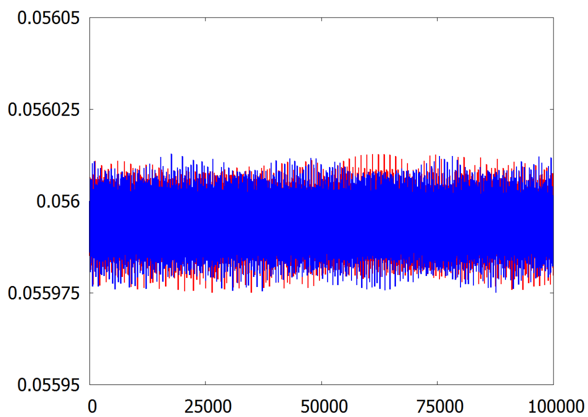
Figure 3.4: The long-term secular evolution of eccentricities and semi-major axis of the planets of the extrasolar system **HD 12661**.



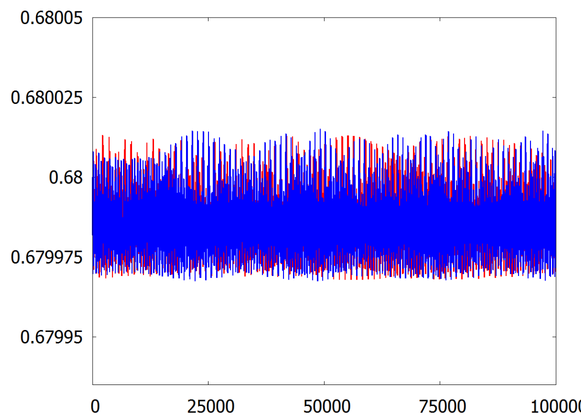
(a) Eccentricity planet BD 082823 b



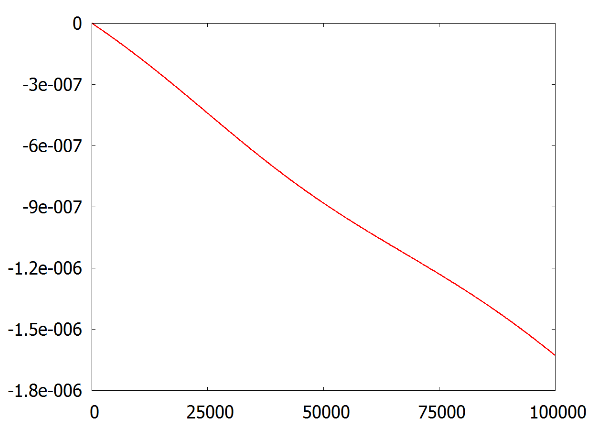
(b) Eccentricity planet BD 082823 c



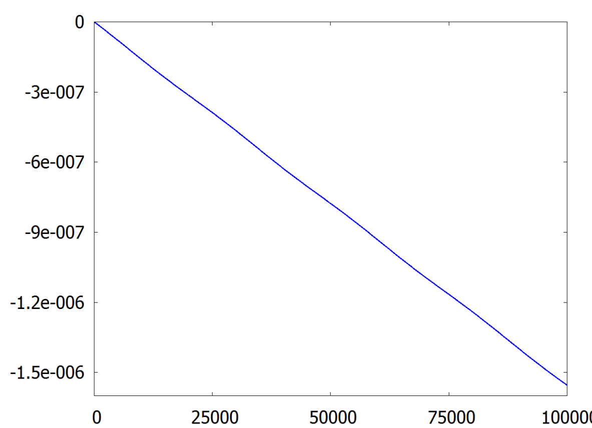
(c) Semi-major axis planet BD 082823 b



(d) Semi-major axis planet BD 082823 c

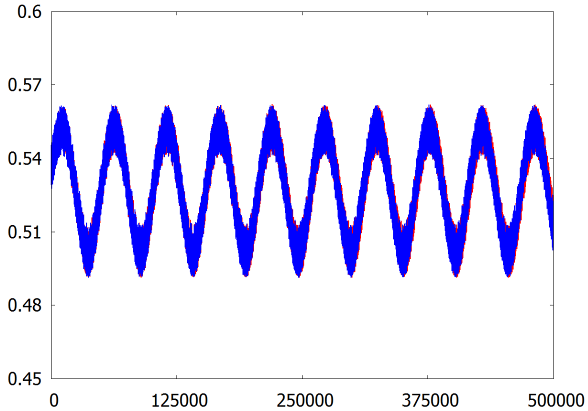


(e) The relative error of the Hamiltonian in the Newtonian case - order: -1.7×10^{-6}

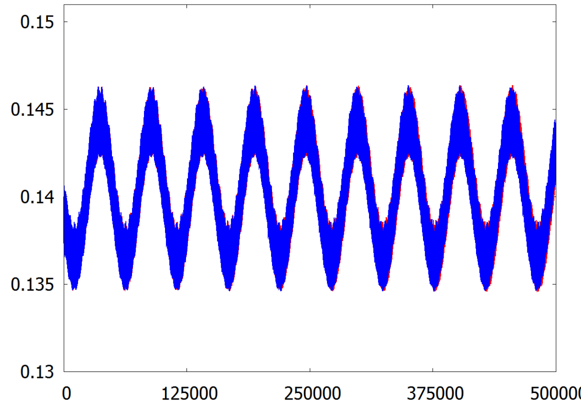


(f) The relative error of the Hamiltonian in the relativistic case - order: -1.6×10^{-6}

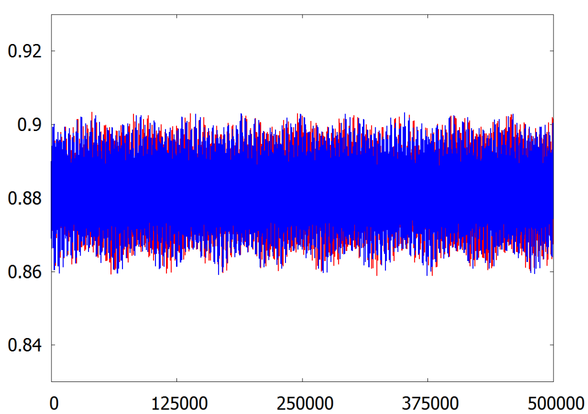
Figure 3.5: The long-term secular evolution of eccentricities and semi-major axis of the planets of the extrasolar system **BD 082823**.



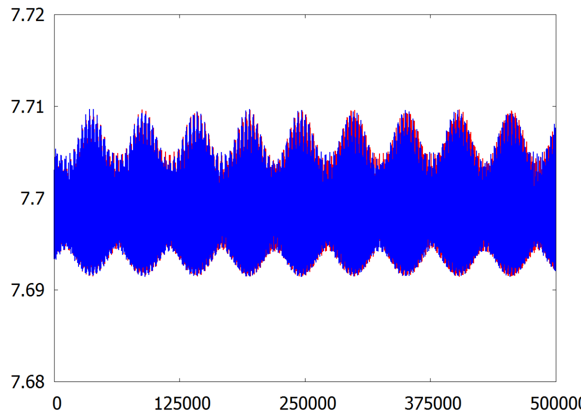
(a) Eccentricity planet HIP 5158 b



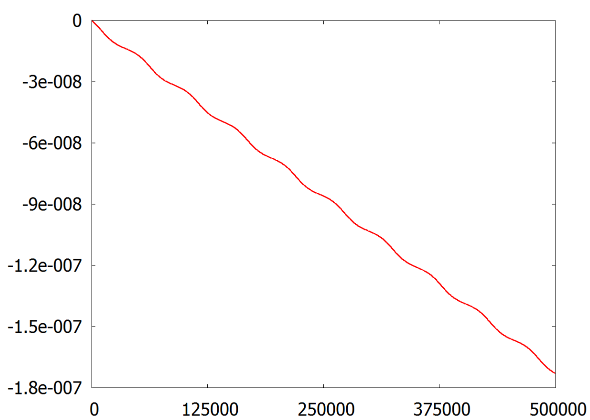
(b) Eccentricity planet HIP 5158 c



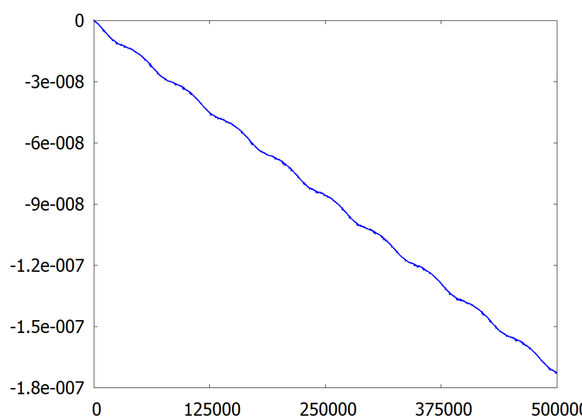
(c) Semi-major axis planet HIP 5158 b



(d) Semi-major axis planet HIP 5158 c



(e) The relative error of the Hamiltonian in the Newtonian case - order: -1.8×10^{-7}



(f) The relative error of the Hamiltonian in the relativistic case - order: -1.8×10^{-7}

Figure 3.6: The long-term secular evolution of eccentricities and semi-major axis of the planets of the extrasolar system **HIP 5158**.

As one can see, in some cases the difference between predictions of General Relativity and by the classical model are significant, in some other cases both theories give practically the same outcome. In particular we can do the following considerations. As expected, if both planets are very far from the central star (as for example in the cases HD 169830, HD 12661 and HIP 5158), then the relativistic corrections are insignificant. Instead, if the inner planet is close to the star (as for example in the cases HD 190360, HD 11964 and BD 082823), then the relativistic corrections are very important, in particular for the inner-most planet.

As regards the semi-major axis, it is important to note that its value always oscillates around the initial value, i.e. the average value remains constant[5]. This fact, together with the average principle, will be used in following chapters to simplify the two Hamiltonians.

In all systems, it is interesting to note the periodic evolution in eccentricities which shows an almost perfect coupling between them, i.e. both elements oscillates with the same period. This is a simple consequence of the conservation of the total angular momentum $\mathbf{M}$ of the system, which is equal to the summation of the angular momentum of the planets in the heliocentric reference system (see (3.18) - we remember that $\mathbf{p}_0 = \mathbf{0}$). In particular, it is simple to prove that (see section 5.2)

$$\parallel \mathbf{M} \parallel = \mu_1 \sqrt{\mathcal{G}(m_0 + m_1)a_1} \sqrt{1 - e_1^2} \cos i_1 + \mu_2 \sqrt{\mathcal{G}(m_0 + m_2)a_2} \sqrt{1 - e_2^2} \cos i_2.$$

The conservation of $\parallel \mathbf{M} \parallel$, the fact that $i_1 = i_2 = 0$ and the fact that a_1 and a_2 are almost constants up to the second order in the masses (they oscillate slightly around an average value) explain the coupling between the eccentricities.

Looking at the element of the examined systems, we may conclude that the corrections become very important for systems with the innermost planet close to the star, with other body relatively distant. The results show in fact that the relativistic effects can accumulate over time to induce substantial changes in the dynamics. In particular, we note that where the relativistic effects are important, they seem to provide "stability" to the system. In fact, if we consider the case of the eccentricity, we note that in the relativistic case it remains closer to the average value than in the Newtonian case. Furthermore, the oscillation frequency of the eccentricity in the relativistic case is greater than that in the classical case.

These results are in agreement with those obtained by Laskar (*Laskar (2008)*) in the case of the Solar system. Quoting Laskar: *"I have thus repeated the previous simulations in absence of relativity, expecting to find a more stable system. But the result was the opposite, ...". "Indeed, as we have demonstrated here, the contribution of general relativity changes in a considerable manner the behavior of the Solar System dynamics." "The contribution of GR is thus essential in order to ensure the relative stability of Mercury."*.

Finally, we can do some considerations on the numerical method that we have used. As already mentioned previously, to obtain good results, and in particular to have a very small drift of energy, it is necessary to use a very small time step. In this case, the numerical integration becomes CPU consuming and the time required to integrate a system becomes very long (see Table 6.1). For this reason, in the next chapters, we look for a "semi-analytical" integrations of the Hamiltonian equations, using the tools provided by Hamiltonian system. Before doing this, however, we must try to simplify the relativistic equations.

[5] It can be proven that if there are no resonances between the mean motion frequencies of the planets, then the mean values of the semi-major axis are constant up to the second order in the masses. For more details, see *Theorem of Poisson* in section 6.1.

3.7 Simplification of the relativistic Hamiltonian

As already mentioned, we want to simplify the relativistic problem skipping the relativistic corrections due to the mutual interactions between the two planetary masses, i.e. we assume that the mutual interaction between the two planets is a Newtonian interaction. We will see that this assumption leads to a substantial simplification of the relativistic Hamiltonian. At the same time, the dynamic obtained by this simplified Hamiltonian is very similar to that described by the real one, at least numerically in the systems that we have considered (see *Brumberg (1991)* and *Migaszewski and Goździewski (2008)* for more details).

Before proceeding, we need to discuss briefly the nature of the canonical momenta (3.3) appearing in the Hamiltonian (3.2). As we have seen, Post-Newtonian Hamiltonians are derived from Lagrangian (2.73) in which the generalized velocities appear not only in the kinetic terms of the Newtonian portion of the Lagrangian, but also in the relativistic perturbation. As a consequence, after switching to the Hamiltonian formulation via Legendre transformation procedure, the post-Newtonian canonical Hamiltonian momenta will differ from the Newtonian ones by terms of order $1/c^2$. This discrepancy will carry over to any subsequent canonical transformation, including the introduction of Delaunay elements. However, in the present work we are concerned with the secular variations of orbital elements for which the discrepancy above is of little consequence (see *Richardson* & *Kelly (1988)* and *Heimberger* & *al. (1990)* for a detailed analysis of the connection between Newtonian and Post-Newtonian Delaunay orbital elements). For this reason, in the following we will consider the relativistic momenta as Newtonian ones.

The simplification of the problem can be justified as follows. As we have seen, the Hamiltonian that describes the motion of the bodies can be written as the sum of two parts: a main part, which takes account of the Newtonian interactions between the planets, and a perturbation part, which takes into account the relativistic corrections to the Newtonian gravity. In this second part, we can notice that the relativistic corrections due to the mutual interactions between the star and the two planets are greater than those due to mutual interaction between the two planets, because $m_1, m_2 \ll m_0$ and because the mutual distance $\parallel \mathbf{r}_1 - \mathbf{r}_2 \parallel$ never becomes small. For this reason, we have decided to simplify the problem assuming that that the mutual interactions between the star and the two planets are of relativistic type (i.e. we consider the relativistic corrections to the Newtonian gravity) and that the mutual interaction between the two planets is only of Newtonian type (i.e. we skip the relativistic corrections caused by the two planetary masses). In other words, we assume that the motion of three bodies is not affected by the relativistic corrections to the Newtonian gravity due to the mutual interactions between the two planets.

It is important to emphasize that this problem is a simplification of a real problem, because the relativistic corrections to the Newtonian gravity due to the mutual interactions between the two planets are present even if they are very small.

To construct the simplified relativistic Hamiltonian function, which describes the dynamics of the planetary system of which we are interested, we first need the relativistic Hamiltonian of the two bodies in heliocentric coordinates. The relativistic Hamiltonian of a system of two bodies having mass m_0 and m_1 is:

$$\begin{aligned} H =& \frac{1}{2m_0} y_0^2 + \frac{1}{2m_1} y_1^2 - \mathcal{G}\frac{m_0 m_1}{r_{01}} - \frac{1}{8c^2}\left(\frac{y_0^4}{m_0^3} + \frac{y_1^4}{m_1^3}\right) - \frac{3\mathcal{G}}{2c^2 r_{01}}\left(\frac{m_1 y_0^2}{m_0} + \frac{m_0 y_1^2}{m_1}\right) + \\ & - \frac{\mathcal{G}}{2c^2 r_{01}}\left[7(\mathbf{y}_0 \cdot \mathbf{y}_1) + (\mathbf{y}_0 \cdot \mathbf{n}_{01})(\mathbf{y}_1 \cdot \mathbf{n}_{01})\right] + \frac{\mathcal{G}^2 m_0 m_1 (m_0 + m_1)}{2c^2 r_{01}^2}. \end{aligned} \tag{3.62}$$

Eliminating from it the motion of the center of inertia (i.e. $m_0\mathbf{x}_0 + m_1\mathbf{x}_1 = \mathbf{0}$ and $\mathbf{y}_0 + \mathbf{y}_1 = \mathbf{0}$) and passing in heliocentric coordinates:

$$\begin{aligned}
\mathbf{r}_0 =&\mathbf{x}_0, \qquad \mathbf{r} = \mathbf{x}_1 - \mathbf{x}_0, \qquad \mathbf{p}_0 = \mathbf{y}_0 + \mathbf{y}_1 = 0,\\
\mathbf{p} = \mathbf{y}_1 =& \frac{m_0 m_1}{m_0+m_1}\mathbf{v} + \frac{1}{c^2}\Bigg\{\frac{m_0 m_1}{m_0+m_1}\Bigg[\left(\frac{1}{2m_0^3}+\frac{1}{2m_1^3}\right)v^2\mathbf{v}+\\
&+\frac{\mathcal{G}}{\parallel \mathbf{r} \parallel}\left(\frac{3m_1}{m_0}+\frac{3m_0}{m_1}+7\right)\mathbf{v}+\frac{\mathcal{G}}{\parallel \mathbf{r} \parallel}(\mathbf{n}\cdot\mathbf{v})\mathbf{v}\Bigg]\Bigg\},
\end{aligned} \tag{3.63}$$

where $\mathbf{n} = \mathbf{r}/\parallel \mathbf{r} \parallel$, $\mathbf{p}$ is the momentum of the relative motion and $\mathbf{v} = \dot{\mathbf{r}}$ is the astrocentric velocity of the body 1, the Hamiltonian function becomes

$$\begin{aligned}
H =& \frac{p^2}{2}\left(\frac{1}{m_0}+\frac{1}{m_1}\right) - \mathcal{G}\frac{m_0 m_1}{\parallel \mathbf{r} \parallel} - \frac{p^4}{8c^2}\left(\frac{1}{m_0^3}+\frac{1}{m_1^3}\right) - \frac{3\mathcal{G}p^2}{2c^2\parallel \mathbf{r} \parallel}\left(\frac{m_1}{m_0}+\frac{m_0}{m_1}\right)+\\
&-\frac{\mathcal{G}}{2c^2\parallel \mathbf{r} \parallel}\left(7p^2+(\mathbf{p}\cdot\mathbf{n})^2\right)+\frac{\mathcal{G}^2 m_0 m_1 (m_0+m_1)}{2c^2\parallel \mathbf{r} \parallel^2}.
\end{aligned} \tag{3.64}$$

To simplify the notation, for each planet-star pair we define the following quantities:

$$\begin{aligned}
\mu_i &= \frac{m_0 m_i}{m_0+m_i}, \qquad \beta_i = \mathcal{G}(m_0+m_i), \qquad \upsilon_i = \frac{m_0 m_i}{(m_0+m_i)^2}\\
\gamma_{1,i} &= \frac{1-3\upsilon_i}{8}, \qquad \gamma_{2,i} = \frac{\beta_i(3+\upsilon_i)}{2}, \qquad \gamma_{3,i} = \frac{\beta_i\upsilon_i}{2}, \qquad \gamma_{4,i} = \frac{\beta_i^2}{2}.
\end{aligned} \tag{3.65}$$

where $i = 1, 2$.

The simplified relativistic Hamiltonian can be construct by adding to the classic Hamiltonian (3.22) the relativistic corrections to the Newtonian gravity due to the mutual interaction between the star and the two planets. In particular, we assume that the relativistic corrections to the Newtonian gravity, due to the mutual interactions between the star and one of the two planets, are precisely the relativistic correction present in the relativistic Hamiltonian (3.64) of the two bodies in heliocentric coordinates. Thus, in a barycentric inertial reference system, the Hamiltonian which we are looking for is expressed as the sum of three terms

$$H = H_0 + \varepsilon H_1 + \frac{1}{c^2}H_2, \tag{3.66}$$

where H_0 describes the Keplerian motion of the 2 planets around the star, εH_1 is for the mutual Newtonian point-mass interactions between planets and $\frac{1}{c^2}H_2$ is for the general (Post-Newtonian) relativistic corrections to the Newtonian gravity (the relativistic corrections due to the masses of the planets are skipped).

The terms H_0, εH_1 and $\frac{1}{c^2}H_2$ are expressed by

$$\begin{aligned}
H_0 &= \sum_{i=1}^{2}\left(\frac{1}{2\mu_i}p_i^2 - \mathcal{G}\frac{m_0 m_i}{\parallel \mathbf{r}_i \parallel}\right),\\
\varepsilon H_1 &= \left(\frac{\mathbf{p}_1\cdot\mathbf{p}_2}{m_0} - \mathcal{G}\frac{m_1 m_2}{\parallel \mathbf{r}_1 - \mathbf{r}_2 \parallel}\right),\\
\frac{1}{c^2}H_2 &= \frac{1}{c^2}\sum_{i=1}^{2}\left[-\frac{\gamma_{1,i}}{\mu_i^3}P_i^4 - \frac{\gamma_{2,i}}{\mu_i}\frac{P_i^2}{\parallel \mathbf{r}_i \parallel} - \frac{\gamma_{3,i}}{\mu_i}\frac{(\mathbf{n}_i\cdot\mathbf{P}_i)^2}{\parallel \mathbf{r}_i \parallel} + \gamma_{4,i}\mu_i\frac{1}{\parallel \mathbf{r}_i \parallel^2}\right],
\end{aligned} \tag{3.67}$$

where

$$\begin{aligned} \mathbf{p}_i &= m_i(\dot{\mathbf{r}}_i + \dot{\mathbf{r}}_0), \\ \mathbf{P}_i &= \mu_i \dot{\mathbf{r}}_i + \frac{1}{c^2}\mu_i \left[4\gamma_{1,i}\dot{r}_i^2\dot{\mathbf{r}}_i + \frac{2\gamma_{2,i}}{\| \mathbf{r}_i \|}\dot{\mathbf{r}}_i + \frac{2\gamma_{3,i}}{\| \mathbf{r}_i \|}(\mathbf{n}_i \cdot \dot{\mathbf{r}}_i)\dot{\mathbf{r}}_i\right], \end{aligned} \tag{3.68}$$

with $i = 1, 2$, and $P_i^2 = \mathbf{P}_i \cdot \mathbf{P}_i$.

Hence, in the Hamiltonian (3.66), we put

$$\mathbf{P}_i = \mu_i \dot{\mathbf{r}}_i \tag{3.69}$$

with the accuracy of $O(c^{-2})$ and the Hamiltonian is conserved up to the order $O(c^{-4})$.

Obviously, the Hamiltonian H_2 can be read as the sum of 2 Hamiltonians

$$H_2 = H_2^1 + H_2^2, \tag{3.70}$$

where H_2^i contains only the relativistic term due to the interaction between the star and the i-th planet.

Using the fact that $\mathbf{p}_0 = m_0\dot{\mathbf{r}}_0 + m_1(\dot{\mathbf{r}}_0 + \dot{\mathbf{r}}_1) + m_2(\dot{\mathbf{r}}_0 + \dot{\mathbf{r}}_2) = \mathbf{0}$, i.e.

$$\dot{\mathbf{r}}_0 = -\frac{m_1\dot{\mathbf{r}}_1 + m_2\dot{\mathbf{r}}_2}{m_0 + m_1 + m_2}, \tag{3.71}$$

it is simple to prove that

$$\mathbf{P}_i = \mathbf{p}_i + \frac{\mu_i}{m_0}\mathbf{p}_{3-i} \tag{3.72}$$

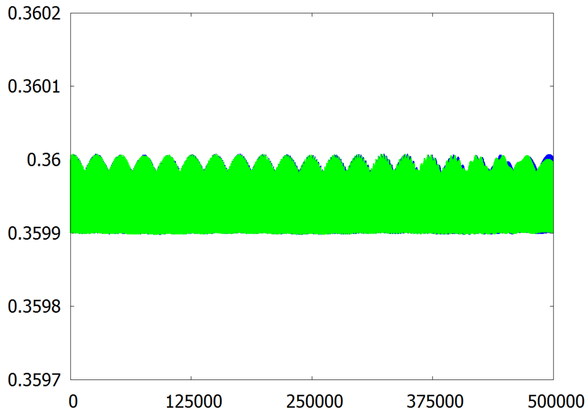
for $i = 1, 2$. Then, the Hamilton's equation for the Hamiltonian (3.66) are

$$\begin{aligned} \dot{\mathbf{r}}_i =& \frac{\mathbf{p}_i}{\mu_i} + \frac{\mathbf{p}_{3-i}}{m_0} - \frac{2}{c^2}\Bigg\{\mathbf{P}_i\left[\frac{2\gamma_{1,i}}{\mu_i^3}P_i^2 + \frac{\gamma_{2,i}}{\mu_i \| \mathbf{r}_i \|}\right] + \frac{\mathbf{P}_{3-i}}{m_0}\Bigg[\frac{2\gamma_{1,3-i}}{\mu_{3-i}^2}P_{3-i}^2 + \\ &+ \frac{\gamma_{2,3-i}}{\| \mathbf{r}_{3-i} \|}\Bigg] + \frac{\gamma_{3,i}}{\mu_i}\frac{(\mathbf{n}_i \cdot \mathbf{P}_i)\mathbf{n}_i}{\| \mathbf{r}_i \|} + \frac{\gamma_{3,3-i}}{m_0}\frac{(\mathbf{n}_{3-i} \cdot \mathbf{P}_{3-i})\mathbf{n}_{3-i}}{\| \mathbf{r}_{3-i} \|}\Bigg\}, \\ \dot{\mathbf{p}}_i =& -\frac{\mathcal{G}(m_0 + m_i)\mu_i\mathbf{r}_i}{\| \mathbf{r}_i \|^3} - \mathcal{G}m_i m_{3-i}\frac{\mathbf{r}_i - \mathbf{r}_{3-i}}{\| \mathbf{r}_i - \mathbf{r}_{3-i} \|^3} - \frac{1}{c^2}\Bigg\{\mathbf{r}_i\Bigg[\frac{\gamma_{2,i}}{\mu_i}\frac{P_i^2}{\| \mathbf{r}_i \|^3} + \\ &- \frac{2\gamma_{4,i}\mu_i}{\| \mathbf{r}_i \|^4} + \frac{3\gamma_{3,i}}{2\mu_i}\frac{(\mathbf{n}_i \cdot \mathbf{P}_i)^2}{\| \mathbf{r}_i \|^3}\Bigg] - \frac{2\gamma_{3,i}}{\mu_i}\frac{\mathbf{n}_i \cdot \mathbf{P}_i}{\| \mathbf{r}_i \|^3}\mathbf{P}_i\Bigg\}, \end{aligned} \tag{3.73}$$

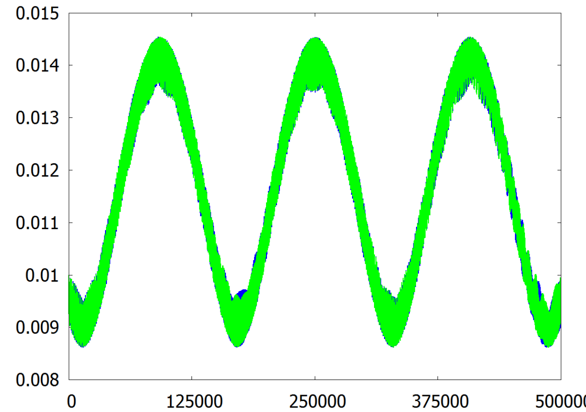
where $\mathbf{P}_i$ is given in (3.72).

We want to compare the dynamics described by these equations of motion respects to the real one. Thus, we integrate equations (3.73) with the classical Runge-Kutta method, using the same initial condition used previously. In the following figures, the blue curves are for the relativistic motions described by equations (3.31)-(3.32), while the green curves are for the simplified relativistic motions described by equations (3.73).

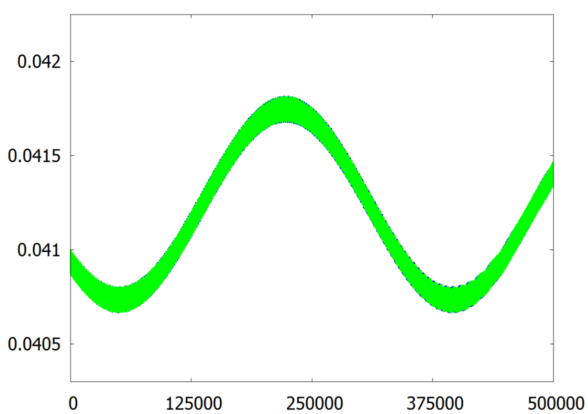
As one can see, in all systems, there is practically no difference between the two cases. Moreover, the simplified relativistic Hamiltonian, although less exact than the relativistic Hamiltonian, is computationally much more affordable (see Table 6.1). These facts justify the simplification that we made earlier.



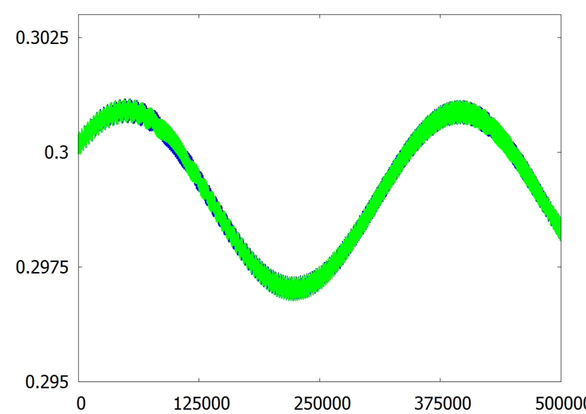
(a) Eccentricity planet HD 190360 b



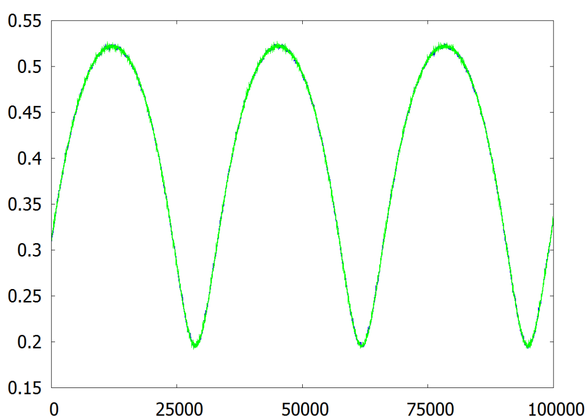
(b) Eccentricity planet HD 190360 c



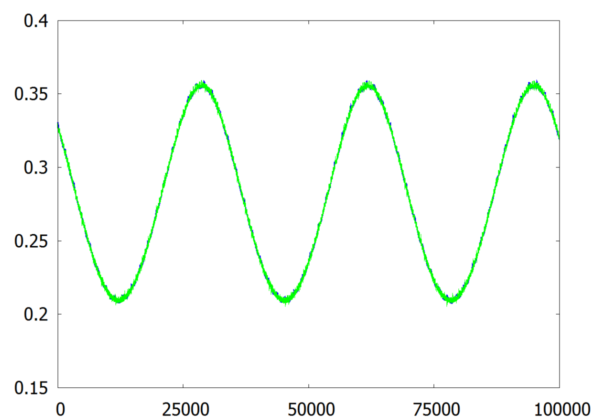
(c) Eccentricity planet HD 11964 b



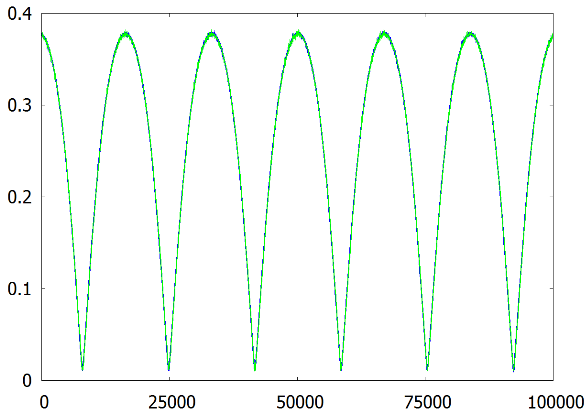
(d) Eccentricity planet HD 11964 c



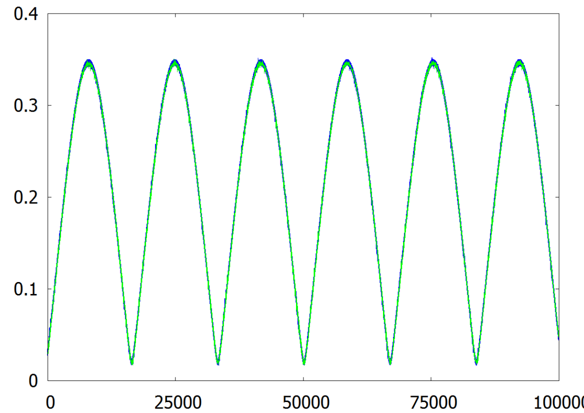
(e) Eccentricity planet HD 169830 b



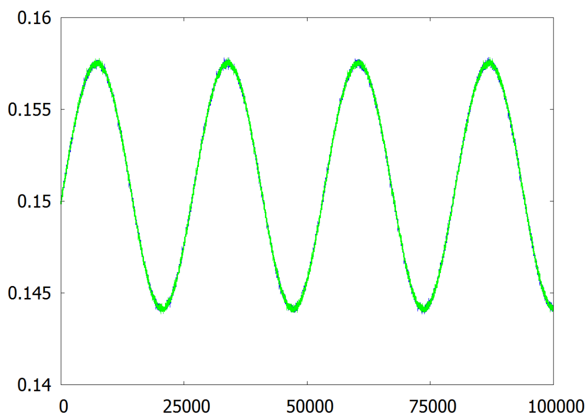
(f) Eccentricity planet HD 169830 c



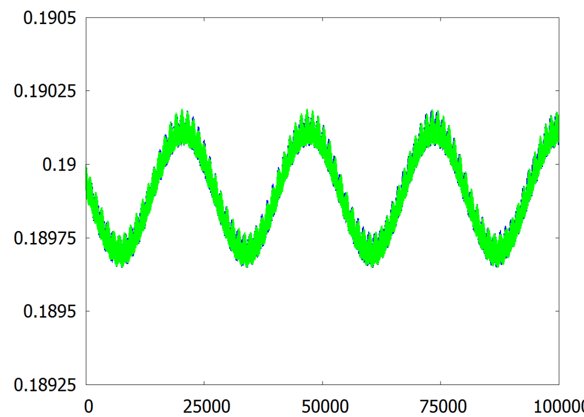
(g) Eccentricity planet HD 12661 b



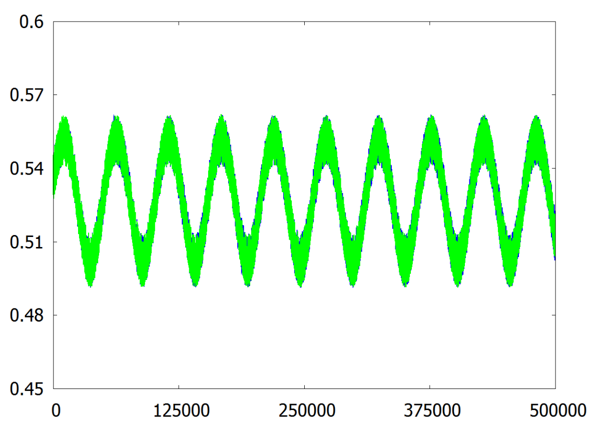
(h) Eccentricity planet HD 12661 c



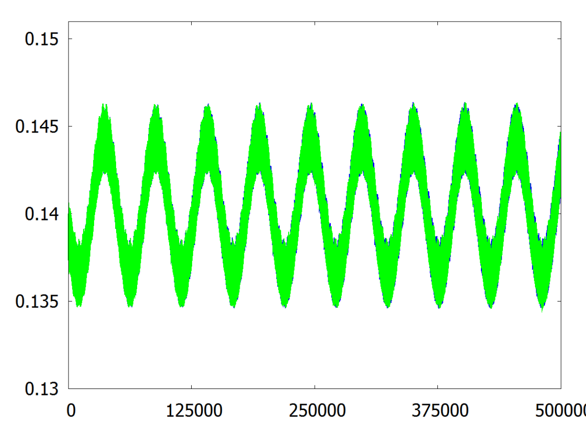
(i) Eccentricity planet BD 082823 b



(j) Eccentricity planet BD 082823 c

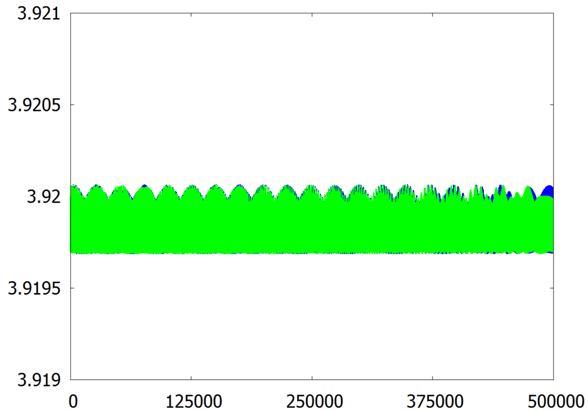


(k) Eccentricity planet HIP 5158 b

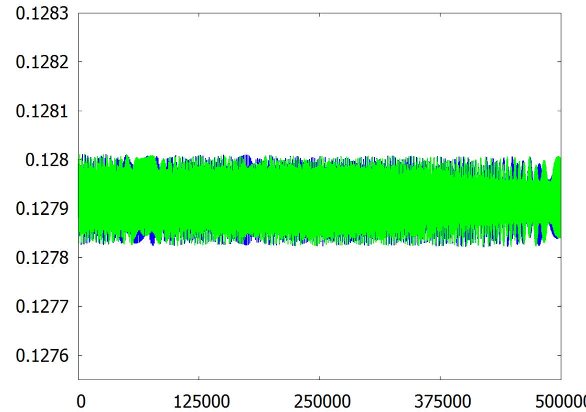


(l) Eccentricity planet HIP 5158 c

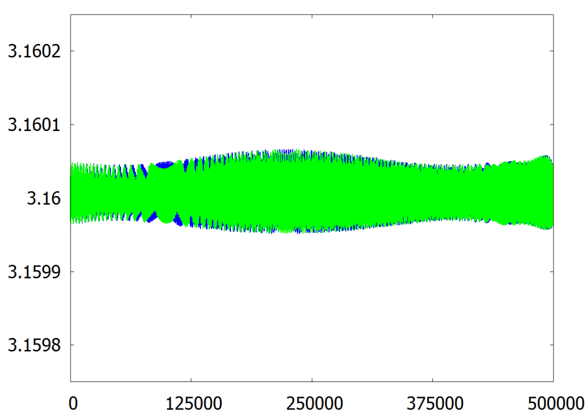
Figure 3.7: Comparison of the evolution of the eccentricity in the case of the relativistic Hamiltonian (3.29) and in the case of the simplified one (3.66). See text for more details.



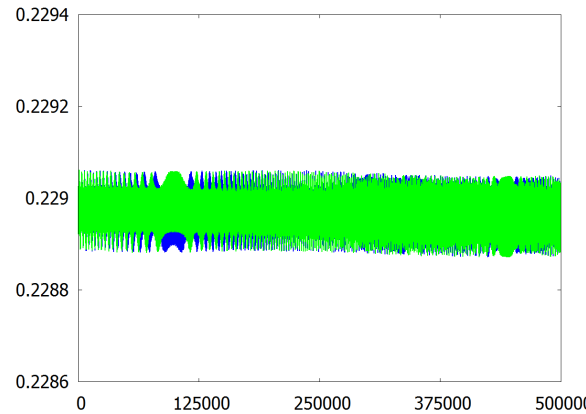
(a) Semi-major axis planet HD 190360 b



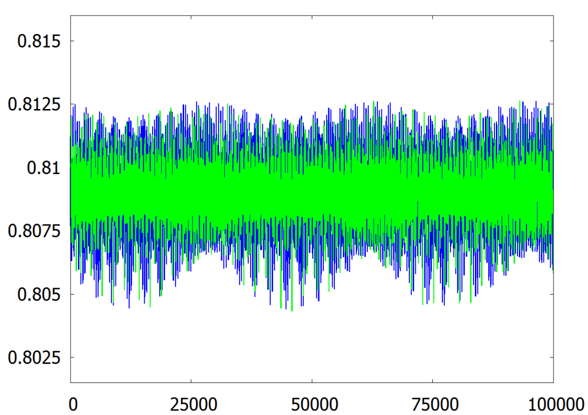
(b) Semi-major axis planet HD 190360 c



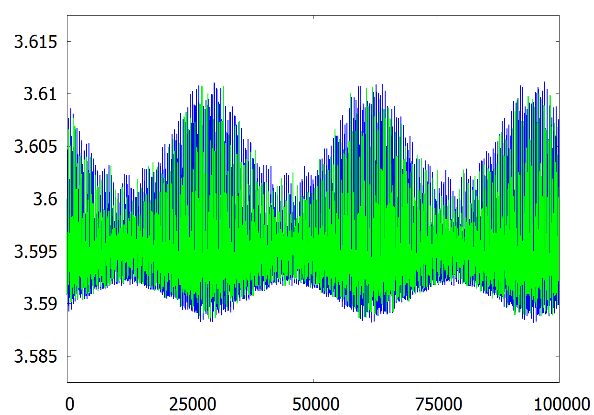
(c) Semi-major axis planet HD 11964 b



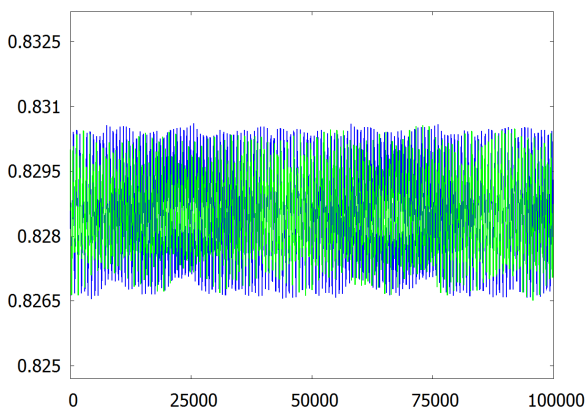
(d) Semi-major axis planet HD 11964 c



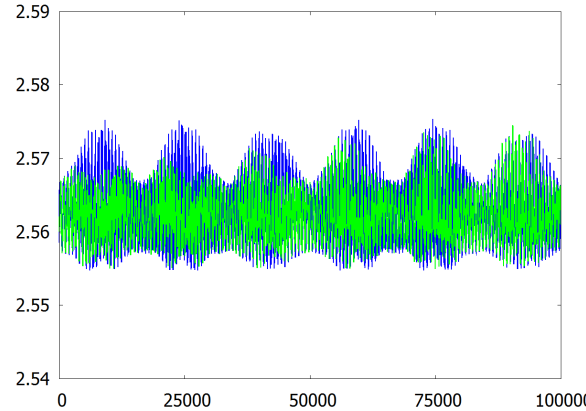
(e) Semi-major axis planet HD 169830 b



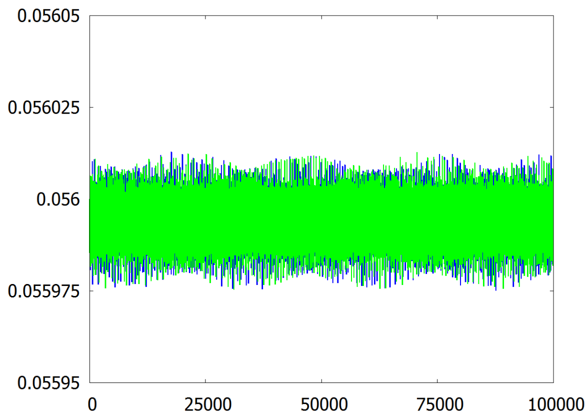
(f) Semi-major axis planet HD 169830 c



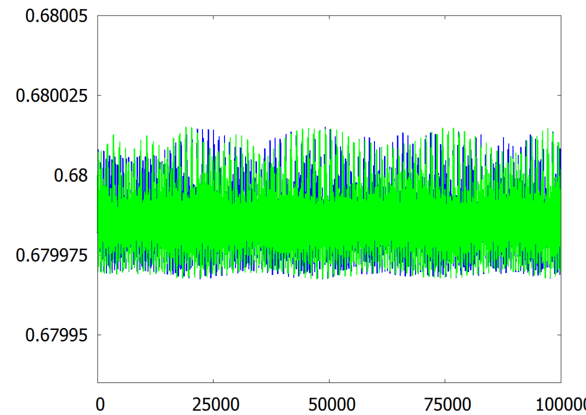
(g) Semi-major axis planet HD 12661 b



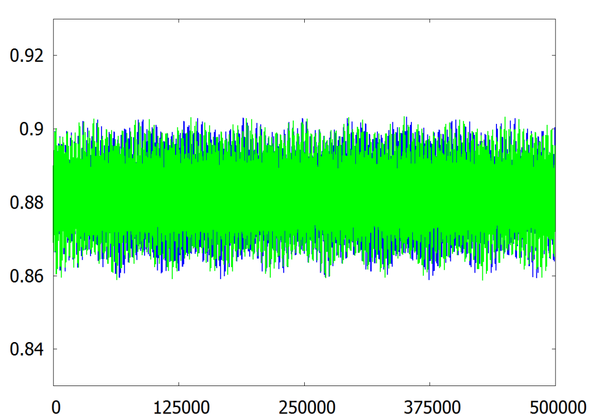
(h) Semi-major axis planet HD 12661 c



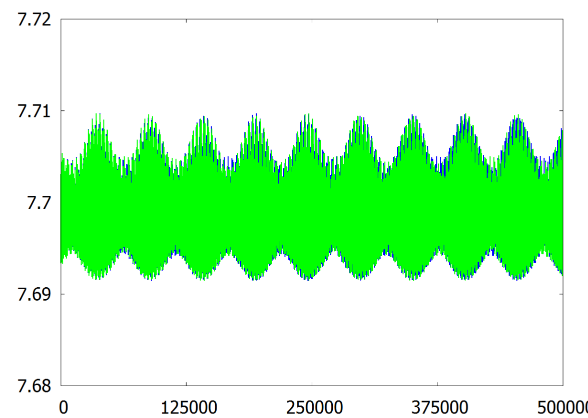
(i) Semi-major axis planet BD 082823 b



(j) Semi-major axis planet BD 082823 c



(k) Semi-major axis planet HIP 5158 b



(l) Semi-major axis planet HIP 5158 c

Figure 3.8: Comparison of the evolution of the semi-major axis in the case of the relativistic Hamiltonian (3.29) and in the case of the simplified one (3.66). See text for more details.

Chapter 4

Hamiltonian systems

In the previous chapter, we have integrated the Hamilton's equations both in the Newtonian case that in the Relativistic one, using the classical Runge-Kutta method. Furthermore, we have found a simplification of the relativistic Hamiltonian (3.29). In particular, we have seen that, at least numerically in the systems that we have considered, the dynamic described by the simplified Hamiltonian (3.66) is very similar to the real one. For these reasons, in the following we construct the semi-analytical solution only for the classical and for the simplified relativistic Hamiltonian (i.e. we don't study analytically the real relativistic Hamiltonian (3.29)).

However, to obtain satisfactory results it is necessary to use a very small integration step and this leads to a technical problem: the direct integrations over the secular time-scale are CPU intensive.

To solve this problem, the idea is to integrate the equations "semi-analytically", using the various tools provided by Hamiltonian's theory and by perturbation theory. In this chapter we present the tools that we will use in following chapters.

4.1 Integrable Hamiltonian

The solution of the system of differential equations

$$\frac{\mathrm{d}r_i}{\mathrm{d}t} = F_i(\mathbf{r}), \tag{4.1}$$

with $i = 1, ..., n$ and $\mathbf{r} = (r_1, ..., r_n)$, can be written in the implicit form as a system of integral equations

$$\int_{\mathbf{r}(0)}^{\mathbf{r}(t)} \frac{\mathrm{d}r_i}{F_i(\mathbf{r})} = \int_0^t \mathrm{d}t. \tag{4.2}$$

The system (4.1) is said to be *integrable by quadratures*, if the integrals on the left-hand side of (4.2) can be explicitly computed, and the resulting relationships $\widetilde{F}_i(\mathbf{r}(t)) - \widetilde{F}_i(\mathbf{r}(0)) = t$, where $\widetilde{F}_i$ are the primitives of $1/F_i$, can be inverted, giving $\mathbf{r}(t)$ as an explicit function of t.

From the general theory of differential equation we learn that the knowledge of enough *first integrals*[1] allows us to perform a complete integration. In particular, for Hamiltonian system, the

[1]A dynamical variable $\Phi(\mathbf{q}, \mathbf{p})$ is said to be a *first integral* if it keeps its value constant under the Hamiltonian flow generated by $H(\mathbf{q}, \mathbf{p})$, i.e. $\dot{\Phi} = \{\Phi, H\} = 0$.

theorem of Liouville states that an n-degree of freedom Hamiltonian is integrable if it admits n independent and in involution first integral $\Phi_1, ..., \Phi_n$ [2].

Theorem 1. *Assume that an autonomous canonical system with n degrees of freedom and with Hamiltonian $H(\mathbf{q}, \mathbf{p})$ possesses n independent first integral $\{\Phi_1(\mathbf{q}, \mathbf{p}), ..., \Phi_n(\mathbf{q}, \mathbf{p})\}$ forming a complete involution system. Then the system is integrable by quadratures. More precisely, one can construct a canonical transformation $(\mathbf{q}, \mathbf{p}) \to (\boldsymbol{\alpha}, \boldsymbol{\Phi})$ such that the transformed Hamiltonian depends only to the new momenta $\Phi_1, ..., \Phi_n$, and the solutions are expressed as*

$$\Phi_j(t) = \Phi_{j,0}, \qquad \alpha_j(t) = \alpha_{j,0} + t\left.\frac{\partial H}{\partial \Phi}\right|_{(\Phi_{1,0},...,\Phi_{n,0})}, \qquad 1 \le j \le n, \tag{4.3}$$

with $\alpha_{j,0}$ and $\Phi_{j,0}$ determined by the initial data.

The local canonical transformation to the new variables $(\boldsymbol{\alpha}, \Phi)$ is constructed by quadratures in the following way. With the non restrictive hypothesis

$$\det\left(\frac{(\Phi_1, ..., \Phi_n)}{(p_1, ..., p_n)}\right) \neq 0, \tag{4.4}$$

the wanted canonical transformation is implicitly defined by

$$\alpha_j = \frac{\partial S}{\partial \Phi_j}(\Phi, \mathbf{q}), \qquad p_j = \frac{\partial S}{\partial q_j}(\Phi, \mathbf{q}), \qquad 1 \le j \le n \tag{4.5}$$

where

$$S(\Phi, \mathbf{q}) = \int \sum_j p_j(\Phi, \mathbf{q}) \mathrm{d}q_j \tag{4.6}$$

and where $p_1(\Phi, \mathbf{q}), ...p_1(\Phi, \mathbf{q})$ are obtained by inversion of $\Phi_1(\mathbf{q}, \mathbf{p}), ..., \Phi_n(\mathbf{q}, \mathbf{p})$.

On the other hand, although it is easier to find constants of motion than to actually solve the Hamiltonian equations, there is no general recipe on how all constant of motion can be found. In particular, if only m constants of motion are known with $m < n$, it is hard to know if additional constants of motion are still to be found or really don't exist. In effect, as stated by Poincaré, integrability itself turns out to be an exceptional property.

4.1.1 The theorem of Arnold-Jost

The *Arnold-Jost theorem* is an extension of Liouville's theorem. Arnold proved that, in the hypothesis of Liouville's theorem and if the $n-$dimensional surface implicitly defined by the constant of motion $\Phi_1, ..., \Phi_n$ is *compact*, it is possible to introduce canonical momenta $\mathbf{I}$ (called *action* of the system) and coordinates $\boldsymbol{\theta}$ such that the coordinates $\theta_1, ..., \theta_n$ are *angles* cyclically defined on the interval $[0, 2\pi]$. A set of canonical variables $(\mathbf{I}, \boldsymbol{\theta})$, where the coordinates $\boldsymbol{\theta}$ are angles, will be generically called *action-angle variables.*

[2] A system of r independent function $\{\Phi_1(\mathbf{q}, \mathbf{p}), ..., \Phi_r(\mathbf{q}, \mathbf{p})\}$ is said to be an *involution system* if the Poisson bracket between any two functions vanishes, i.e. $\{\Phi_j, \Phi_k\} = 0$ for $j, k = 1, ..., r$. It can be proven that an involution system contains at most n independent functions, where n is the number of degrees of freedom.

Theorem 2. *Let the Hamiltonian $H(\mathbf{q},\mathbf{p})$ possesses an involution system $\Phi_1,...,\Phi_n$ of first integrable so that it is integrable in Liouville's sense. Let $\mathbf{c}=(c_1,...,c_n)\in\mathbb{R}^n$ be such that the level surface determined by the equations $\Phi_1(\mathbf{q},\mathbf{p})=c_1,...,\Phi_n(\mathbf{q},\mathbf{p})=c_n$ contains a compact and connected component M_c. Then in a neighborhood U of M_c there are canonical action-angle coordinates $(I,\theta)\colon \mathscr{G}\times\mathbb{T}^n\to U(M_c)$, where $\mathscr{G}\subset\mathbb{R}^n$ is an open neighborhood of $\mathbf{c}$, such that the Hamiltonian depends only on $I_1,...,I_n$, i.e. $H=H(I_1,...,I_n)$, and the corresponding canonical equations are*

$$I_j(t)=I_{j,0},\qquad \theta_j(t)=\theta_{j,0}+t\omega_j(I_{1,0},...,I_{n,0}),\qquad 1\le j\le n \tag{4.7}$$

where $\theta_{j,0}$ and $I_{j,0}$ are the initial data, and $\omega_j=\frac{\partial H}{\partial I_j}$.

The existence of n constants of motion for a $n-$degree of freedom Hamiltonian system ensures that the motion evolves on an $n-$dimensional surface $M_\mathbf{c}$ embedded in $2n-$dimensional phase space. The fact that $\Phi_1,...,\Phi_n$ are a complete involution system ensures that the motion can be decomposed in n independent flows (which do commute) generated by the functions $\Phi_1,...,\Phi_n$, each considered as a one-degree of freedom Hamiltonian. Moreover, the condition that the surface $M_\mathbf{c}$ is compact implies that $M_\mathbf{c}$ is diffeomorphic to $\mathbb{T}^n$ and that the individual flows of $\Phi_1,...,\Phi_n$, and hence global motion, can be decomposed into independent[3] periodic cycles, which we denote by $\gamma_1,...,\gamma_n$.

To construct the action-angle coordinates, we proceed as follows.
As first thing, we must find the cycles γ_i ($i=1,...,n$). This is expected to be the hardest part, because it requires in principle an integration of the system via Liouville's algorithm applied to the involution system $\Phi_1,...,\Phi_n$. The action $\mathbf{I}$ are then introduced by[4]

$$I_i=\frac{1}{2\pi}\oint_{\gamma_i}\sum_{j=1}^n p_j\mathrm{d}q_j,\qquad 1\le i\le n, \tag{4.10}$$

where $p_1(\Phi,\mathbf{q}),...,p_n(\Phi,\mathbf{q})$ are obtained by inversion of $\Phi_1=\Phi_1(\mathbf{q},\mathbf{p}),...,\Phi_n=\Phi_n(\mathbf{q},\mathbf{p})$ with respect to $\Phi_1,...,\Phi_n$. The resulting function $I_1,...,I_n$ depend of course only on $\Phi_1,...,\Phi_n$.

In order to find the angle variables $\theta_1,...,\theta_n$, we apply the theorem of Liouville to the new involution system $I_1,...,I_n$. Thus, writing $\mathbf{p}$ as a function of $\mathbf{I}$ and $\mathbf{q}$, we define the integral generating function as

$$S(\mathbf{I},\mathbf{q})=\int\sum_{j=1}^n p_j(\mathbf{I},\mathbf{q})dq_j \tag{4.11}$$

and the new coordinates $\boldsymbol{\theta}$ are introduced as

$$\theta_k=\frac{\partial S}{\partial I_k}(\mathbf{I},\mathbf{q}),,\qquad 1\le k\le n. \tag{4.12}$$

[3]The cycles γ_j,γ_k are said *independent* if γ_j can not be continuously deformed into γ_k, for $j\neq k$.
[4]It can be proven that the transformation $(\mathbf{q},\mathbf{p})\to(\boldsymbol{\theta},\mathbf{I})$ is canonical if and only if

$$\oint_{\gamma_k}\sum_{j=1}^n p_j\mathrm{d}q_j=\oint_{\gamma_k}\sum_{j=1}^n I_j\mathrm{d}\theta_j,\qquad 1\le k\le n. \tag{4.8}$$

Because all actions $I_1,...,I_n$ and all angles θ_j with $j\neq k$ are constant on γ_k, then the integral on the r.h.s. gives

$$\oint_{\gamma_k}\sum_{j=1}^n p_j\mathrm{d}q_j=I_k\int_0^{2\pi}d\theta_k=2\pi I_k,\qquad 1\le k\le n. \tag{4.9}$$

It can be shown that the transformation $(\mathbf{q},\mathbf{p}) \to (\boldsymbol{\theta},\mathbf{I})$ so defined is canonical and the Hamiltonian H is dependent on the action $\mathbf{I}$ only. Moreover, one can prove that $\theta_1, ..., \theta_n$ are angles, namely θ_i is increased by 2π when a complete cycle γ_i is followed.

The action-angle variables defined above are not unique: indeed, they depend at least on the choice of the cycles $\gamma_1, ..., \gamma_n$, which is not unique. It can be proven the following Lemma:

Lemma 1. *Let $\boldsymbol{\theta}, \mathbf{I}$ be action angle variables. New action-angles variables $\boldsymbol{\psi}, \mathbf{J}$ are constructed by composition of the following canonical transformations:*

- *translation of the action variables and translation of the origin of the angles by a quantity depending on the torus, namely*

$$I_k = J_k + c_k, \qquad \theta_k = \psi_k + \frac{\partial S}{\partial J_k}(\mathbf{J}), \qquad 1 \le k \le n \tag{4.13}$$

  where $\mathbf{c} \equiv (c_1, ..., c_n) \in \mathbb{R}^n$ and where $S(\mathbf{J})$ is an arbitrary differentiable function of the action variables;

- *linear transformation of the action-angle variables by a unimodular matrix[5] A:*

$$\boldsymbol{\psi} = A\boldsymbol{\theta}, \qquad \mathbf{J} = \left(A^T\right)^{-1}\mathbf{I}. \tag{4.14}$$

4.1.2 Periodic and quasi-periodic motion on a torus

For completeness, we study briefly the dynamics of a system which is integrable in Arnold-Jost sense.

We consider the phase space $\mathscr{F} = \mathscr{G} \times \mathbb{T}^n$, where $\mathscr{G} \subset \mathbb{R}^n$ is an open set, $\mathbf{q} = (q_1, ..., q_n) \in \mathbb{T}^n$ are angle variables and $\mathbf{p} = (p_1, ..., p_n) \in \mathscr{G}$ are action variables, and a Hamiltonian function $H = H(p_1, ..., p_n)$. As we have seen, the equations of motion are simply

$$q_j(t) = \omega_j(p_{1,0}, ..., p_{n,0})t + q_{j,0}, \qquad p_j(t) = p_{j,0}, \qquad 1 \le j \le n, \tag{4.15}$$

where $p_{1,0}, ..., p_{n,0}$ and $q_{1,0}, ..., q_{n,0}$ are the initial data and where ω_j are the frequencies.

The motion of the angles on a torus depends on the frequencies $\omega_j(\mathbf{p})$. This leads to introduce the concept of *resonance module*.

Definition 2. *The resonance module $\mathcal{M}_{\boldsymbol{\omega}}$ associated to H is defined by*

$$\mathcal{M}_{\boldsymbol{\omega}} = \{\mathbf{k} = (k_1, ..., k_n) \in \mathbb{Z}^n \,|\, \mathbf{k} \cdot \boldsymbol{\omega} = \sum_{j=1}^{n} k_j \omega_j = 0\}. \tag{4.16}$$

A relation $\mathbf{k} \cdot \boldsymbol{\omega}$ is called *resonance* and the number $\dim \mathcal{M}_{\boldsymbol{\omega}}$ is called the *multiplicity of resonance*[6]. The latter number actually represents the number of independent resonances to which $\boldsymbol{\omega}$ is subjected. The extreme case are $\dim \mathcal{M}_{\boldsymbol{\omega}} = 0$ and $\dim \mathcal{M}_{\boldsymbol{\omega}} = n - 1$.

Proposition 1. *Consider the orbit $\mathbf{q}(t)$ given in (4.15) with initial point $\mathbf{q}_0$ on $\mathbb{T}^n$, and let $\mathcal{M}_{\boldsymbol{\omega}}$ be the resonance module associated to the frequency vector $\boldsymbol{\omega}$. Then the orbit (4.15) is dense on a torus $\mathbb{T}^{n-\dim \mathcal{M}_{\boldsymbol{\omega}}} \subset \mathbb{T}^n$.*

[5] A matrix A is said *unimodular* if it has integer entries and if $\det A = \pm 1$.

[6] The dimension $\dim \mathcal{M}_{\boldsymbol{\omega}}$ of the resonance module is invariant under linear unimodular changes of the angular coordinates.

In particular, if $\mathcal{M}_\omega = \{0\}$ then the frequencies are said to be *nonresonant* and the motion is called *quasi-periodic*, while if $\dim \mathcal{M}_\omega = n - 1$ then the motion on the torus is *periodic* and the frequencies are said to be *completely resonant*.

Consider now a linear transformation

$$\mathbf{q}' = M\mathbf{q} \tag{4.17}$$

with a unimodular matrix M. This transformation changes the system $\dot{\mathbf{q}} = \boldsymbol{\omega}$ into

$$\dot{\mathbf{q}}'_j = \boldsymbol{\omega}', \qquad \boldsymbol{\omega}' = M\boldsymbol{\omega} \tag{4.18}$$

so that the change of coordinates induces a change of the frequencies.

Lemma 2. *Let $\boldsymbol{\omega} \in \mathbb{R}^n$ be given, and let $\dim \mathcal{M}_\omega > 0$. Then there is a unimodular matrix M such that $\boldsymbol{\omega}' = M\boldsymbol{\omega}$ has exactly $\dim \mathcal{M}_\omega$ vanishing components, while the remaining $n - \dim \mathcal{M}_\omega$ components form a non-resonant vector.*

In general, the frequencies depend on the torus we consider, namely on the values of the actions $\mathbf{p}$: in this case, the system is said to be *anisochronous*. Instead, in the case in which the frequencies are independent of the value of the actions $\mathbf{p}$, the system is said to be *isochronous*.

4.2 Delaunay variables

A remarkable application of the Arnold-Jost theorem is the calculation of action-angle variables for the integrable Hamiltonian of the classical two-body problem. They will be the variables that we will be later use to study the dynamics of the planetary problem using Hamiltonian perturbation techniques.

As we have seen, the classical Hamiltonian of a system of two bodies having mass m_0 and m_1 is

$$H = \frac{1}{2m_0} y_0^2 + \frac{1}{2m_1} y_1^2 - \mathcal{G}\frac{m_0 m_1}{\| \mathbf{x}_0 - \mathbf{x}_1 \|}, \tag{4.19}$$

where $y_i^2 = \mathbf{y}_i \cdot \mathbf{y}_i$, for $i = 0, 1$. Eliminating from it the motion of the center of inertia (i.e. imposing that $m_0\mathbf{x}_0 + m_1\mathbf{x}_1 = \mathbf{0}$ and $m_0\dot{\mathbf{x}}_0 + m_1\dot{\mathbf{x}}_1 = \mathbf{0}$) and passing in heliocentric coordinates

$$\mathbf{r}_0 = \mathbf{x}_0, \qquad \mathbf{r} = \mathbf{x}_1 - \mathbf{x}_0, \tag{4.20}$$

$$\mathbf{p}_0 = \mathbf{y}_0 + \mathbf{y}_1 = 0, \qquad \mathbf{p} = \frac{m_0 m_1}{m_0 + m_1}\dot{\mathbf{r}}, \tag{4.21}$$

where we have used $\dot{\mathbf{r}}_0 = -\frac{m_1\dot{\mathbf{r}}}{m_0+m_1}$, the Hamiltonian function becomes

$$H = \frac{p^2}{2\mu} - \mathcal{G}\frac{\mu(m_0 + m_1)}{\| \mathbf{r} \|}, \tag{4.22}$$

where $\mu = \frac{m_0 m_1}{m_0+m_1}$ and where $p^2 = \mathbf{p} \cdot \mathbf{p}$. We can eliminate the mass μ through the transformation

$$\mathbf{p} = \mu\mathbf{p}', \qquad \mathbf{r} = \mathbf{r}'. \tag{4.23}$$

To make the transformation (4.23) canonical, we have to divide the Hamiltonian for μ: in this way, we obtain the Hamiltonian of a point of mass $m_0 + m_1$ in a central field of force.

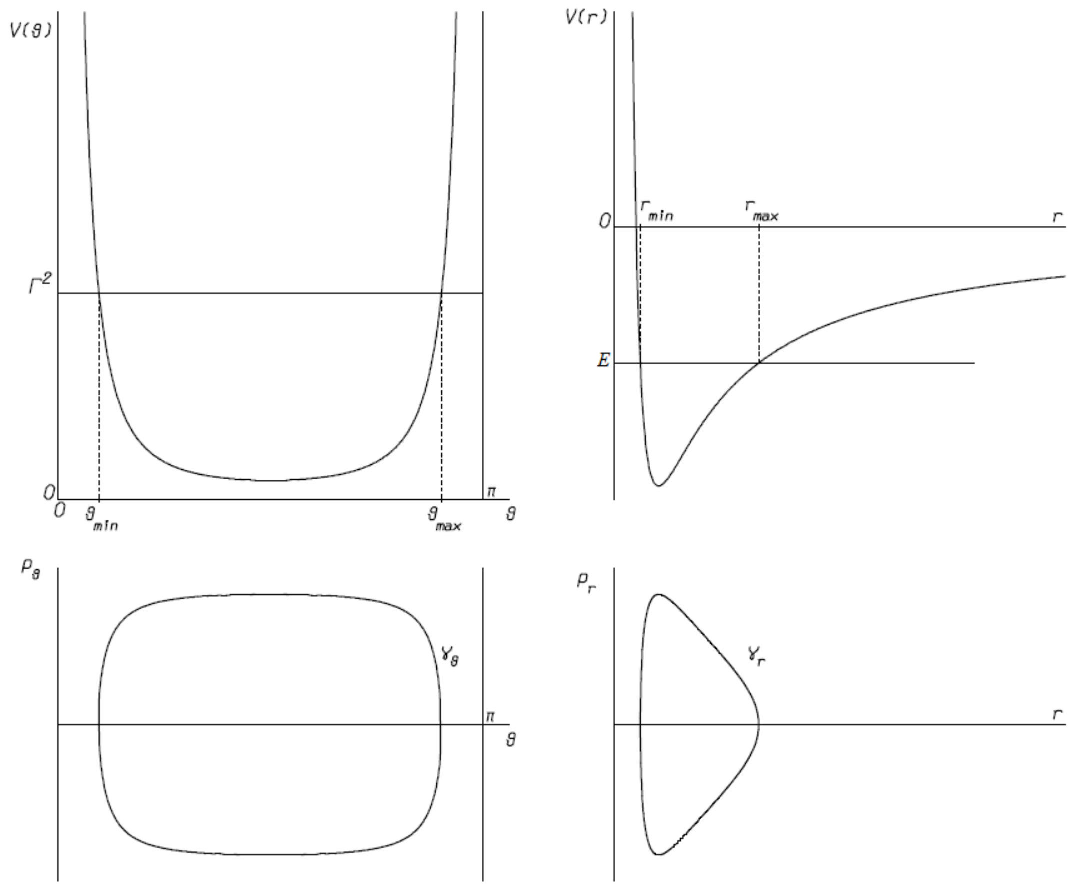


Figure 4.1: The cycle γ_θ and γ_r defined by the constants of motion Γ^2 and E. Reprinted from Fig. 3.4 of Giorgilli, *Hamiltonian systems*.

The Hamiltonian of a point of mass m in a central field of force can be rewritten in polar coordinates as

$$H = \frac{1}{2}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2 \sin^2\theta}\right) - \frac{\mathcal{G}m}{r} \tag{4.24}$$

where r, θ, ϕ are the usual spherical coordinates respect to the center of force, and $p_r = \dot{r}, p_\theta = r^2\dot{\theta}$ and $p_\phi = r^2 \sin^2\theta\dot{\phi}$ are the corresponding conjugate momenta[7].

To calculate the Delaunay variables, we use the following complete involution system of first integrals

$$J = p_\phi, \qquad \Gamma^2 = p_\theta^2 + \frac{J^2}{\sin^2\theta}, \qquad E = \frac{1}{2}\left(p_r^2 + \frac{\Gamma^2}{r^2}\right) - \frac{\mathcal{G}m}{r}, \tag{4.25}$$

where Γ is the norm of the angular momentum of the system, J is the component of the angular momentum vector along the z axis, and E is the total energy of the system (obviously $E = H$).

[7]I remember that the relationship between the spherical coordinates r, θ, ϕ and the Cartesian coordinates x, y, z are $x = r\sin\theta\cos\phi, y = r\sin\theta\sin\phi$ and $z = r\cos\theta$.

By inversion of (4.25), we get

$$p_\phi = J, \qquad p_\theta = \left(\Gamma^2 - \frac{J^2}{\sin^2\theta}\right)^{1/2}, \qquad p_r = \left[2m\left(E + \frac{\mathcal{G}m}{r}\right) - \frac{\Gamma^2}{r^2}\right]^{1/2}. \tag{4.26}$$

Then we proceed to identify the cycles. The Keplerian case is easier than the general case, because each of the constants of motion (4.25) defines a cycle.

The function J defines a cycle γ_ϕ on the (p_ϕ, ϕ) plane, which is given by $p_\phi = J$ and $\phi \in [0, 2\pi]$.

The function Γ^2 can be considered as the Hamiltonian of a point with unit mass, moving on the segment $[0, \pi]$ under the action of the potential $V(\theta) = J^2/\sin^2\theta$. For $\Gamma^2 > \Gamma^2_{min} = J^2$ the orbit in the phase plane (p_θ, θ) is a closed line, giving the second cycle γ_θ (see Fig. 4.1).

The third function can be considered as the Hamiltonian of a point moving on the half line $r > 0$ under the action of the potential $V^*(r) = \frac{\Gamma^2}{2mr^2} - \frac{\mathcal{G}m}{r}$. In particular, the motion on the half line r is bounded for $E_{min} < E < 0$, with $E_{min} = -m\mathcal{G}/(2\Gamma^2)$, while for $E \geq 0$ is unbounded. In the first case the orbit in the phase plane (p_r, r) is a closed curve, and this given the third cycle γ_r (see Fig. 4.1). Conversely, no cycle can be found for $E \geq 0$, and the invariant surface in phase space for the complete problem is actually the product $\mathbb{T}^2 \times \mathbb{R}$. In the latter case the angular variables can be introduced only for the cycles γ_ϕ and γ_θ.

The following step is to introduce the actions of the system, using (4.10). Because on the cycle γ_ϕ only ϕ evolves, $\mathrm{d}\theta$ and $\mathrm{d}r$ are zero, and the sum in (4.10) reduces to the sole term $p_\phi \mathrm{d}\phi$; an analogous situation happens for the cycles γ_θ and γ_r. Therefore, the actions become

$$\begin{aligned}
I_\phi &= \frac{1}{2\pi}\oint_{\gamma_\phi} p_\phi \mathrm{d}\phi = J \\
I_\theta &= \frac{1}{2\pi}\oint_{\gamma_\theta} p_\theta \mathrm{d}\theta = \Gamma - |J| \\
I_r &= \frac{1}{2\pi}\oint_{\gamma_r} p_r \mathrm{d}r = -\Gamma + \mathcal{G}\sqrt{-\frac{m^2}{2E}}.
\end{aligned} \tag{4.27}$$

Using these actions, by inversion of (4.27), we can calculate the Hamiltonian as

$$H = -\frac{m^2\mathcal{G}^2}{2(|I_\phi| + I_\theta + I_r)^2}. \tag{4.28}$$

It is immediately seen that the Hamiltonian actually depends on the sum of the action variables. This implies that the three frequencies of the system actually coincides, which justifies the fact that in the Keplerian description of the planetary motion only one frequency does actually appear.

A better set of action variables is constructed by introducing the *variables of Delaunay* $L, G, \mathcal{H}$ defined by the linear transformation (see Lemma 4.1)

$$\begin{aligned}
L &= |I_\phi| + I_\theta + I_r, \\
G &= |I_\phi| + I_\theta, \\
\mathcal{H} &= |I_\phi|.
\end{aligned} \tag{4.29}$$

It is immediate to notice that G and $\mathcal{H}$ coincide with Γ and J, respectively. Since the transformation is performed via a unimodular matrix, the corresponding transformation on the angles preserves the periods. The Hamiltonian in Delaunay's variables takes the form

$$H = -\frac{m^2\mathcal{G}^2}{2L^2}. \tag{4.30}$$

The canonical transformation should now be completed by constructing the angle variables l, g and h associated respectively to the actions L, G and $\mathcal{H}$. To do this, we first define the generating function

$$S = \int (p_r \mathrm{d}_r + p_\theta \mathrm{d}\theta + p_\phi \mathrm{d}\phi), \tag{4.31}$$

where the expression of $p_r(L, G, \mathcal{H})$, $p_\phi(L, G, \mathcal{H})$ and $p_\theta(L, G, \mathcal{H})$ are obtained by inverting (4.29) and using the relations (4.26) and (4.27). The conjugate angle will be then:

$$l = \frac{\partial S}{\partial L}, \qquad g = \frac{\partial S}{\partial G}, \qquad h = \frac{\partial S}{\partial \mathcal{H}}. \tag{4.32}$$

Using the Delaunay variables $L, G, \mathcal{H}, l, g, h$, the Hamiltonian's equations become

$$\dot{l} = \frac{m^2 \mathcal{G}^2}{L^3}, \qquad \dot{g} = \dot{h} = \dot{G} = \dot{L} = \dot{\mathcal{H}} = 0, \tag{4.33}$$

and the motion is periodic with a single frequency

$$\omega(L) = \frac{m^2 \mathcal{G}^2}{L^3}. \tag{4.34}$$

It also may be useful to recall the relationship between the Delaunay variables and the orbital elements

$$\begin{aligned} L &= \sqrt{m\mathcal{G}a}, & l &= M, \\ G &= L\sqrt{1-e^2}, & g &= \omega, \\ \mathcal{H} &= G \cos i, & h &= \Omega. \end{aligned} \tag{4.35}$$

To avoid the problem that the angles l, g, h are not well defined when the inclination and/or the eccentricity are zero, the following *modified Delaunay variables* are also often used:

$$\begin{aligned} \Lambda &= L, & \lambda &= l + g + h, \\ P &= L - G, & p &= -g - h, \\ Q &= G - \mathcal{H}, & q &= -h. \end{aligned} \tag{4.36}$$

These variables have the advantage that λ is always well defined, while p and q are not defined only when the conjugate actions P and Q are respectively equal to zero. The relationship between the modified Delaunay variables and the orbital elements is given by

$$\begin{aligned} \Lambda &= \sqrt{m\mathcal{G}a}, & \lambda &= M + \varpi, \\ P &= L(1 - \sqrt{1-e^2}), & p &= -\varpi, \\ Q &= 2L\sqrt{1-e^2}\sin^2 i, & q &= -\Omega. \end{aligned} \tag{4.37}$$

Note that, for small eccentricities and inclinations, P is proportional to e^2 and Q to i^2.

In some cases, the singularities $P = 0$ and $Q = 0$ become uncomfortable. To eliminate them, we introduce *Poincaré variables* defined as

$$\begin{aligned} \Lambda &= \sqrt{m\mathcal{G}a}, & \lambda &= M + \varpi, \\ \eta_1 &= \sqrt{2P}\cos p, & \xi_1 &= \sqrt{2P}\sin p, \\ \eta_2 &= \sqrt{2Q}\cos q & \xi_2 &= \sqrt{2Q}\sin q. \end{aligned} \tag{4.38}$$

It's simple to prove the following equalities

$$\eta_1 = \sqrt{2\Lambda}\sqrt{1-\sqrt{1-e^2}}\cos\varpi, \qquad \xi_1 = -\sqrt{2\Lambda}\sqrt{1-\sqrt{1-e^2}}\sin\varpi,$$
$$\eta_2 = \sqrt{2\Lambda}\sqrt{\sqrt{1-e^2}(1-\cos i)}\cos\Omega, \qquad \xi_2 = -\sqrt{2\Lambda}\sqrt{\sqrt{1-e^2}(1-\cos i)}\sin\Omega. \tag{4.39}$$

4.3 Quasi-integrable Hamiltonian

A Hamiltonian system is said to be *quasi-integrable* if, using a suitable set of canonical action-angle variables, its Hamiltonian function can be written as

$$H(\mathbf{p},\mathbf{q}) = H_0(\mathbf{p}) + \varepsilon H_\varepsilon(\mathbf{p},\mathbf{q},\varepsilon), \tag{4.40}$$

where

$$\varepsilon H_\varepsilon(\mathbf{p},\mathbf{q},\varepsilon) = \varepsilon H_1(\mathbf{p},\mathbf{q}) + \varepsilon^2 H_2(\mathbf{p},\mathbf{q}) + \varepsilon^3 H_3(\mathbf{p},\mathbf{q}) + ... \tag{4.41}$$

and where $\mathbf{p} = (p_1, ..., p_n) \in \mathscr{G} \subset \mathbb{R}^n$ are action variables, $\mathbf{q} = (q_1, ..., q_n) \in \mathbb{T}^n$ are angle variables, ε is a small parameter and $\mathrm{grad}_{\mathbf{p}} H_0$, H_1, H_2, ... are intended to be order unity. The Hamiltonian is assumed to be an analytic function of all its arguments; in particular, we shall assume that it can be expanded in power series of ε in a neighborhood of the origin.

The smallness of the parameter ε means that the system is a small perturbation of an integrable one. In particular, it is natural to regard H_0 as the *integrable approximation*, and $\varepsilon H_1 + \varepsilon^2 H_2 + ...$ as its *perturbation*. The Hamilton's equations corresponding to (4.40) are:

$$\begin{aligned} \dot{p}_j &= -\varepsilon \frac{\partial H_\varepsilon(\mathbf{p},\mathbf{q},\varepsilon)}{\partial q_j} \\ \dot{q}_j &= \omega_j(\mathbf{p}) + \varepsilon \frac{\partial H_\varepsilon(\mathbf{p},\mathbf{q},\varepsilon)}{\partial p_j} \end{aligned} \tag{4.42}$$

for $j = 1, ..., n$, where $\omega_j(\mathbf{p}) = \frac{\partial H_0(\mathbf{p})}{\partial p_j}$. The question is raised whether the dynamics remain similar to that of the unperturbed system.

4.3.1 Lindstedt method

For simplicity, suppose we have to study a nearly integrable system of the form

$$\dot{x} = g(x) + \varepsilon f(x) \tag{4.43}$$

where $x \in \mathbb{R}$, ε is a small parameter (i.e. $\varepsilon \ll 1$) and $g\colon \mathbb{R} \to \mathbb{R}$ and $f\colon \mathbb{R} \to \mathbb{R}$ are two analytic functions. A very elementary idea to solve (4.43) is to look for a solution of the form

$$x(t) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + ... \tag{4.44}$$

where $x_0(t), x_1(t), x_2(t), ...$ are functions to be determined. Replacing (4.44) in (4.43) and using the fact that f and g are an analytic functions, i.e. f and g can be expanded in Taylor series in ε

$$g(x_0+\varepsilon x_1+\varepsilon^2 x_2+...) = \sum_{n=0}^{\infty} \varepsilon^n g_n(x_0, ..., x_n), \quad f(x_0+\varepsilon x_1+\varepsilon^2 x_2+...) = \sum_{n=0}^{\infty} \varepsilon^n f_n(x_0, ..., x_n), \tag{4.45}$$

we obtain immediately, for comparison of the coefficients of the development in ε, the system of equations

$$\begin{aligned} \dot{x}_0 &= g_0(x_0) \\ \dot{x}_1 &= g_1(x_0, x_1) + f_0(x_0) \\ &\dots \\ \dot{x}_s &= g_s(x_0, ..., x_s) + f_{s-1}(x_0, ..., x_{s-1}) \\ &\dots \end{aligned} \tag{4.46}$$

where $f_{s-1}(x_0, ..., x_{s-1})$ is completely determined by the previous steps. We can note the recursive nature of this system of equations.

In a similar way, we can look for the system (4.42) a solution of the form

$$\begin{aligned} \mathbf{q}(t) &= \mathbf{q}_0(t) + \varepsilon \mathbf{q}_1(t) + \varepsilon^2 \mathbf{q}_2(t) + ... \\ \mathbf{p}(t) &= \mathbf{p}_0(t) + \varepsilon \mathbf{p}_1(t) + \varepsilon^2 \mathbf{p}_2(t) + ... \end{aligned} \tag{4.47}$$

with of course (in the Hamiltonian case) $\mathbf{p}_0(t) = \mathbf{p}(0)$ and $\mathbf{q}_0(t) = \mathbf{q}(0) + \boldsymbol{\omega}(\mathbf{p}(0))t$. Indeed, inserting such expressions into the equations of motion (4.42) and using the fact that $H(\mathbf{q}, \mathbf{p})$ is an analytic function, we obtain immediately, for comparison of the coefficients of the development in ε, a infinite system of equations which we may attempt to solve recursively.

Such a procedure is called a *Lindstedt method* and series (4.47) are called *Lindstedt series*. Though conceptually simple, however, the method is not simple to use in practice (at least not to obtain mathematically rigorous results) due to the huge amount of terms which are rapidly generated by raising the order. Another important problem concerns the convergence of the series (4.47).

An example of application of the Lindstedt method consists in looking for first integrals for the Hamiltonian (4.40) in order to use the theorem of Arnold-Jost. The idea is to look for a first integral Φ such that $\{H, \Phi\} = 0$ which are in some sense continuations of first integrals of the unperturbed systems, namely in the form

$$\Phi(\mathbf{p}, \mathbf{q}, \varepsilon) = \Phi_0(\mathbf{p}, \mathbf{q}) + \varepsilon \Phi_1(\mathbf{p}, \mathbf{q}) + \varepsilon^2 \Phi_2(\mathbf{p}, \mathbf{q}) + ... \tag{4.48}$$

where Φ_0 is a first integral of H_0, i.e. $\{H_0, \Phi_0\} = 0$, and $\Phi_0, \Phi_1, ...$ are analytic functions of their argument. To solve the equation $\{H, \Phi\} = 0$, we replace the expression in powers of ε for both the Hamiltonian and the function Φ. Then, collecting together all coefficients of the same power of ε, we get the infinite system of equations of the form

$$\{H_0, \Phi_s\} = -\sum_{i=1}^{s} \{H_i, \Phi_{s-i}\} \tag{4.49}$$

where the r.h.s. is a known function, because it depends only on H and on $\Phi_0, ..., \Phi_{s-1}$ which are either know or determined by the equations of preceding steps. Hence we may attempt a recursive solution of the system.

4.4 The averaging principle

We will ignore for a while that the system is Hamiltonian and we consider an arbitrary system of $m \times n$ differential equations

$$\begin{aligned} \dot{\boldsymbol{\varphi}} &= \boldsymbol{\omega}(\mathbf{I}) + \varepsilon \mathbf{f}(\mathbf{I}, \boldsymbol{\varphi}) \\ \dot{\mathbf{I}} &= \varepsilon \mathbf{g}(\mathbf{I}, \boldsymbol{\varphi}) \end{aligned} \tag{4.50}$$

where $\boldsymbol{\varphi} \in \mathbb{T}^m$, $\mathbf{I} \in \mathscr{G} \subset \mathbb{R}^n$ and

$$\mathbf{f}(\mathbf{I}, \boldsymbol{\varphi} + \mathbf{2\pi}) \equiv \mathbf{f}(\mathbf{I}, \boldsymbol{\varphi}), \qquad \mathbf{g}(\mathbf{I}, \boldsymbol{\varphi} + \mathbf{2\pi}) \equiv \mathbf{g}(\mathbf{I}, \boldsymbol{\varphi}). \tag{4.51}$$

The *averaging principle* for system (4.50) consists of its replacement by another system, called the averaged system:

$$\dot{\mathbf{J}} = \varepsilon \bar{\mathbf{g}}(\mathbf{J}) \tag{4.52}$$

in the n dimensional region $\mathscr{G} \subset \mathbb{R}^n$, where

$$\bar{\mathbf{g}}(\mathbf{J}) = \frac{1}{(2\pi)^m} \int_0^{2\pi} ... \int_0^{2\pi} \mathbf{g}(\mathbf{J}, \boldsymbol{\varphi}) \mathrm{d}\varphi_1 ... \mathrm{d}\varphi_m. \tag{4.53}$$

We claim that the system (4.52) is a "good approximation" to the system (4.46).

This approach has been critically considered by Arnold, quoting his book (i.e. *Arnold (1989)*, Chapter 10) *"this principle is neither a theorem, an axiom, nor a definition, but rather a physical proposition, i.e. a vaguely formulated and, strictly speaking, untrue assertions. Such assertions are often fruitful sources of mathematical theorems."*. Indeed, a satisfactory description of the connection between the solutions of the system (4.50) and the system (4.52) in the general case has not yet been found.

In replacing system (4.50) by system (4.52) we discard the term $\varepsilon \tilde{\mathbf{g}}(\mathbf{I}, \boldsymbol{\varphi}) = \varepsilon \mathbf{g}(\mathbf{I}, \boldsymbol{\varphi}) - \varepsilon \bar{\mathbf{g}}(\mathbf{I})$ on the right-hand side. This term has order ε as does the remaining term $\varepsilon \bar{\mathbf{g}}(\mathbf{I})$. The average principle is based on the assertion that in the general case the motion of system (4.50) can be divided into the "evolution" (4.52), which is represented by the term $\varepsilon \bar{\mathbf{g}}$, and into the small oscillations, which is represented by the term $\varepsilon \tilde{\mathbf{g}}$ (see Fig. 4.2). In this general form, this assertions is invalid and the principle itself is untrue.

In effects, it is not difficult to find systems, Hamiltonian or general, for which the averaging principle is satisfied, as well as system for which it is not.

If we apply the principle to the Hamiltonian system (4.40), we find that the average of the perturbation vanishes due to the periodicity of the Hamiltonian:

$$\bar{\mathbf{g}} = \frac{1}{(2\pi)^n} \int_0^{2\pi} ... \int_0^{2\pi} \frac{\partial H_\varepsilon(\mathbf{p}, \mathbf{q}, \varepsilon)}{\partial \mathbf{q}} \mathrm{d}q_1 ... \mathrm{d}q_n = \frac{H_\varepsilon(\mathbf{p}, \mathbf{2\pi}, \varepsilon) - H_\varepsilon(\mathbf{p}, \mathbf{0}, \varepsilon)}{(2\pi)^n} = 0. \tag{4.54}$$

In other words, there is no evolution in an averaged Hamiltonian system, i.e. $\mathbf{p}(t) = \mathbf{p}_0$ for any t.

It can be proven a theorem justifying this principle in the very particular case $m = 1$ (see *Arnold (1989)* for details).

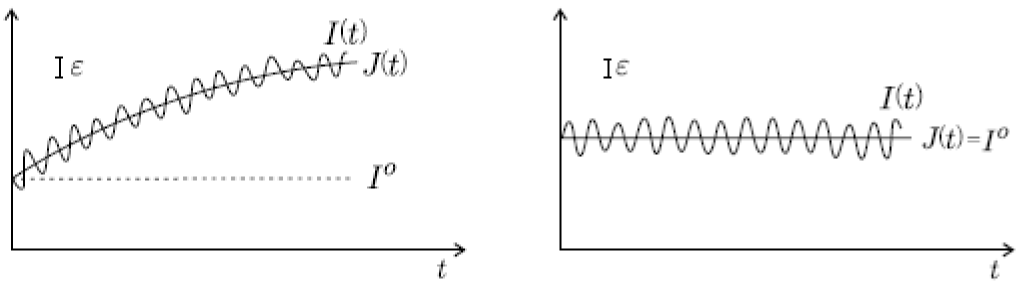


Figure 4.2: Illustrating the averaging principle, in the general case (left) and in the Hamiltonian case (right). Reprinted from Fig. 10 of Benettin, *The elements of Hamiltonian perturbation theory.*

4.5 Introduction to perturbation theory

The power of the Hamiltonian formalism is that, instead of looking for approximations of the real dynamics by working on the equations of motion - which would be very cumbersome - one can work directly on the Hamiltonian function. Then, an alternative idea to study nearly integrable system is to produce small canonical transformations, i.e. transformations near to the identity, which move the perturbation to higher order in ε (see *Morbidelli (2011)* and *Benettin* for more details).

For the following, we only study the case of isochronous Hamiltonian, i.e. of the form

$$H(\mathbf{p},\mathbf{q}) = \boldsymbol{\omega}\cdot\mathbf{p} + \varepsilon H_\varepsilon(\mathbf{p},\mathbf{q},\varepsilon), \tag{4.55}$$

where $\boldsymbol{\omega} \in \mathbb{R}^n \setminus \mathbf{0}$ are fixed and where $H_\varepsilon(\mathbf{p},\mathbf{q},\varepsilon)$ is defined in (4.41). An example of isochronous systems is a system of weakly coupled harmonic oscillators.

Moreover, we assume that there does not exist an integer vector $\mathbf{k}^* \neq \mathbf{0}$ such that the quantity $\mathbf{k}^*\cdot\boldsymbol{\omega}$ vanishes, i.e.

$$\mathbf{k}^*\cdot\boldsymbol{\omega} \neq 0, \qquad \forall \mathbf{k}^* \in \mathbb{Z}^n \setminus \mathbf{0}. \tag{4.56}$$

The general strategy of every perturbation approach for these Hamiltonian systems is to look for a canonical transformation *close to the identity* of the form

$$\mathbf{p} = \mathbf{p}^1 + \varepsilon\mathbf{f}_1(\mathbf{p}^1,\mathbf{q}^1), \qquad \mathbf{q} = \mathbf{q}^1 + \varepsilon\mathbf{g}_1(\mathbf{p}^1,\mathbf{q}^1), \tag{4.57}$$

such that, by substituting (4.57) in (4.55), the latter becomes

$$H(\mathbf{p}^1(\varepsilon),\mathbf{q}^1(\varepsilon)) \equiv H^1(\mathbf{p}^1,\mathbf{q}^1) = \boldsymbol{\omega}\cdot\mathbf{p}^1 + \varepsilon\bar{H}_1(\mathbf{p}^1) + \varepsilon^2 H_2(\mathbf{p}^1,\mathbf{q}^1) + ... \tag{4.58}$$

with some new functions $\bar{H}_1, H_2, ...$ of order unity, where $\bar{H}_1$ is the average of H_1 over the angles $\mathbf{q}$, i.e.

$$\bar{H}_1(\mathbf{p}) = \frac{1}{(2\pi)^n}\int_0^{2\pi} ... \int_0^{2\pi} H_1(\mathbf{p},\mathbf{q})\mathrm{d}q_1...\mathrm{d}q_n. \tag{4.59}$$

If this operation is successful, then $H_0 + \varepsilon\bar{H}_1$ is the integrable approximation of order ε^2 of the real dynamics. In principle, this procedure can be iterated, looking for a sequence of canonical transformations close to the identity:

$$\mathbf{p}^{r-1} = \mathbf{p}^r + \varepsilon^r\mathbf{f}_r(\mathbf{p}^r,\mathbf{q}^r), \qquad \mathbf{q}^{r-1} = \mathbf{q}^r + \varepsilon^r\mathbf{g}_r(\mathbf{p}^r,\mathbf{q}^r), \tag{4.60}$$

which move the perturbation to higher orders in ε, i.e. such that the Hamiltonian in the action-angle variables $\mathbf{p}^r, \mathbf{q}^r$ becomes

$$H^r(\mathbf{p}^r, \mathbf{q}^r) = \boldsymbol{\omega} \cdot \mathbf{p}^r + \varepsilon \bar{H}_1(\mathbf{p}^r) + ... + \varepsilon^r \bar{H}_r(\mathbf{p}^r) + \varepsilon^{r+1} H_{r+1}(\mathbf{p}^r, \mathbf{q}^r) + ..., \tag{4.61}$$

thus obtaining a sequence of increasingly better approximations of the real dynamics. One could hope to use this procedure indefinitely, thus transforming the original $H(\mathbf{p}, \mathbf{q})$ into the integrable $H^\infty(\mathbf{p}^\infty)$. However, we know from the work of Poincaré that in general the procedure cannot be successful up to infinite order. On the other hand, chosen $r \geq 1$ in a way that we consider suitable (e.g. in a way depending on ε and on the properties of the system), one can stop the procedure at order r, obtaining, less than errors of order $O(\varepsilon^{r+1})$, an integrable approximation of the real dynamics, which is $\mathbf{p}^r = \mathbf{p}^r(0)$ and $\mathbf{q}^r = \boldsymbol{\omega}^r t + \mathbf{q}^r(0)$, with $\boldsymbol{\omega}^r = \mathrm{grad}_{\mathbf{p}^r}[\boldsymbol{\omega} \cdot \mathbf{p}^r + ... + \varepsilon^r \bar{H}_r(\mathbf{p}^r)]$. The motion in the original variables $\mathbf{p}, \mathbf{q}$ can be obtained by composing all the sequence of canonical transformations: in particular, because we have done canonical transformations close to the identity, $\mathbf{p}(t), \mathbf{q}(t)$ have oscillations of size ε around the values $\mathbf{p}^r, \mathbf{q}^r(t)$.

4.5.1 Lie series approach

In the procedure sketched in the previous section, a big problem is to select, among all possible transformations $(\mathbf{p}, \mathbf{q}) \to (\mathbf{p}^1, \mathbf{q}^1)$ that transform (4.55) to (4.58), the one that is canonical.

To look for a canonical transformation ε-near the identity (for ε small), we use a method based on the Lie series. On a $2n$-dimensional phase space endowed with canonical coordinates $(\mathbf{p}, \mathbf{q})$, we consider an analytic function $\chi(\mathbf{p}, \mathbf{q})$, that will be called a *generating function*. We define the *Lie derivative* as the time derivative along the Hamiltonian vector field generated by χ, i.e.

$$L_\chi \cdot = \{\cdot, \chi\}. \tag{4.62}$$

We also define the *Lie series operator* as the exponential of L_χ, namely

$$\exp(\varepsilon L_\chi) = \sum_{i=0}^{\infty} \frac{\varepsilon^i}{i!} L_\chi^i, \tag{4.63}$$

which represents the time one evolution of the canonical flow generated by the autonomous Hamiltonian χ.

Using the Lie series, the near identity canonical transformation $(\mathbf{p}, \mathbf{q}) \to (\mathbf{p}^1, \mathbf{q}^1)$ is defined as

$$\begin{aligned} \mathbf{p} = \exp(\varepsilon L_\chi)\mathbf{p}^1 = \mathbf{p}^1 + \varepsilon \frac{\partial \chi}{\partial \mathbf{q}}\Big|_{\mathbf{p}^1,\mathbf{q}^1} + \frac{\varepsilon^2}{2} L_\chi \frac{\partial \chi}{\partial \mathbf{q}}\Big|_{\mathbf{p}^1,\mathbf{q}^1} + ... \\ \mathbf{q} = \exp(\varepsilon L_\chi)\mathbf{q}^1 = \mathbf{q}^1 + \varepsilon \frac{\partial \chi}{\partial \mathbf{p}}\Big|_{\mathbf{p}^1,\mathbf{q}^1} + \frac{\varepsilon^2}{2} L_\chi \frac{\partial \chi}{\partial \mathbf{p}}\Big|_{\mathbf{p}^1,\mathbf{q}^1} + ... \end{aligned} \tag{4.64}$$

where we consider $\chi(\mathbf{p}^1, \mathbf{q}^1)$ as a function of the new variables $(\mathbf{p}^1, \mathbf{q}^1)$. This is a near the identity canonical transformation in the sense that it depends analytically on ε as a parameter, and for $\varepsilon = 0$ is the identity. It is also an easy matter to write the inverse transformation: it is enough to replace χ by $-\chi$ or, equivalently, ε by $-\varepsilon$.

Having defined a coordinate transformation we may ask how a function is transformed. In particular, having given a function $f(\mathbf{q}, \mathbf{p})$ we should calculate the transformed function $f^1(\mathbf{p}^1, \mathbf{q}^1)$ by substitution of the transformation (4.64). In perturbation theory we also need to expand the transformed function in powers of the parameter ε, which is clearly a long and cumbersome procedure since we should apply the Taylor's formula. A remarkable property is the following.

Lemma 3 (The exchange theorem)**.** *Let a generating function* $\chi(\mathbf{p}, \mathbf{q})$ *and a function* $f(\mathbf{q}, \mathbf{p})$ *be given. Then the following equality holds true:*

$$f(\mathbf{p}, \mathbf{q})\Big|_{[\mathbf{p}=\exp(\varepsilon L_\chi)\mathbf{p}^1, \mathbf{q}=\exp(\varepsilon L_\chi)\mathbf{q}^1]} = [\exp(\varepsilon L_\chi) f(\mathbf{p}, \mathbf{q})]\Big|_{\mathbf{p}=\mathbf{p}^1, \mathbf{q}=\mathbf{q}^1} . \tag{4.65}$$

The claim is that the series expansion in ε of the transformed function may be calculated by applying the exponential operator of the Lie series directly to the function, with no need of making a substitution of variables.

An elegant and effective representation of the operation of transforming a function is found as follows. Assume, as is typical in perturbation theory, that the function to be transformed is expanded in power series of the parameter ε, namely $f(\mathbf{p}, \mathbf{q}, \varepsilon) = f_0(\mathbf{p}, \mathbf{q}) + \varepsilon f_1(\mathbf{p}, \mathbf{q}) + \varepsilon^2 f_2(\mathbf{p}, \mathbf{q}) + ...$ and that we want to write the transformed function $g = \exp(\varepsilon L_\chi) f$ as a power series $g_0 + \varepsilon g_1 + \varepsilon^2 g_2 + ...$ in ε. Working at a formal level we may use the linearity of the Lie series operator, thus writing

$$g = \exp(\varepsilon L_\chi) f_0 + \varepsilon \exp(\varepsilon L_\chi) f_1 + \varepsilon^2 \exp(\varepsilon L_\chi) f_2 + ... \tag{4.66}$$

That is, we apply the Lie series to every term of the expansion of f . The action of the operator is represented by the triangular diagram

$$\begin{array}{cccccc}
g_0 & f_0 & & & & \\
 & \downarrow & & & & \\
g_1 & L_\chi f_0 & f_1 & & & \\
 & \downarrow & \downarrow & & & \\
g_2 & \frac{1}{2!} L_\chi^2 f_0 & L_\chi f_1 & f_2 & & \\
 & \downarrow & \downarrow & \downarrow & & \\
g_2 & \frac{1}{3!} L_\chi^3 f_0 & \frac{1}{2!} L_\chi^2 f_1 & L_\chi f_2 & f_3 & \\
 & \downarrow & \downarrow & \downarrow & \downarrow & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array} \tag{4.67}$$

where terms of the same order in ε are aligned on the same line. The calculation may be performed by columns: if the function f and the generating function χ are known, then every column may be calculated proceeding up-down until the line corresponding to the wanted order in ε is reached. Then it is enough to add together all terms appearing on the same line, and this gives every term of the expansion of g up to the wanted order. A compact form of the diagram is given by the recursive formula

$$g_0 = f_0, \qquad g_r = \sum_{j=0}^{r} \frac{1}{j!} L_\chi^j f_{r-j} \quad \text{for} \quad r > 0. \tag{4.68}$$

Composition of Lie Series

As we have seen, the Lie Series defines a near the identity transformation, which is also canonical. Anyhow, more general transformations can be constructed by *composition of Lie Series*.

Let us consider the sequence of generating function $\boldsymbol{\chi} = \{\chi_1, \chi_2, \chi_3, ...\}$, and let the sequence of operators $S^{(1)}, S^{(2)}, S^{(3)}, ...$ be defined as

$$S^{(1)} = \exp(\varepsilon L_{\chi_1}), \qquad S^{(k)} = \exp(\varepsilon^k L_{\chi_k}) S^{(k-1)} \quad \text{for} \quad k \geq 1. \tag{4.69}$$

We can also defined the operator

$$S_{\boldsymbol{\chi}} = ... \circ \exp(\varepsilon^3 L_{\chi_3}) \circ \exp(\varepsilon^2 L_{\chi_2}) \circ \exp(\varepsilon L_{\chi_1}). \tag{4.70}$$

Working at a formal level, we assume that every operator is well defined and that the sequence tends to some limit.

Inverting the operator so defined is not difficult. Let us define the sequence $\widetilde{S}^{(1)}, \widetilde{S}^{(2)}, ..., \widetilde{S}^{(r)}$ as

$$\widetilde{S}^{(1)} = \exp(-\varepsilon L_{\chi_1}), \qquad \widetilde{S}^{(k)} = \widetilde{S}^{(k-1)} \exp(-\varepsilon^k L_{\chi_k}) \quad \text{for} \quad k \geq 1. \tag{4.71}$$

and let us assume again that the sequence tends to some limit. It is simple to prove that

$$\widetilde{S}^{(k)} \circ S^{(k)} = \mathrm{Id} \tag{4.72}$$

and that the inverse of the operator S_χ is

$$S_{\boldsymbol{\chi}}^{(-1)} = \exp(-\varepsilon L_{\chi_1}) \circ \exp(-\varepsilon^2 L_{\chi_2}) \circ \exp(-\varepsilon^3 L_{\chi_3}) \circ ... \tag{4.73}$$

Lemma 4. *Consider a near the identity canonical transformation in the form*

$$\mathbf{q} = \mathbf{q}' + \varepsilon \boldsymbol{\varphi}_1(\mathbf{p}', \mathbf{q}') + \varepsilon^2 \boldsymbol{\varphi}_2(\mathbf{p}', \mathbf{q}') + ..., \qquad \mathbf{p} = \mathbf{p}' + \varepsilon \boldsymbol{\psi}_1(\mathbf{p}', \mathbf{q}') + \varepsilon^2 \boldsymbol{\psi}_2(\mathbf{p}', \mathbf{q}') + ... \tag{4.74}$$

Then there exists a sequence $\chi_1(\mathbf{p}', \mathbf{q}'), \chi_2(\mathbf{p}', \mathbf{q}'), ...$ of generating functions such that

$$\mathbf{q} = S_{\boldsymbol{\chi}} \mathbf{q}', \qquad \mathbf{p} = S_{\boldsymbol{\chi}} \mathbf{p}'. \tag{4.75}$$

4.6 Birkhoff normal form

The problem is now reduced to finding a suitable generating Hamiltonian χ such that, if H has the form (4.55), H^1 has the form (4.58).

As we have seen, using Lemma 4.3 to write the Hamiltonian in the new variables, one gets

$$H^1(\mathbf{p}^1, \mathbf{q}^1) = [\exp(\varepsilon L_\chi) H(\mathbf{p}, \mathbf{q})] \Big|_{\mathbf{p}=\mathbf{p}^1, \mathbf{q}=\mathbf{q}^1}. \tag{4.76}$$

If we write (4.76) in an explicit form of order ε^2, we obtain

$$\begin{aligned} H^1(\mathbf{p}^1, \mathbf{q}^1) = &\boldsymbol{\omega} \cdot \mathbf{p}^1 + \varepsilon H_1(\mathbf{p}^1, \mathbf{q}^1) + \varepsilon\{\boldsymbol{\omega} \cdot \mathbf{p}^1, \chi(\mathbf{p}^1, \mathbf{q}^1)\} + \varepsilon^2 H_2(\mathbf{p}^1, \mathbf{q}^1) + \\ &+ \varepsilon^2\{H_1(\mathbf{p}^1, \mathbf{q}^1), \chi(\mathbf{p}^1, \mathbf{q}^1)\} + \frac{\varepsilon^2}{2}\{\{\boldsymbol{\omega} \cdot \mathbf{p}^1, \chi(\mathbf{p}^1, \mathbf{q}^1)\}, \chi(\mathbf{p}^1, \mathbf{q}^1)\} + O(\varepsilon^3) \end{aligned} \tag{4.77}$$

Moreover, we assume that ε is small enough such that, if χ and H are analytic, the series (4.76) is absolutely convergent. From (4.77), we immediately see that the term of H^1 of order zero in ε is $\boldsymbol{\omega} \cdot \mathbf{p}^1$; therefore H^1 will have the form (4.58) if and only if the first-order term in ε will be a function of the sole actions $\mathbf{p}^1$, namely if and only if the equation

$$H_1(\mathbf{p}^1, \mathbf{q}^1) + \{\boldsymbol{\omega} \cdot \mathbf{p}^1, \chi(\mathbf{p}^1, \mathbf{q}^1)\} = \bar{H}_1(\mathbf{p}^1), \tag{4.78}$$

which is also called the *homological equation*, is satisfied by some functions $\chi(\mathbf{p}^1, \mathbf{q}^1)$ and $\bar{H}_1(\mathbf{p}^1)$. To solve (4.78) we can use the fact that the coordinates $\mathbf{q}^1$ are angles and that the Hamiltonian H is periodic in $\mathbf{q}^1$. We expand H_1 in a Fourier series as

$$H_1(\mathbf{p}^1, \mathbf{q}^1) = \sum_{\mathbf{k} \in \mathbb{Z}^n} c_{\mathbf{k}}(\mathbf{p}^1) \exp(\sqrt{-1}\mathbf{k} \cdot \mathbf{q}^1) \tag{4.79}$$

and then we then look for a solution χ of (4.78) of a similar form

$$\chi(\mathbf{p}^1, \mathbf{q}^1) = \sum_{\mathbf{k} \in \mathbb{Z}^n} d_{\mathbf{k}}(\mathbf{p}^1) \exp(\sqrt{-1}\mathbf{k} \cdot \mathbf{q}^1). \tag{4.80}$$

It is simple to prove that

$$\{\boldsymbol{\omega} \cdot \mathbf{p}^1, \chi(\mathbf{p}^1, \mathbf{q}^1)\} = -\sqrt{-1} \sum_{\mathbf{k} \in \mathbb{Z}^n} d_{\mathbf{k}}(\mathbf{p}^1) \mathbf{k} \cdot \boldsymbol{\omega} \exp(\sqrt{-1}\mathbf{k} \cdot \mathbf{q}^1). \tag{4.81}$$

With the assumption (4.56), the solution of equation (4.78) is given by a generating function χ of the form (4.80) with coefficients $d_{\mathbf{k}}$ given by

$$d_{\mathbf{0}} = 0, \qquad d_{\mathbf{k}}(\mathbf{p}^1) = -\sqrt{-1} \frac{c_{\mathbf{k}}(\mathbf{p}^1)}{\mathbf{k} \cdot \boldsymbol{\omega}} \quad \text{for} \quad \mathbf{k} \neq \mathbf{0} \tag{4.82}$$

and a function $\bar{H}_1$ that is simply

$$\bar{H}_1(\mathbf{p}^1) = c_{\mathbf{0}}(\mathbf{p}^1). \tag{4.83}$$

However, this solution of (4.78) is only a *formal* solution, because we have to prove that the generating Hamiltonian χ is well defined as an analytic function, i.e. that its Fourier series (4.80) with coefficients given by (4.82) is absolutely convergent. Obviously, if we show that the series is absolutely convergent, then the generating function χ, and so the canonical transformation relating the original variables $\mathbf{p}, \mathbf{q}$ to the new ones $\mathbf{p}^1, \mathbf{q}^1$, is determined.

In this case, the terms of the new Hamiltonian $H^1(\mathbf{p}^1, \mathbf{q}^1)$ can be simply computed using the complete expression (4.77). In particular H^1 has the form

$$H^1(\mathbf{p}^1, \mathbf{q}^1) = \boldsymbol{\omega} \cdot \mathbf{p}^1 + \varepsilon \bar{H}_1(\mathbf{p}^1) + \sum_{j=2}^{\infty} \varepsilon^j H_j^{(1)}(\mathbf{p}^1, \mathbf{q}^1), \tag{4.84}$$

where $H_j^{(1)}(\mathbf{p}^1, \mathbf{q}^1)$ is given by the recursive formula (4.68) for $j \geq 1$. The Hamiltonian in the form (4.84) is said to be in *Birkhoff normal form* to first order in ε.

The most elementary situation in which the series (4.80) is absolutely convergent is the case in which the Hamiltonian H_1 is a Fourier polynomial, i.e.

$$H_1(\mathbf{p}^1, \mathbf{q}^1) = \sum_{\mathbf{k} \in \mathcal{N}} c_{\mathbf{k}}(\mathbf{p}^1) \exp(\sqrt{-1}\mathbf{k} \cdot \mathbf{q}^1) \tag{4.85}$$

where $\mathcal{N} \subset \mathbb{Z}^n$ is a finite subset. In this case, χ is obviously given by

$$\chi(\mathbf{p}^1, \mathbf{q}^1) = \sum_{\mathbf{k} \in \mathcal{N} \setminus \mathbf{0}} -\sqrt{-1} \frac{c_{\mathbf{k}}(\mathbf{p}^1)}{\mathbf{k} \cdot \boldsymbol{\omega}} \exp(\sqrt{-1}\mathbf{k} \cdot \mathbf{q}^1), \tag{4.86}$$

which is analytic because none of the denominators $\mathbf{k} \cdot \boldsymbol{\omega}$ vanish and because the Fourier expansion contains only a finite number of terms.

4.7 The small divisors problem

If the Hamiltonian H_1 has a full Fourier series, it is not sufficient to assume that the divisor is not zero, because it can assume arbitrarily small values making convergence of the Fourier series (4.80) doubtful. For this reason, it is necessary that the Fourier coefficients $d_{\mathbf{k}}$ decrease sufficiently fast by increasing (the norm of) of $\mathbf{k}$, i.e. that the denominators $\mathbf{k} \cdot \boldsymbol{\omega}$ should never be too close to zero for large $\mathbf{k}$.

Definition 3. *The frequencies $\boldsymbol{\omega}$ are said to be (γ, τ)-Diophantine, if there exist real constants $\gamma > 0$ and $\tau > n-1$ such that*

$$|\mathbf{k} \cdot \boldsymbol{\omega}| > \frac{\gamma}{|\mathbf{k}|^\tau}, \qquad \forall \mathbf{k} \in \mathbb{Z}^n \setminus \mathbf{0}, \tag{4.87}$$

with $|\mathbf{k}| = |k_1| + |k_2| + ... + |k_n|$.

For the following we need to introduce a complexification of domains. Consider the common case of a phase space $\mathscr{G} \times \mathbb{T}^n$, where $\mathscr{G} \subset \mathbb{R}^n$, endowed with action-angle variables $\mathbf{q} \in \mathbb{T}^n$ and $\mathbf{p} \in \mathscr{G}$. For $\sigma > 0$, we define the complexification $\mathbb{T}^n_\sigma$ of the n-torus as

$$\mathbb{T}^n_\sigma = \{\boldsymbol{\varphi} \in \mathbb{C}^n \mid \quad \mathrm{Re}(\varphi_j) \in [0, 2\pi], \, |\mathrm{Im}(\varphi_j)| < \sigma\}. \tag{4.88}$$

For any real analytic function $F\colon \mathbb{T}^n_\sigma \to \mathbb{C}$, periodic of real period 2π in each argument, we denote by $\| F \|_{\infty,\sigma}$ the supremum norm of F in the strip, and we define the "Fourier norm" as

$$\| F \|_\sigma = \sum_{\mathbf{k} \in \mathbb{Z}^n} |F_{\mathbf{k}}| \exp(|\mathbf{k}|\sigma), \tag{4.89}$$

where $F_{\mathbf{k}}$ is the $\mathbf{k}$-th Fourier coefficient of F.

Lemma 5. *If F is analytic in $\mathbb{T}^n_\sigma$ and $\| F \|_{\infty,\sigma} < \infty$, then the amplitude of the Fourier components $F_{\mathbf{k}}$ decrease exponentially with $|\mathbf{k}|$, according to*

$$|F_{\mathbf{k}}| \leq \| F \|_{\infty,\sigma} \exp(-|\mathbf{k}|\sigma). \tag{4.90}$$

The complexification $\mathscr{G}_\rho$ of the real domain $\mathscr{G}$ is defined as

$$\mathscr{G}_\rho = \bigcup_{\mathbf{p} \in \mathscr{G}} \Delta_\rho(\mathbf{p}), \qquad \Delta_\rho(\mathbf{p}) = \{\mathbf{z} \in \mathbb{C}^n \mid |\mathbf{z} - \mathbf{p}| < \rho\}, \tag{4.91}$$

where $|\mathbf{z}| = \max_j |z_j|$. The phase space becomes

$$D_{\rho,\sigma} = \mathscr{G}_\rho \times \mathbb{T}^n_\sigma. \tag{4.92}$$

For any analytic function $F\colon D_{\rho,\sigma} \to \mathbb{C}$, we extend the definition of the "Fourier norm" in the following way

$$\| F \|_{\rho,\sigma} = \sup_{\mathbf{I}\in\mathscr{G}_\rho} \| F(\mathbf{I},\cdot) \|_\sigma \tag{4.93}$$

We are now ready to introduce a proposition which states the possibility of performing one perturbation step for the isochronous system (4.55), in the assumption that the frequencies are (γ,τ)-Diophantine and that the perturbation H_ε is analytic in the above introduced domain $D_{\rho,\sigma}$.

Proposition 2. *Let H as in (4.55), with (γ,τ)-Diophantine frequency $\boldsymbol{\omega}$. Assume H_ε is analytic in $D_{\rho,\sigma}$ for some (ρ,σ), and $\| H_\varepsilon \|_{\rho,\sigma} < \infty$. Let*

$$E = \gamma\rho\sigma^{\tau+1}. \tag{4.94}$$

For any $0 < \eta < 1$, if

$$\varepsilon < \varepsilon_0 = \frac{c\eta^{\tau+2}E}{\| H_1 \|_{\rho,\sigma}}, \tag{4.95}$$

where c is a constant depending only on τ, there exists an analytical transformation $S_\chi^\varepsilon\colon D_{(1-\eta)(\rho,\sigma)} \to D_{(\rho,\sigma)}$ such that the new Hamiltonian $H^1 = S_\chi^\varepsilon H\colon D_{(1-\eta)(\rho,\sigma)} \to \mathbb{C}$ has the form

$$H^1(\mathbf{p}^1,\mathbf{q}^1) = \boldsymbol{\omega}\cdot\mathbf{p}^1 + \varepsilon\bar{H}_1(\mathbf{p}^1) + \sum_{j=2}^{\infty}\varepsilon^j H_j^{(1)}(\mathbf{p}^1,\mathbf{q}^1). \tag{4.96}$$

Another way to solve the problem of the small divisors takes advantage of the analytic properties of H_1. Using the exponential decay of its coefficients (see Lemma 4.5), the idea is to separate the Fourier expansion of H_1 in two parts, namely $H_1 = H_1^{<K} + H_1^{\geq K}$ with

$$H_1^{<K} = \sum_{\mathbf{k}\in\mathbb{Z}^n\setminus\mathbf{0};|\mathbf{k}|<K} c_{\mathbf{k}}(\mathbf{p})\exp(\sqrt{-1}\mathbf{k}\cdot\mathbf{q}), \qquad H_1^{\geq K} = \sum_{\mathbf{k}\in\mathbb{Z}^n;|\mathbf{k}|\geq K} c_{\mathbf{k}}(\mathbf{p})\exp(\sqrt{-1}\mathbf{k}\cdot\mathbf{q}) \tag{4.97}$$

choosing K large enough that $H_1^{\geq K}$ is of order ε with respect to $H_1^{<K}$. In this case, the Hamiltonian becomes

$$H(\mathbf{p},\mathbf{q}) = \boldsymbol{\omega}\cdot\mathbf{p} + \varepsilon H_1^{<K}(\mathbf{p},\mathbf{q}) + \varepsilon^2\widetilde{H}_2(\mathbf{p},\mathbf{q}) + ... \tag{4.98}$$

where

$$\widetilde{H}_2(\mathbf{p},\mathbf{q}) = H_2(\mathbf{p},\mathbf{q}) + \frac{1}{\varepsilon}H_1^{\geq K}(\mathbf{p},\mathbf{q}). \tag{4.99}$$

For the following, I recall $\widetilde{H}_2$ as H_2 and I remember that $\bar{H}_1^{<K} = \bar{H}_1$.

Then, in the expression (4.77) only $\varepsilon H_1^{<K}$ appears at order ε, and equation (4.78) admits the solution

$$\chi(\mathbf{p}^1,\mathbf{q}^1) = \sum_{\mathbf{k}\in\mathbb{Z}^n\setminus 0;|\mathbf{k}|<K} -\sqrt{-1}\frac{c_{\mathbf{k}}(\mathbf{p}^1)}{\mathbf{k}\cdot\boldsymbol{\omega}}\exp(\sqrt{-1}\mathbf{k}\cdot\mathbf{q}^1). \tag{4.100}$$

Now χ is obviously analytic because the Fourier expansion contains only a finite number of terms and because none of the denominators $\mathbf{k}\cdot\boldsymbol{\omega}$ vanish.

4.8 Beyond the first order

After the elimination of the harmonics with coefficients of order ε, we would like to iterate the procedure, in order to eliminate also the nonresonant harmonics with coefficients of higher order in ε.

The idea is to proceed recursively and perform any finite number $r \geq 1$ of steps, so as to push the remainder to order ε^{r+1}. The input of a typical step is a Hamiltonian H^{r-1} in a "Birkhoff normal form" up to order $r-1$:

$$H^{r-1}(\mathbf{p}^{r-1},\mathbf{q}^{r-1}) = \boldsymbol{\omega}\cdot\mathbf{p}^{r-1} + \sum_{j=1}^{r-1}\varepsilon^j \bar{H}_j^{(j-1)}(\mathbf{p}^{r-1}) + \sum_{j=r}^{\infty}\varepsilon^j H_j^{(r-1)}(\mathbf{p}^{r-1},\mathbf{q}^{r-1}), \tag{4.101}$$

where $\bar{H}_1^{(0)} \equiv \bar{H}_1$. The output is a new Hamiltonian H^r of exactly the same form, but with $r-1$ replaced by r everywhere and r replaced by $r+1$.

The step from $r-1$ to r can be performed looking for a generating function χ_r and a canonical transformation

$$\mathbf{p}^{r-1} = \exp(\varepsilon^r L_{\chi_r})\mathbf{p}^r, \qquad \mathbf{q}^{r-1} = \exp(\varepsilon^r L_{\chi_r})\mathbf{q}^r. \tag{4.102}$$

The new Hamiltonian is immediately found to be

$$H^r(\mathbf{p}^r,\mathbf{q}^r) = [\exp(\varepsilon^r L_{\chi_r})H^{r-1}(\mathbf{p}^{r-1},\mathbf{q}^{r-1})]\Big|_{\mathbf{p}^{r-1}=\mathbf{p}^r,\mathbf{q}^{r-1}=\mathbf{q}^r}, \tag{4.103}$$

that is

$$\begin{aligned} H^r(\mathbf{p}^r,\mathbf{q}^r) =& \boldsymbol{\omega}\cdot\mathbf{p}^r + \varepsilon\bar{H}_1(\mathbf{p}^r) + \ldots + \varepsilon^{r-1}\bar{H}_{r-1}^{(r-2)}(\mathbf{p}^r) + \\ &+ \varepsilon^r H_r^{(r-1)}(\mathbf{p}^r,\mathbf{q}^r) + \varepsilon^r\{\boldsymbol{\omega}\cdot\mathbf{p}^r, \chi_r(\mathbf{p}^r,\mathbf{q}^r)\} + O(\varepsilon^{r+1}). \end{aligned} \tag{4.104}$$

As before, the goal is achieved if χ_r satisfies the homological equation

$$H_r^{(r-1)}(\mathbf{p}^r,\mathbf{q}^r) + \{\boldsymbol{\omega}\cdot\mathbf{p}^r, \chi_r(\mathbf{p}^r,\mathbf{q}^r)\} = \bar{H}_r^{(r-1)}(\mathbf{p}^r). \tag{4.105}$$

It is of crucial importance that the equation to be solved at each step is the same.

Obviously, if $H_r^{(r-1)}$ is a Fourier polynomial, we can always solve the equation (4.105) and χ_r is obviously analytic.

If the Hamiltonian $H_r^{(r-1)}$ has a full Fourier series, we have to distinguish the case in which $\boldsymbol{\omega}$ satisfies the Diophantine condition from the case in which it does not satisfy.

If $\boldsymbol{\omega}$ satisfies the Diophantine condition, we can solve the equation (4.105) and in particular we can prove the following iterative proposition:

Proposition 3. *Let H^{r-1} as in (4.101), with (γ,τ)-Diophantine $\boldsymbol{\omega}$, and assume H^{r-1} is analytical in $D_{(1-(r-1)\eta)(\rho,\sigma)}$ with $\eta < \frac{1}{2r}$. If ε is sufficiently small, namely*

$$\varepsilon^r < \frac{c\eta^{\tau+2}E}{\| H_r^{(r-1)} \|_{\rho,\sigma}} \tag{4.106}$$

with E as in (4.94), then there exists an analytic canonical transformation $S_{\chi_r}^{\varepsilon^r}\colon D_{(1-r\eta)(\rho,\sigma)} \to D_{(1-(r-1)\eta)(\rho,\sigma)}$ such that the new Hamiltonian $H^r = S_{\chi_r}^{\varepsilon^r}H^{r-1}$ has the form (4.101) with r in place of $r-1$ and $r+1$ in place of r.

If $\boldsymbol{\omega}$ does not satisfy the Diophantine condition, as before we decompose the perturbation $H_r^{(r-1)}(\mathbf{p}^r, \mathbf{q}^r)$ as

$$H_r^{(r-1)}(\mathbf{p}^r, \mathbf{q}^r) = H_r^{<K_r}(\mathbf{p}^r, \mathbf{q}^r) + H_r^{\geq K_r}(\mathbf{p}^r, \mathbf{q}^r), \tag{4.107}$$

where $H_r^{<K_r}$ and $H_r^{\geq K_r}$ are defined as in (4.97), with K_r instead of K. In this procedure, K_r must be large enough so that all harmonics of order larger than K_r have coefficients smaller than ε^r. It can be proven that a good choice for K_r is $K_r \geq rK$.

Then, χ_r is chosen in order to eliminate the part of the perturbation of order ε^r, i.e. χ_r is the solution of the equation

$$H_r^{<K_r}(\mathbf{p}^r, \mathbf{q}^r) + \{\boldsymbol{\omega} \cdot \mathbf{p}^r, \chi_r(\mathbf{p}^r, \mathbf{q}^r)\} = \bar{H}_r^{<K_r}(\mathbf{p}^r). \tag{4.108}$$

In this case, there is in general no hope of constructing Hamiltonians H^r in Birkhoff normal form to order ε^r with arbitrarily large r, namely to transform the original Hamiltonian into an integrable $H^\infty(\mathbf{p}^\infty)$.

4.8.1 Practical calculation

In a practical calculation, we truncate the series to some order ε^r, for a chosen order $r \geq 1$, in such a way that all equalities written above for infinite series remain valid up to terms of order $O(\varepsilon^{r+1})$. In this case the Hamiltonian is transformed to the form

$$H^r(\mathbf{p}^r, \mathbf{q}^r) = \boldsymbol{\omega} \cdot \mathbf{p}^r + \varepsilon \bar{H}_1(\mathbf{p}^r) + ... + \varepsilon^r \bar{H}_r^{(r-1)}(\mathbf{p}^r) + O(\varepsilon^{r+1}). \tag{4.109}$$

This should be considered as the *optimal normal form*, in the sense that we have chosen the optimal finite order $r \geq 1$ which minimizes the size of the *non-normalized remainder*.

Thus, if we neglect the remainder, we obtain, less than errors of order $O(\varepsilon^{r+1})$, an integrable approximation of the real dynamics, which is $\mathbf{p}^r =$ constant and $\mathbf{q}^r = \boldsymbol{\omega}^r t + \mathbf{q}^r(0)$, with $\boldsymbol{\omega}^r = \mathrm{grad}_{\mathbf{p}^r}[\boldsymbol{\omega} \cdot \mathbf{p}^r + ... + \varepsilon^r \bar{H}_r^{(r-1)}(\mathbf{p}^r)]$.

In particular, suppose we want to construct the transformed Hamiltonian up to degree r in ε of the truncated Hamiltonian $H = H_0 + \varepsilon H_1 + ... + \varepsilon^r H_r$. With a little attention we realize that it is enough to construct every diagram of the form (4.67) until we reach the line corresponding to the power ε^r, and in particular we need to know only the generating functions $\chi_1, ..., \chi_r$. Moreover, it is sufficient to include all terms up to order ε^r.

Similar considerations apply to the inverse transformation. In particular, let us consider the sequence of generating function $\{\chi_1, ..., \chi_r\}$, and let the operators $S_\chi^{(r)}$ and $\widetilde{S}_\chi^{(r)}$ be defined as in (4.69) an (4.71). Let us calculate the transformation

$$\begin{aligned} q &= S_\chi^{(r)} q' = q' + \varepsilon \varphi_1(q', p') + ... + \varepsilon^r \varphi_r(q', p'), \\ p &= S_\chi^{(r)} p' = p' + \varepsilon \psi_1(q', p') + ... + \varepsilon^r \psi_r(q', p'), \end{aligned} \tag{4.110}$$

up to degree r in ε, where the functions $\varphi_1(q', p'), ..., \varphi_r(q', p')$ and $\psi_1(q', p'), ..., \psi_r(q', p')$ may be explicitly calculated. We may then consider the inverse transformation

$$\begin{aligned} q' &= \widetilde{S}_\chi^{(r)} q = q + \varepsilon \widetilde{\varphi}_1(q, p) + ... + \varepsilon^r \widetilde{\varphi}_r(q, p), \\ p' &= \widetilde{S}_\chi^{(r)} p = p + \varepsilon \widetilde{\psi}_1(q, p) + ... + \varepsilon^r \widetilde{\psi}_r(q, p), \end{aligned} \tag{4.111}$$

which could be explicitly calculated. We can notice that, because we work with truncated series at order r, if we substitute the expressions (4.111) into (4.110) the result is the identity up to a term of order ε^{r+1}. This is the best we can expect in a practical calculation.

Chapter 5

Development of the Hamiltonian in Poincaré variables

To integrate the Hamilton's equations "semi-analytically", using the various tools provided by Hamiltonian's theory and by perturbation theory, we first need to rewrite the classical Hamiltonian (3.22) and the simplified relativistic one (3.66) in Delaunay variables or in Poicaré variables, in order to obtain two quasi-integrable Hamiltonians of the form (4.40).

In this chapter, we present a method for the expansion of the classical and relativistic Hamiltonian in Poincaré variables, which can be implemented in a straightforward manner on a computer. To do this, we follow in particular the works of *Laskar (1989)* and of *Duriez (1989)*.

Because we are interested in the problem of three bodies, we use the conservation of the total angular momentum to simplify the problem.

Moreover, because we are concerned with the secular variations of the eccentricities, we can simplify the problem using the averaging principle. Indeed, the averaging makes it possible to reduce the number of degrees of freedom, because averaged Hamiltonians don't depend on the fast angles M_i (for $i = 1, 2$) and the conjugate momenta to the fast angles become first integrals of the secular problem, i.e. the semi-major axes a_i are constants. In particular, if there are no resonances between the mean motion frequencies of the planets, the use of the averaging principle is justified by Poisson's theorem (see section 6.1). Thus, assuming that no strong mean motion resonances are present, we can obtain qualitative information on the long-term changes of the eccentricities.

5.1 Hamilton equations in Poincaré variables

The problem is to rewrite the Hamiltonians (3.22) and (3.66) in modified Delaunay's variables[1]

$$\begin{aligned} \Lambda_k &= \mu_k \sqrt{\mathcal{G}(m_0 + m_k) a_k} & \lambda_k &= M_k + \omega_k + \Omega_k \\ P_k &= \Lambda_k (1 - \sqrt{1 - e_k^2}) & p_k &= -\omega_k - \Omega_k = -\varpi_k \\ Q_k &= 2\Lambda_k \sqrt{1 - e_k^2} \sin^2 i_k & q_k &= -\Omega_k. \end{aligned} \tag{5.1}$$

[1] We have constructed the action-angle variables using the Hamiltonian in the form (4.22) rather than (4.19). Had we started the construction of action-angle variables from (4.22), the resulting actions would have been multiplied by μ_k with respect to those defined in (4.37)-(4.39), and the resulting Hamiltonian would have been multiplied by μ_k^3 with respect to (4.28).

or in Poincaré variables

$$
\begin{aligned}
\Lambda_k &= \mu_k\sqrt{\mathcal{G}(m_0+m_k)a_k} & \lambda_k &= M_k+\varpi_k \\
\eta_k &= \sqrt{2\Lambda_k}\sqrt{1-\sqrt{1-e_k^2}}\cos\varpi_k & \xi_k &= -\sqrt{2\Lambda_k}\sqrt{1-\sqrt{1-e_k^2}}\sin\varpi_k \\
\eta_{2k} &= \sqrt{2\Lambda_k}\sqrt{\sqrt{1-e_k^2}(1-\cos i_k)}\cos\Omega_k & \xi_{2k} &= -\sqrt{2\Lambda_k}\sqrt{\sqrt{1-e_k^2}(1-\cos i_k)}\sin\Omega_k
\end{aligned}
\tag{5.2}
$$

for $k = 1, 2$.

In Poincaré variables, the Hamiltonians (3.22) and (3.66) become

$$
\begin{aligned}
H_{\text{new}} &= \sum_{j=1}^{2} -\frac{\mathcal{G}^2(m_0+m_j)^2\mu_j^3}{2\Lambda_j^2} + \varepsilon H_1(\boldsymbol{\Lambda},\boldsymbol{\lambda},\boldsymbol{\eta},\boldsymbol{\xi}) \\
H_{\text{rel}} &= \sum_{j=1}^{2} -\frac{\mathcal{G}^2(m_0+m_j)^2\mu_j^3}{2\Lambda_j^2} + \varepsilon H_1(\boldsymbol{\Lambda},\boldsymbol{\lambda},\boldsymbol{\eta},\boldsymbol{\xi}) + \frac{1}{c^2}H_2(\boldsymbol{\Lambda},\boldsymbol{\lambda},\boldsymbol{\eta},\boldsymbol{\xi}),
\end{aligned}
\tag{5.3}
$$

where H_1 and H_2 can be explicitly written as a function of the Poincaré variables (5.2) by direct substitution in (3.22) and (3.66). Then, Hamilton's equations will be:

$$
\begin{aligned}
\dot{\Lambda}_k &= -\varepsilon\frac{\partial H_1}{\partial \lambda_k} & \dot{\lambda}_k &= \frac{\mathcal{G}^2(m_0+m_j)^2\mu_j^3}{\Lambda_j^3} + \varepsilon\frac{\partial H_1}{\partial \Lambda_k} \\
\dot{\eta}_k &= -\varepsilon\frac{\partial H_1}{\partial \xi_k} & \dot{\xi}_k &= \varepsilon\frac{\partial H_1}{\partial \eta_k} \\
\dot{\eta}_{2k} &= -\varepsilon\frac{\partial H_1}{\partial \xi_{2k}} & \dot{\xi}_{2k} &= \varepsilon\frac{\partial H_1}{\partial \eta_{2k}},
\end{aligned}
\tag{5.4}
$$

and in the relativistic case

$$
\begin{aligned}
\dot{\Lambda}_k &= -\varepsilon\frac{\partial H_1}{\partial \lambda_k} - \frac{1}{c^2}\frac{\partial H_2}{\partial \lambda_k} & \dot{\lambda}_k &= \frac{\mathcal{G}^2(m_0+m_j)^2\mu_j^3}{2\Lambda_j^3} + \varepsilon\frac{\partial H_1}{\partial \Lambda_k} + \frac{1}{c^2}\frac{\partial H_2}{\partial \Lambda_k} \\
\dot{\eta}_k &= -\varepsilon\frac{\partial H_1}{\partial \xi_k} - \frac{1}{c^2}\frac{\partial H_2}{\partial \xi_k} & \dot{\xi}_k &= \varepsilon\frac{\partial H_1}{\partial \eta_k} + \frac{1}{c^2}\frac{\partial H_2}{\partial \eta_k} \\
\dot{\eta}_{2k} &= -\varepsilon\frac{\partial H_1}{\partial \xi_{2k}} - \frac{1}{c^2}\frac{\partial H_2}{\partial \xi_{2k}} & \dot{\xi}_{2k} &= \varepsilon\frac{\partial H_1}{\partial \eta_{2k}} + \frac{1}{c^2}\frac{\partial H_2}{\partial \eta_{2k}},
\end{aligned}
\tag{5.5}
$$

for $k = 1, 2$.

Obviously, in these Hamiltonians, λ_k plays the role of the fast angles, because $\dot{\lambda}_k = O(1)$, while the argument of perihelion and the longitudes of node are slow angles, because $\dot{\Lambda}_k, \dot{\xi}_k, \dot{\xi}_{2k}, \dot{\eta}_k, \dot{\eta}_{2k} = O(\varepsilon)$. Indeed, in the planetary system with a dominant stellar mass, the orbits slowly rotate due to mutual interactions and to relativistic effects, and the real motion is distinguishable from the Keplerian one only on observation conducted for times of the order of centuries.

In the following, when we do not want to distinguish the classical case from the relativistic one, we will refer for simplicity to a Hamiltonian function of the form

$$
H = \sum_{j=1}^{2} -\frac{\mathcal{G}^2(m_0+m_j)^2\mu_j^3}{2\Lambda_j^2} + H_\epsilon(\boldsymbol{\Lambda},\boldsymbol{\lambda},\boldsymbol{\eta},\boldsymbol{\xi}),
\tag{5.6}
$$

where H_ϵ depends on the case, i.e. H_ϵ is equal to εH_1 in the classical case and it is equal to $\varepsilon H_1 + \frac{1}{c^2}H_2$ in the relativistic case.

5.1.1 D'Alembert rules

In the classical studies of the dynamics of planetary problems, the function H_ϵ is usually expanded in Fourier series of the angles $\boldsymbol{\lambda}$, $\mathbf{p}$ and $\mathbf{q}$ and in power series of $\mathbf{P}^{1/2}$ and $\mathbf{Q}^{1/2}$, or equivalently, using the Poincaré variables, in Fourier series of $\boldsymbol{\lambda}$ and in power series of $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$. Let us denote by $\Lambda_j, P_j, Q_j, \lambda_j, p_j, q_j$ $(j = 1, 2)$ the modified Delaunay variables of the 2 bodies and by $\alpha_j, \beta_j, k_j, m_j, s_j$ integer numbers. Moreover, we generically denote by $\boldsymbol{\Lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{k}, \mathbf{m}$ and $\mathbf{s}$ the vectors whose components are $\Lambda_j, \alpha_j, \beta_j, k_j m_j$ and s_j respectively. The most general form of the Fourier series expansion in the angles and power series expansion in $P_j^{1/2}, Q_j^{1/2}$ is therefore:

$$H_\epsilon = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{Z}_+^2} \sum_{\mathbf{k},\mathbf{m},\mathbf{s}\in\mathbb{Z}^2} c_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{k},\mathbf{m},\mathbf{s}}(\boldsymbol{\Lambda}) \left(\prod_{j=1}^{2} P_j^{\alpha_j/2} Q_j^{\beta_j/2} \right) \exp[\sqrt{-1}(\boldsymbol{\lambda}\cdot\mathbf{k} + \mathbf{p}\cdot\mathbf{m} + \mathbf{q}\cdot\mathbf{s})] \tag{5.7}$$

where $c_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{k},\mathbf{m},\mathbf{s}}(\boldsymbol{\Lambda})$ are suitable coefficients and $\boldsymbol{\lambda}\cdot\mathbf{k} = \sum_j \lambda_j k_j$.

With obvious notations, in Poincaré variables H_ϵ has the form

$$H_\epsilon = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{Z}_+^4} \sum_{\mathbf{k}\in\mathbb{Z}^2} c_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{k}}(\boldsymbol{\Lambda}) \left(\prod_{j=1}^{4} \xi_j^{\alpha_j/2} \eta_j^{\beta_j/2} \right) \exp[\sqrt{-1}(\boldsymbol{\lambda}\cdot\mathbf{k})] \tag{5.8}$$

where $c_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{k}}(\boldsymbol{\Lambda})$ are suitable coefficients.

In reality, these trigonometric series, resulting from the developments, do not contain all possible combinations of angles: consideration of the symmetries and analytic properties of H_ϵ allow the easily derivation of the so-called *D'Alembert rules*. In the case of modified Delaunay variables, they are:

1. H_ϵ must be invariant under a simultaneous change of sign of all the angles λ_j, p_j, q_j; therefore the Fourier series expansion must contain only cosine terms, namely $c_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{k},\mathbf{m},\mathbf{s}} = c_{\boldsymbol{\alpha},\boldsymbol{\beta},-\mathbf{k},-\mathbf{m},-\mathbf{s}}$ and all coefficients are real;

2. H_ϵ must be invariant under an arbitrary rotation of the reference frame around the z axis. A rotation of the reference frame by an angle θ increments the longitudes λ_j, p_j, q_j by θ. Because $p_j = -\varpi_j$ and $q_j = -\Omega_j$, the invariance of H_ϵ implies that $\sum_j (k_j - m_j - s_j) = 0$;

3. H_ϵ must be invariant under a simultaneous change of sign of all inclinations, namely by a transformation $Q_j^{1/2} \to -Q_j^{1/2}$, $\forall j$. This implies that $\sum_j \beta_j/2$ must be an integer number;

4. H_ϵ must be analytical function of the variables $\boldsymbol{\xi}, \boldsymbol{\eta}$ in a neighborhood of $\boldsymbol{\xi} = \boldsymbol{\eta} = \mathbf{0}$. This implies $|m_j| \leq \alpha_j$ and $|s_j| \leq \beta_j$ for each $j = 1, 2$ and, moreover, m_j has the same parity of α_j and s_j has the same parity of β_j, i.e. m_j can only take the values $-\alpha_j, -\alpha_j + 2, ..., \alpha_j - 2, \alpha_j$ and s_j can only take the values $-\beta_j, -\beta_j + 2, ..., \beta_j - 2, \beta_j$.

In the case of Poincaré variables, they are:

1. the coefficients $c_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{k}}$ are complex and they satisfy the condition $c_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{k}} = \bar{c}_{\boldsymbol{\alpha},\boldsymbol{\beta},-\mathbf{k}}$; therefore the Fourier series expansion must contain only sine terms if $|\boldsymbol{\beta}|$ is odd and only cosine terms if $|\boldsymbol{\beta}|$ is even;

2. the sum $\alpha_3 + \alpha_4 + \beta_3 + \beta_4$ must be an even number;

3. the sum $k_1 + k_2$ satisfy the condition $k_1 + k_2 \leq \alpha_1 + ... + \alpha_4 + \beta_1 + ... + \beta_4$ and the two sums must have the same parity.

I emphasize again that D'Alembert rules apply both in the classical case that in the relativistic one, because they are obtained only by considerations of the symmetries and analytic properties of H_ϵ.

5.2 Conservation of the angular momentum and the reduced three-body problem

We know that in classical mechanics, for a closed system, in addition to conservation of energy and total linear momentum, there is the conservation of the total angular momentum. It is simple to prove that also in general relativity the total angular momentum is conserved: this is a consequence of the fact that, because of the isotropy of the space, the Lagrangian of a closed system does not change under the rotation of the system as a whole.

Let $\mathbf{C}$ the angular momentum of a system of three bodies in the heliocentric reference system (see (3.18) - we remember that $\mathbf{p}_0 = \mathbf{0}$):

$$\mathbf{C} = [C_x, C_y, C_z] = \mathbf{C}_1 + \mathbf{C}_2 = \mathbf{r}_1 \times \mathbf{p}_1 + \mathbf{r}_2 \times \mathbf{p}_2. \tag{5.9}$$

Using the Delaunay variables

$$\begin{aligned} L_k &= \mu_k \sqrt{\mathcal{G}(m_0 + m_k) a_k} & l_k &= M_k \\ G_k &= L_k \sqrt{1 - e_k^2} & g_k &= \omega_k \\ \mathcal{H}_k &= G_k \cos i_k & h_k &= \Omega_k, \end{aligned} \tag{5.10}$$

for $k = 1, 2$, it is simple to prove that the angular momentum is equal to[2]

$$C_x = \sum_{j=1}^{2} \sqrt{G_j^2 - \mathcal{H}_j^2} \sin h_j, \qquad C_y = -\sum_{j=1}^{2} \sqrt{G_j^2 - \mathcal{H}_j^2} \cos h_j, \qquad C_z = \sum_{j=1}^{2} \mathcal{H}_j. \tag{5.11}$$

In the case of a system of three bodies and in the heliocentric reference system, the conservation of the angular momentum induces two basic properties (see *Laskar (1989)* and *Laskar and Roboutel (1995)* for more details).

The first property induced by the conservation of the angular momentum is that the ascending node of the two planets have a common motion on the *invariant plane* (plane perpendicular to the angular momentum vector and includes the central star). Indeed, let $\mathbf{C}$ the total angular momentum in the heliocentric reference system (i.e. $\mathbf{C} = \mathbf{r}_1 \times \mathbf{p}_1 + \mathbf{r}_2 \times \mathbf{p}_2$) and let $\mathbf{V}$ the vector directed toward the intersection of the orbits of the two bodies, as shown in Fig. 5.1. Because $\mathbf{V}$ lies both in the plane $(\mathbf{r}_1, \mathbf{p}_1)$ and in the plane $(\mathbf{r}_2, \mathbf{p}_2)$, we have that the vector $\mathbf{V}$ is perpendicular to the angular momentum $\mathbf{C}$:

$$\mathbf{V} \cdot \mathbf{C} = \mathbf{V} \cdot (\mathbf{r}_1 \times \mathbf{p}_1) + \mathbf{V} \cdot (\mathbf{r}_2 \times \mathbf{p}_2) = 0. \tag{5.12}$$

[2] Remember that $G_j = \| \mathbf{C}_j \|$ and that the orbital plane of the j-th body is perpendicular to $\mathbf{C}_j$, for $j = 1, 2$.

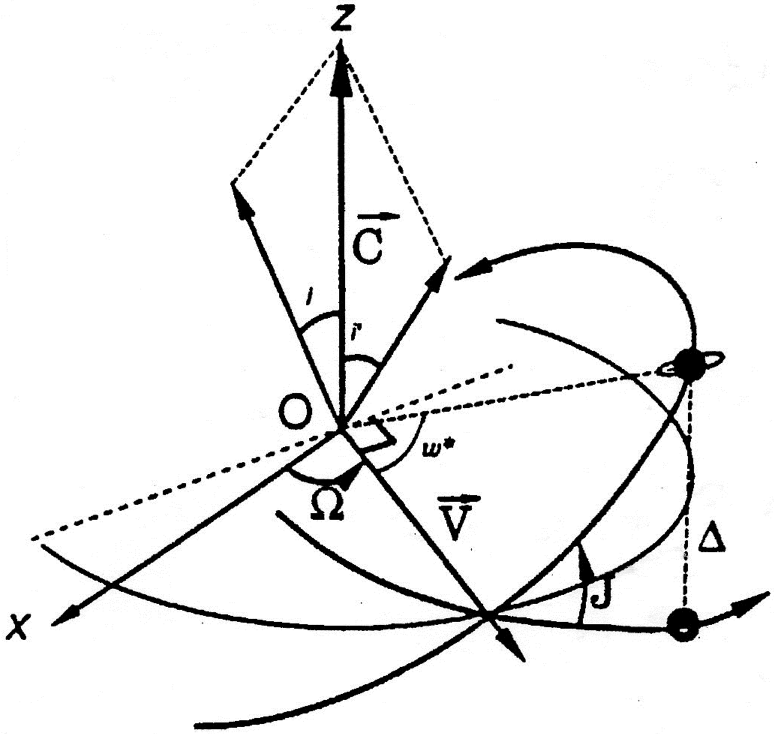


Figure 5.1: Invariant plane. Reprinted from Fig. 2 of Laskar, *Systèmes de variables et éléments.*

Because $\mathbf{C}$ is constant, then $\mathbf{V}(t)$ always remains in the plane perpendicular to $\mathbf{C}$. Thus, by definition of $\mathbf{V}$, this implies that the intersection of the two orbits remains always on the same plane perpendicular to the angular momentum $\mathbf{C}$. If we choose this plane as the reference plane, we have that the two points of intersection of the orbits coincide with the ascending node of the two planets, and then the thesis.

The second point is that the Hamiltonian depends on the longitudes of the nodes uniquely by their difference (in fact it is invariant under an arbitrary rotation of the reference frame around the z axis), i.e. $H = H_0(\mathbf{\Lambda}) + H_\epsilon(\mathbf{\Lambda}, \boldsymbol{\lambda}, \mathbf{G}, \mathbf{g}, \boldsymbol{\mathcal{H}}, h_2 - h_1)$. Then, if the osculating elements are defined with respect to the invariant plane, the difference of the longitude of the node is always equal to the constant π (i.e. $|h_1 - h_2| = \pi$), and the nodes vanish from the Hamiltonian.

The components of the angular momentum in the plane perpendicular to the vector $\mathbf{V}$ are given by

$$\begin{aligned} G_1 \cos i_1 + G_2 \cos i_2 &= C \\ G_1 \sin i_1 + G_2 \sin i_2 &= 0, \end{aligned} \tag{5.13}$$

where $C = \| \mathbf{C} \|$ is constant. The equations (5.13) can be rewritten as

$$\begin{aligned} \mathcal{H}_1 + \mathcal{H}_2 &= C \\ G_1^2 - \mathcal{H}_1^2 &= G_2^2 - \mathcal{H}_2^2 \end{aligned} \tag{5.14}$$

from which we can derive the following relationship

$$\begin{aligned} \mathcal{H}_1 &= \frac{C}{2} + \frac{G_1^2}{2C} - \frac{G_2^2}{2C} \\ \mathcal{H}_2 &= \frac{C}{2} - \frac{G_1^2}{2C} + \frac{G_2^2}{2C}. \end{aligned} \tag{5.15}$$

We now consider the following change of coordinates

$$\begin{aligned} \mathcal{H}_1 \to \Psi_1 &= \mathcal{H}_1 + \mathcal{H}_2, \qquad h_1 \to \psi_1 = \frac{h_1 + h_2}{2}, \\ \mathcal{H}_2 \to \Psi_2 &= \mathcal{H}_2 - \mathcal{H}_1, \qquad h_1 \to \psi_2 = \frac{h_2 - h_1}{2}, \end{aligned} \tag{5.16}$$

which is a canonical transformation. Without loss of generality, we can put $h_2 = h_1 + \pi$ and we obtain

$$\Psi_1 = C, \qquad \psi_1 = h_1 + \frac{\pi}{2}, \qquad \psi_2 = \frac{\pi}{2} \tag{5.17}$$

where Ψ_1 and ψ_2 are two constants of motion. The equations of motion become

$$\begin{aligned} &\Psi_1(t) = C, && \frac{\mathrm{d}\psi_1(t)}{\mathrm{d}t} = \frac{\partial H(\mathbf{L}, \boldsymbol{\lambda}, \mathbf{G}, \mathbf{g}, \Psi_1, \psi_2)}{\partial \Psi_1}, \\ &\frac{\mathrm{d}\Psi_2(t)}{\mathrm{d}t} = -\frac{\partial H(\mathbf{L}, \boldsymbol{\lambda}, \mathbf{G}, \mathbf{g}, \Psi_1, \psi_2)}{\partial \psi_2}, && \psi_2(t) = \frac{\pi}{2}. \end{aligned} \tag{5.18}$$

Thus H depends only on the variables $L_1, L_2, l_1, l_2, G_1, G_2, g_1, g_2$, because ψ_2 and Ψ_2 are constant and H is independent of ψ_1 and Ψ_2 (indeed $\dot{\psi}_2 = \frac{\partial H}{\partial \Psi_2} = 0$). In this way, we have reduced of 2 the number of degrees of freedom of the system and we obtain an Hamiltonian system which possesses 4 degrees of freedom.

Thus, we can introduce the modified planar Delaunay's variables

$$\begin{aligned} \Lambda_k &= L_k & \lambda_k^\star &= M_k + \omega_k = l_k + g_k \\ P_k &= L_k - G_k & p_k^\star &= -\omega_k = -g_k = -\varpi_k^\star, \end{aligned} \tag{5.19}$$

where the orbital elements $(M_j, \omega_j, \Omega_j)$ are obviously defined with respect to the invariant plane and not to an arbitrary reference frame.

We can also introduce the planar Poincaré variables

$$\begin{aligned} \Lambda_k &= \mu_k \sqrt{\mathcal{G}(m_0 + m_k) a_k}, \qquad \lambda_k^\star = M_k + \omega_k = l_k + \varpi_k^\star \\ \xi_k^\star &= \sqrt{2P_k} \cos p_k^\star = \sqrt{2\Lambda_k} \sqrt{1 - \sqrt{1 - e_k^2}} \cos \omega_k \\ \eta_k^\star &= \sqrt{2P_k} \sin p_k^\star = -\sqrt{2\Lambda_k} \sqrt{1 - \sqrt{1 - e_k^2}} \sin \omega_k. \end{aligned} \tag{5.20}$$

The equations of motion become in the classical case:

$$\begin{aligned} \dot{\Lambda}_k &= -\varepsilon \frac{\partial H_1}{\partial \lambda_k^\star} & \dot{\lambda}_k^\star &= \frac{\mathcal{G}^2 (m_0 + m_k)^2 \mu_k^3}{\Lambda_k^3} + \varepsilon \frac{\partial H_1}{\partial \Lambda_k} \\ \dot{\eta}_k^\star &= -\varepsilon \frac{\partial H_1}{\partial \xi_k^\star} & \dot{\xi}_k^\star &= \varepsilon \frac{\partial H_1}{\partial \eta_k^\star}, \end{aligned} \tag{5.21}$$

and in the relativistic case:

$$
\begin{aligned}
\dot{\Lambda}_k &= -\varepsilon\frac{\partial H_1}{\partial \lambda_k^\star} - \frac{1}{c^2}\frac{\partial H_2}{\partial \lambda_k^\star} & \dot{\lambda}_k^\star &= \frac{\mathcal{G}^2(m_0+m_k)^2\mu_k^3}{\Lambda_k^3} + \varepsilon\frac{\partial H_1}{\partial \Lambda_k} + \frac{1}{c^2}\frac{\partial H_2}{\partial \Lambda_k^\star} \\
\dot{\eta}_k^\star &= -\varepsilon\frac{\partial H_1}{\partial \xi_k^\star} - \frac{1}{c^2}\frac{\partial H_2}{\partial \xi_k^\star} & \dot{\xi}_k^\star &= \varepsilon\frac{\partial H_1}{\partial \eta_k^\star} + \frac{1}{c^2}\frac{\partial H_2}{\partial \eta_k^\star},
\end{aligned} \tag{5.22}
$$

for $k = 1, 2$, where $H = H(\boldsymbol{\Lambda}, \boldsymbol{\lambda}^\star, \boldsymbol{\eta}^\star, \boldsymbol{\xi}^\star; C)$.

For the following, it is also useful to introduce a complex form of the planar Poincaré variables. Given two canonical variables (a, b), to obtain a complex form we can use the following canonical transformation

$$
(a, b) \to \left(\frac{1}{\sqrt{2}}(a - \sqrt{-1}b), -\frac{\sqrt{-1}}{\sqrt{2}}(a + \sqrt{-1}b)\right). \tag{5.23}
$$

Using (5.23) in the case of planar Poincaré variables, we obtain the complex variables $(\chi_k, -\sqrt{-1}\bar{\chi}_k)$ defined as

$$
\chi_k = \frac{1}{\sqrt{2}}(\eta_k - \sqrt{-1}\xi_k) = \sqrt{\Lambda_k}\sqrt{1 - \sqrt{1-e^2}}\exp(\sqrt{-1}\varpi_k^\star). \tag{5.24}
$$

Finally, if the two inclinations are zero (i.e. $i_1 = i_2 = 0$), then the orbital plane coincides with the plane (x, y) and the planar Poincaré variables (5.20) coincide with the Poincaré variables (5.2), i.e.

$$
\lambda_k^\star = \lambda_k, \qquad \varpi_k^\star = \varpi_k, \qquad \xi_k^\star = \xi_k, \qquad \eta_k^\star = \eta_k. \tag{5.25}
$$

5.3 Development of the classical Hamiltonian in the reduced problem

We will present now a method for the computation of the expansion of the Hamiltonian in the planar Poincaré variables (see *Laskar (1989)* and *Duriez (1989)* for more details). In particular, in the classical case, we are interested in developing the two terms

$$
\frac{m_1 m_2}{\| \mathbf{r}_1 - \mathbf{r}_2 \|}, \qquad \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{m_0}. \tag{5.26}
$$

To do this, we use the *osculating orbits* introduced in chapter 3. I only remember that the osculating orbit of an object in space at a given moment in time is the gravitational Kepler orbit (i.e. the ellipse in heliocentric coordinate) that it would have about its central body if perturbations were not present, i.e. the orbit that coincides with the current orbital state vectors (position and velocity).

5.3.1 Development of the Keplerian movement

We consider a planet P whose movement around the center O is described by a Keplerian ellipse. We denote by r the distance between P and O. The aim of this section is to develop r, ν and $r^n \exp(\sqrt{-1}m\nu)$ as functions of M.

Development of r and ν

In the two-body problem, the elliptical motion appears to be periodic with period 2π respect to the angle M (the mean anomaly): this allows us to develop the distance r and the true anomaly ν in Fourier series of M.

I remember that the relationship between r and the eccentric anomaly E is

$$\frac{a}{r} = \frac{1}{1 - e\cos E}. \tag{5.27}$$

The function on the right is periodic in M of period 2π (E and M are angles) and it is an even and smooth function. Thus it can be expanded in its Fourier series

$$\frac{1}{1 - e\cos E} = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(kM), \tag{5.28}$$

where

$$a_k = \frac{1}{\pi}\int_0^{2\pi} \frac{\cos(kl)}{1 - e\cos E}\mathrm{d}l, \qquad \forall k \geq 0. \tag{5.29}$$

With a simple change of variable, we obtain

$$a_k = \frac{1}{\pi}\int_0^{2\pi} \cos[k(E - e\sin E)]\mathrm{d}E = 2J_k(ke), \qquad \forall k \geq 0, \tag{5.30}$$

where $J_k(x)$ is the *Bessel function* of order k (see (1.28)). Thus, the Fourier series of a/r is given by

$$\frac{a}{r} = 1 + \sum_{k>0} J_k(ke)\cos(kM). \tag{5.31}$$

It can been proven that this series converges uniformly for $e < 0.66274341...$ (*Tisserand, 1888*).

In a similar way, we can derive the following developments (for more details see *Brouwer and Clemence, 1961*)

$$\begin{aligned}
\frac{r}{a} &= 1 + \frac{e^2}{2} - 2e\sum_{k=1}^{\infty}\frac{1}{k^2}\frac{\mathrm{d}J_k(ke)}{\mathrm{d}e}\cos(kM) \\
\cos\nu &= -e + \frac{2(1-e^2)}{2}\sum_{k=1}^{\infty} J_k(ke)\cos(kM) \\
\sin\nu &= 2\sqrt{1-e^2}\sum_{k=1}^{\infty}\frac{1}{k}\frac{\mathrm{d}J_k(ke)}{\mathrm{d}e}\sin(kM).
\end{aligned} \tag{5.32}$$

If we introduce the variables

$$X = e\exp(\sqrt{-1}M), \qquad \bar{X} = e\exp(-\sqrt{-1}M), \tag{5.33}$$

it is simple to prove that

$$\frac{a}{r} = 1 + R, \qquad R = \sum_{\substack{k>0 \\ h\geq 0}} C_{k,h}(X^k + \bar{X}^k)(X\bar{X})^h \quad \text{and} \quad C_{h,k} = \frac{(-1)^h (k/2)^{k+2h}}{h!(h+k)!}. \tag{5.34}$$

If the eccentricity is small enough, we can truncate the developments at a not too high degree d:

$$\frac{a}{r} = 1 + R, \qquad R = \sum_{k=1}^{d} \sum_{h=0}^{\lfloor (d-k)/2 \rfloor} C_{k,h}(X^k + \bar{X}^k)(X\bar{X})^h + o(e^{d+1}), \tag{5.35}$$

where $\lfloor x \rfloor$ indicates the integer part of x.

Starting from the expression[3] (5.32), to get r/a to the same form of (5.35), we have to introduce appropriate coefficients $C'_{k,h}$ such that

$$\frac{r}{a} = \sum_{k=0}^{d} \sum_{h=0}^{\lfloor (d-k)/2 \rfloor} C'_{k,h}(X^k + \bar{X}^k)(X\bar{X})^h + o(e^{d+1}). \tag{5.36}$$

To express the true anomaly ν, we start from the following expression

$$\frac{\mathrm{d}\nu}{\mathrm{d}M} = \frac{a^2}{r^2}\sqrt{1-e^2} \tag{5.37}$$

which can be rewritten as

$$\frac{\mathrm{d}\nu}{\mathrm{d}M} = (1 + 2R + R^2)(1 - \frac{1}{2}X\bar{X} - \frac{1}{8}X^2\bar{X}^2 + ...). \tag{5.38}$$

Introducing appropriate coefficients $C"_{k,h}$, we obtain

$$\frac{\mathrm{d}\nu}{\mathrm{d}M} = 1 + \sum_{k=1}^{d} \sum_{h=0}^{\lfloor (d-k)/2 \rfloor} C"_{k,h}(X^k + \bar{X}^k)(X\bar{X})^h + o(e^{d+1}), \tag{5.39}$$

which gives:

$$\sqrt{-1}(\nu - M) = \sum_{k=1}^{d} \sum_{h=0}^{\lfloor (d-k)/2 \rfloor} \frac{1}{k} C"_{k,h}(X^k - \bar{X}^k)(X\bar{X})^h + o(e^{d+1}). \tag{5.40}$$

Development of $r^n \exp(\sqrt{-1}m(\nu - M))$

Let

$$\theta = \exp(\sqrt{-1}(\nu - M)). \tag{5.41}$$

To obtain the development of θ in X and $\bar{X}$, it is sufficient to combine the development (5.40) with the development of the exponential function in Taylor series.

We can then write the development of $\left(\frac{r}{a}\right)^n \theta^m$ as

$$\left(\frac{r}{a}\right)^n \theta^m = \sum_{k=-\infty}^{+\infty} X_k^{n,m}(e) \exp(\sqrt{-1}(k-m)M), \tag{5.42}$$

[3] We can deduce r/a in a different way. Starting from the expression (5.34), r/a is equal to

$$\frac{r}{a} = \frac{1}{1+R} = \sum_{j=0}^{d} (-1)^j R^j + o(e^{d+1})$$

where $X_k^{n,m}(e)$ are the Hansen coefficients (these coefficients are defined in Appendix B).

Introducing appropriate coefficients $C_{k,\bar{k}}^{n,m}$, we can write $\left(\frac{r}{a}\right)^n \theta^m$ in the form

$$\begin{aligned}\left(\frac{r}{a}\right)^n \theta^m &= \sum_{k=-d}^{d} X^k \sum_{h=0}^{\lfloor (d-|k|)/2 \rfloor} C_{k,\bar{k}}^{n,m} (X\bar{X})^h + o(e^{d+1}) = \\ &= \sum_{0 \le k+\bar{k} \le d} C_{k,\bar{k}}^{n,m} X^k \bar{X}^{\bar{k}} + o(e^{d+1}),\end{aligned} \tag{5.43}$$

where $X^k \equiv \bar{X}^{|k|}$ if $k < 0$.

From the development (5.43), it is easy to pass to Poincaré coordinates. To do this, we introduce the variables

$$z = e \exp(\sqrt{-1}\varpi^\star), \qquad \bar{z} = e \exp(-\sqrt{-1}\varpi^\star). \tag{5.44}$$

Using $M = \lambda^\star - \varpi^\star$, we obtain

$$X = e \exp(\sqrt{-1}M) = e \exp(\sqrt{-1}(\lambda^\star - \varpi^\star)) = \bar{z} \exp(\sqrt{-1}\lambda^\star) \tag{5.45}$$

and

$$\left(\frac{r}{a}\right)^n \theta^m = \sum_{0 \le k+\bar{k} \le d} C_{k,\bar{k}}^{n,m} \bar{z}^k z^{\bar{k}} \exp(\sqrt{-1}\lambda^\star(k - \bar{k})) + o(e^{d+1}). \tag{5.46}$$

Finally, using the planar complex Poincaré variables (5.24), the relationship between z and χ is

$$z = \chi \sqrt{\frac{\Lambda}{2}} \sqrt{1 - \frac{\chi\bar{\chi}}{2\Lambda}} = \chi \sqrt{\frac{\Lambda}{2}} \left(1 - \frac{\chi\bar{\chi}}{4\Lambda} - \frac{(\chi\bar{\chi})^2}{32\Lambda^2} + ...\right) \tag{5.47}$$

and replacing the expression (5.47) in (5.46), we obtain the desired development.

5.3.2 Development of the inverse of the distance

We consider two planets P_1 and P_2 whose movement around the center O is described by the osculating orbits represented by classical orbital elements $(a_k, e_k, i_k, M_k, \omega_k, \Omega_k)$ in a reference system $R_O = (O; \mathbf{i}_O, \mathbf{j}_O, \mathbf{k}_O)$. We denote by $\mathbf{r}_1 = r_1 \mathbf{u}_1$ and $\mathbf{r}_2 = r_2 \mathbf{u}_2$ the rays vector joining O respectively with P_1 and P_2, where obviously $r_1 = \parallel \mathbf{r}_1 \parallel$ and $r_2 = \parallel \mathbf{r}_2 \parallel$, and S is the angle between the vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, as Fig. 5.2 shows.

Then we have

$$\frac{1}{\Delta} = \frac{1}{\parallel \mathbf{r}_1 - \mathbf{r}_2 \parallel} = (r_1^2 + r_2^2 - 2 r_1 r_2 \cos S)^{-1/2}. \tag{5.48}$$

Without loss of generality, if we suppose that $r_1 < r_2$, we can introduce the quantities ρ and α less than 1:

$$\rho = \frac{r_1}{r_2}, \qquad \alpha = \frac{a_1}{a_2}, \tag{5.49}$$

and we obtain

$$\frac{1}{\Delta} = \frac{1}{r_2}(1 + \rho^2 - 2\rho \cos S)^{-1/2}. \tag{5.50}$$

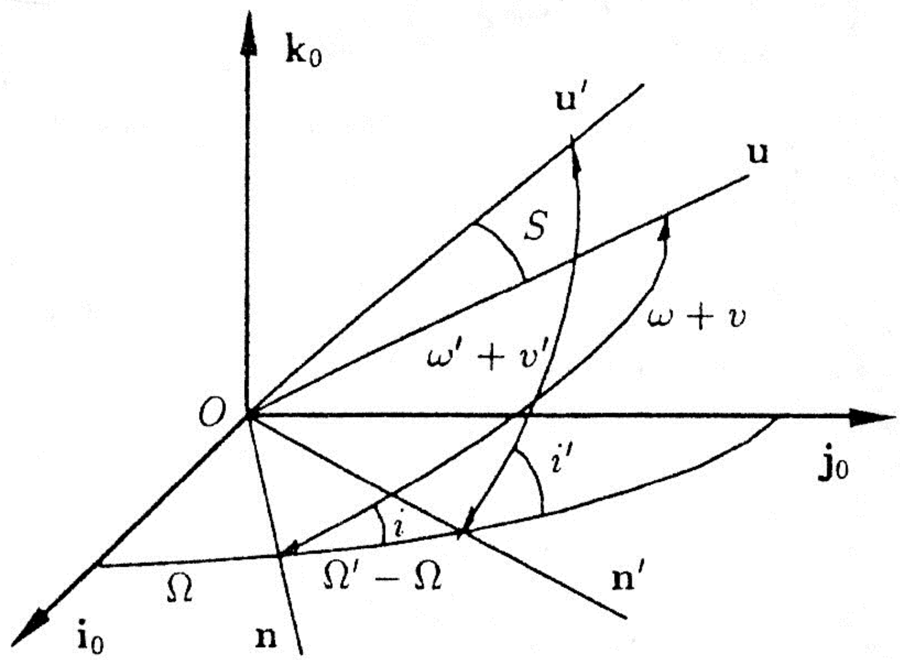


Figure 5.2: Reprinted from Fig.1 of Duriez, *Le Développment de la Fonction Perturbatrice.*

Development of ρ^n

Writing ρ as

$$\rho = \alpha \frac{r_1}{a_1} \frac{a_2}{r_2} \tag{5.51}$$

and using (5.43), the development of ρ^n in the 4 variables $X_1, \bar{X}_1, X_2$ and $\bar{X}_2$ is

$$\rho^n = \alpha^n \sum_{0 \le k+\bar{k}+k'+\bar{k}' \le d} C^{n,0}_{k,\bar{k}} C^{-n,0}_{k',\bar{k}'} X_1^k \bar{X}_1^{\bar{k}} X_2^{k'} \bar{X}_2^{\bar{k}'} + o(e^{d+1}). \tag{5.52}$$

In a similar way, using z_k and $\bar{z}_k$ instead of X_k and $\bar{X}_k$, we obtain

$$\rho^n = \alpha^n \sum_{0 \le k+\bar{k}+k'+\bar{k}' \le d} C^{n,0}_{k,\bar{k}} C^{-n,0}_{k',\bar{k}'} \bar{z}_1^k z_1^{\bar{k}} \bar{z}_2^{k'} z_2^{\bar{k}'} \exp[\sqrt{-1}((k-\bar{k})\lambda_1^\star + (k'-\bar{k}')\lambda_2^\star)] + o(e^{d+1}). \tag{5.53}$$

Finally, replacing the expression (5.47) in (5.53), we obtain the development of ρ^n in the planar Poincaré variables.

Development of $\cos S$

As shown in Fig. 5.2, the orbits lie on the two orbital planes $O\mathbf{n}_1\mathbf{u}_1$ and $O\mathbf{n}_2\mathbf{u}_2$ with respect to R_O (with respect to which the orbital elements are defined). In particular $\mathbf{n}_1$ and $\mathbf{n}_2$ are directed towards the two ascending nodes. The angle between $\mathbf{n}_1$ and $\mathbf{n}_2$ is $\Omega_1 - \Omega_2$ and the orthogonal bases $(\mathbf{n}_1, \mathbf{k}_O \times \mathbf{n}_1, \mathbf{k}_O)$ and $(\mathbf{n}_2, \mathbf{k}_O \times \mathbf{n}_2, \mathbf{k}_O)$ differ from each other by a rotation of an angle $\Omega_1 - \Omega_2$

respect to the axis $\mathbf{k}_O$. Since the angles between $\mathbf{n}_1$ and $\mathbf{u}_1$ and between $\mathbf{n}_2$ and $\mathbf{u}_2$ are respectively $\omega_1 + \nu_1$ and $\omega_2 + \nu_2$, the unit vectors of P_1 and P_2 with respect to two bases are

$$\mathbf{u}_1 = \begin{bmatrix} \cos(\omega_1+\nu_1) \\ \sin(\omega_1+\nu_1)\cos i_1 \\ \sin(\omega_1+\nu_1)\sin i_1 \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} \cos(\omega_2+\nu_2) \\ \sin(\omega_2+\nu_2)\cos i_2 \\ \sin(\omega_2+\nu_2)\sin i_2 \end{bmatrix}. \tag{5.54}$$

Taking account of the rotation between the two bases and the fact that in the reduced problem the two longitudes of nodes Ω_1 and Ω_2 satisfy the relation $\Omega_2 = \Omega_1 + \pi$, we obtain[4]

$$\cos S = \cos w_1^\star \cos w_2^\star + \sin w_1^\star \sin w_2^\star \cos J \tag{5.56}$$

where $J = i_2 - i_1$ is the mutual inclination between the two planets and where

$$w_k^\star = \nu_k + \omega_k \tag{5.57}$$

is the *true longitude*. Note that $w_k^\star = \nu_k + \lambda_k^\star - M_k$.

Using the fact that the total angular momentum C is conserved, the calculation of $\cos J$ is easy. In fact, using (5.13), we derive immediately

$$C^2 = G_1^2 + G_2^2 + 2G_1G_2\cos J \tag{5.58}$$

and

$$\begin{aligned} \cos J &= \frac{C^2 - \Lambda_1^2(1-e_1^2) - \Lambda_2^2(1-e_2^2)}{2\Lambda_1\Lambda_2\sqrt{1-e_1^2}\sqrt{1-e_2^2}} = \\ &= \frac{4C^2 - (2\Lambda_1 - \chi_1\bar{\chi}_1)^2 - (2\Lambda_2 - \chi_2\bar{\chi}_2)^2}{2(2\Lambda_1 - \chi_1\bar{\chi}_1)(2\Lambda_2 - \chi_2\bar{\chi}_2)}, \end{aligned} \tag{5.59}$$

where we have used (5.24).

Replacing the expression (5.59) in (5.56), we obtain the desired development of $\cos S$.

5.3.3 Reduction to the plane problem

Suppose $J = 0$. Then the angle S is simply $S = w_1^\star - w_2^\star$, i.e. the difference between the true longitudes.

If we define $D^2 = 1 + \rho^2 - 2\rho\cos\beta$ where $\beta = w_1^\star - w_2^\star$, we have

$$\frac{1}{\Delta} = \frac{1}{r_2}(1+\rho^2-2\rho\cos\beta)^{-1/2} = \frac{1}{a_2}\frac{a_2}{r_2}D^{-1}. \tag{5.60}$$

If ρ is fixed, the function D is periodic with period 2π with respect to the variable β. Then it admits a development in Fourier series of the form

$$\frac{1}{D} = \frac{1}{2}b_{1/2}^{(0)}(\rho) + \sum_{j=1}^{\infty} b_{1/2}^{(j)}(\rho)\cos(j\beta), \tag{5.61}$$

[4] In the general case, $\cos S$ is equal to:

$$\begin{aligned} \cos S = \mathbf{u}_1\cdot\mathbf{u}_2 = {}& \cos(\Omega_1-\Omega_2)[\cos(\omega_1+\nu_1)\cos(\omega_2+\nu_2) + \sin(\omega_1+\nu_1)\sin(\omega_2+\nu_2)\cos i_1\cos i_2] + \\ &+ \sin(\Omega_1-\Omega_2)[\cos(\omega_1+\nu_1)\sin(\omega_2+\nu_2)\cos i_2 - \sin(\omega_1+\nu_1)\cos(\omega_2+\nu_2)\cos i_1] + \\ &+ \sin(\omega_1+\nu_1)\sin(\omega_2+\nu_2)\sin i_1\sin i_2. \end{aligned} \tag{5.55}$$

and, for a positive integer s, we have

$$D^{-s} = (1+\rho^2-2\rho\cos\beta)^{-s/2} = \sum_{j=-\infty}^{+\infty} \frac{1}{2} b_{s/2}^{(|j|)}(\rho)\exp(\sqrt{-1}j\beta). \tag{5.62}$$

The coefficients $\frac{1}{2}b_{s/2}^{(j)}$ are called *"Laplace coefficients"* and they are defined for $j \geq 0$ as

$$\frac{1}{2}b_{s/2}^{(j)} = \frac{(s/2)_j}{(1)_j}\rho^j F\left(\frac{s}{2},\frac{s}{2}+j,j+1;\rho^2\right), \tag{5.63}$$

where F is the *hypergeometric function of Gauss* defined as

$$F(a,b,c;x) = \sum_{k=0}^{\infty} \frac{(a)_k(b)_k}{(c)_k}\frac{x^k}{k!}, \tag{5.64}$$

which converges if c is not a negative integer for all $|x| < 1$ [5]. Here, $(a)_k$ is the Pochhammer symbol, which is defined by

$$(a)_0 = 1, \qquad (a)_k = a(a+1)...(a+k-1) = (a+k-1)(a)_{k-1}. \tag{5.66}$$

One way to obtain the relationship (5.63) is the following. We write

$$\begin{aligned}(1+\rho^2-2\rho\cos\beta)^{-1/2} &= (1-\rho\exp(\sqrt{-1}\beta)-\rho\exp(-\sqrt{-1}\beta)+\rho^2)^{-1/2} = \\ &= (1-\rho\exp(\sqrt{-1}\beta))^{-1/2}(1-\rho\exp(-\sqrt{-1}\beta))^{-1/2};\end{aligned} \tag{5.67}$$

then we develop the two terms with the formula of the binomial[6] and, after we have done the product of the series, we collect the terms with the common factor $\exp(\sqrt{-1}j\beta)$ and we obtain (5.63).

However, in general ρ is not constant. On the other hand if the eccentricities are low, ρ remains in a neighborhood of a fixed value. In particular, writing

$$\rho^2 = \alpha^2 + \alpha^2\left(\frac{\rho^2}{\alpha^2}-1\right) = \alpha^2+\epsilon \tag{5.69}$$

and using the Taylor expansion of the hypergeometric function

$$F(a,b,c;\alpha^2+\epsilon) = \sum_{m=0}^{\infty}\frac{(a)_m(b)_m}{(c)_m}\frac{\epsilon^m}{m!}F(a+m,b+m,c+m;\alpha^2), \tag{5.70}$$

[5] It can be useful the following property of the hypergeometric function (see *Whittaker* for more details):

$$F(a,b,c;x) = (1-x)^{c-b-a}F(c-a,c-b,c;x). \tag{5.65}$$

[6] The (generalized) binomial expansion of $(1+x)^\alpha$ is

$$(1+x)^\alpha = 1+\sum_{k=1}^{\infty}\frac{\alpha(\alpha-1)...(\alpha-k+1)}{k!}x^k, \qquad \alpha\in\mathbb{R}. \tag{5.68}$$

we obtain for $j \geq 0$

$$\frac{1}{2}b_{s/2}^{(j)} = \sum_{m=0}^{\infty} \varphi_{s,m}^{(j)}(\alpha)\left(\frac{\rho^2}{\alpha^2}-1\right)^m . \tag{5.71}$$

In this case, $\varphi_{s,m}^{(j)}(\alpha)$ depends only on α:

$$\varphi_{s,m}^{(j)}(\alpha) = \frac{(s/2)_j}{(1)_j}\frac{(s/2)_m(s/2+j)_m}{(j+1)_m}\frac{\alpha^{2m+j}}{(1)_m}F\left(\frac{s}{2}+m, \frac{s}{2}+m+j, j+m+1; \alpha^2\right). \tag{5.72}$$

Thus we obtain

$$D^{-s} = \sum_{j=-\infty}^{+\infty}\sum_{m=0}^{\infty} \varphi_{s,m}^{(|j|)}(\alpha)\left(\frac{\rho^2}{\alpha^2}-1\right)^m \left(\frac{\rho}{\alpha}\right)^{|j|}\exp(\sqrt{-1}j\beta), \tag{5.73}$$

which is equal to

$$D^{-s} = \sum_{j=0}^{+\infty}\sum_{m=0}^{\infty} 2\varphi_{s,m}^{(j)}(\alpha)\left(\frac{\rho^2}{\alpha^2}-1\right)^m \left(\frac{\rho}{\alpha}\right)^{j}\cos(j\beta). \tag{5.74}$$

Using (5.43) and the fact that $\exp(\sqrt{-1}j(w_1^\star - w_2^\star)) = \theta^j\bar{\theta}'^j \exp(\sqrt{-1}j(\lambda_1^\star - \lambda_2^\star))$, we obtain

$$D^{-s} = \sum_{j=-\infty}^{+\infty}\sum_{M\in\mathbb{N}_0^4} \phi_{M,j}^{(s)}(\alpha)X_1^{\mu}\bar{X}_1^{\bar{\mu}}X_2^{\mu'}\bar{X}_2^{\bar{\mu}'}\exp(\sqrt{-1}j(\lambda_1^\star - \lambda_2^\star)), \tag{5.75}$$

where M represents the 4-tuple $\{\mu, \bar{\mu}, \mu', \bar{\mu}'\}$. If we want to truncate the development to the degree d in eccentricities, it is sufficient to impose $M \in N^4(d)$, where $N^4(d)$ is the set of positive or null integers defined as

$$N^4(d) = \{(\mu, \bar{\mu}, \mu', \bar{\mu}') \in \mathbb{N}_0^4 \mid \quad 0 \leq \mu + \bar{\mu} + \mu' + \bar{\mu}' \leq d\}. \tag{5.76}$$

Finally, using the relationship (5.45) between X and z and the relationship (5.47) between z and the planar complex Poincaré variables, we obtain the development of D^{-s} in the planar Poincaré variables.

5.3.4 General case

Suppose now that at least one of the two inclinations is different from zero.

Starting from (5.50), we have

$$\begin{aligned}\frac{1}{\Delta} &= \frac{1}{r_2}(1+\rho^2-2\rho\cos\beta-2\rho(\cos S-\cos\beta))^{-1/2} = \\ &= \frac{1}{a_2}\frac{a_2}{r_2}D^{-1}(1-\rho U D^{-2})^{-1/2},\end{aligned} \tag{5.77}$$

where

$$U = 2(\cos S - \cos\beta) = 2\sin w_1^\star \sin w_2^\star(\cos J - 1). \tag{5.78}$$

As before, using the fact that the total angular momentum C is conserved, we derive immediately

$$\cos J - 1 = \frac{C^2 - \left(\Lambda_1\sqrt{1-e_1^2} + \Lambda_2\sqrt{1-e_2^2}\right)^2}{2\Lambda_1\Lambda_2\sqrt{1-e_1^2}\sqrt{1-e_2^2}} = \\ = \frac{4C^2 - (2\Lambda_1 - \chi_1\bar{\chi}_1 + 2\Lambda_2 - \chi_2\bar{\chi}_2)^2}{2(2\Lambda_1 - \chi_1\bar{\chi}_1)(2\Lambda_2 - \chi_2\bar{\chi}_2)}. \tag{5.79}$$

Replacing the expression (5.79) in (5.78), we obtain the development of U in the planar Poincaré variables.

Then, using the formula of the binomial, we have that the inverse of the distance in the reduced problem is given by

$$\frac{1}{\| \mathbf{r}_1 - \mathbf{r}_2 \|} = \frac{1}{a_2}\frac{a_2}{r_2}\sum_{k=0}^{\infty}(-1)^k\frac{(-1/2)_k}{(1)_k}(\rho U)^k D^{-2k-1}. \tag{5.80}$$

If we want to truncate the development to the degree d in eccentricities, it is sufficient in (5.80) to vary k between 0 and $\lfloor d/2 \rfloor$ and to truncate the development of D^{-2k-1} to the degree $(d-2k)$ in the eccentricities.

The transformation to planar Poincaré variables is then straightforward.

5.3.5 Kinetic part

Finally it remains to develop the kinetic part

$$T_1 = \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{m_0} \tag{5.81}$$

in the reduced problem, where $\mathbf{p}_k = \mu_k \dot{\mathbf{r}}_k$ and $\dot{\mathbf{r}}_k = [\dot{r}_{x,k}, \dot{r}_{y,k}, \dot{r}_{z,k}]$.

We assume that the motion of the two bodies is described by the osculating orbits, i.e. we consider a two-body fictitious problem. If we place ourselves in the plane of the orbit $(\mathcal{X}_k, \mathcal{Y}_k)$, the movement, expressed in function of the true anomaly ν_k, is given by

$$\begin{aligned} \mathcal{X}_k &= r_k \cos\nu_k, & \mathcal{Y}_k &= r_k \sin\nu_k, \\ \dot{\mathcal{X}}_k &= -\frac{n_k a_k}{\sqrt{1-e^2}}\sin\nu_k, & \dot{\mathcal{Y}}_k &= \frac{n_k a_k}{\sqrt{1-e_k^2}}(e_k + \cos\nu_k), \end{aligned} \tag{5.82}$$

where n_k is the mean motion of the corresponding Keplerian orbit and where $r_k = \| \mathbf{r}_k \|$. In the fixed reference frame, velocities are thus given by

$$\begin{bmatrix} \dot{r}_{x,k} \\ \dot{r}_{y,k} \\ \dot{r}_{z,k} \end{bmatrix} = \mathcal{R}_1(i_k) \times \mathcal{R}_3(\omega_k) \times \begin{bmatrix} \dot{\mathcal{X}}_k \\ \dot{\mathcal{Y}}_k \\ 0 \end{bmatrix}, \tag{5.83}$$

where

$$\mathcal{R}_1(i_k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos i_k & -\sin i_k \\ 0 & \sin i_k & \cos i_k \end{bmatrix}; \qquad \mathcal{R}_3(\omega_k) \begin{bmatrix} \cos\omega_k & -\sin\omega_k & 0 \\ \sin\omega_k & \cos\omega_k & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.84}$$

Thus we have

$$\frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{\mu_1\mu_2} = \mathcal{R}_1(i_1) \times \mathcal{R}_3(\omega_1) \times \begin{bmatrix} \dot{\mathcal{X}}_1 \\ \dot{\mathcal{Y}}_1 \\ 0 \end{bmatrix} \cdot \mathcal{R}_1(i_2) \times \mathcal{R}_3(\omega_2) \times \begin{bmatrix} \dot{\mathcal{X}}_2 \\ \dot{\mathcal{Y}}_2 \\ 0 \end{bmatrix}. \tag{5.85}$$

If we call

$$\begin{bmatrix} \hat{\mathcal{X}}_k \\ \hat{\mathcal{Y}}_k \\ 0 \end{bmatrix} = \mathcal{R}_3(\omega_k) \times \begin{bmatrix} \dot{\mathcal{X}}_k \\ \dot{\mathcal{Y}}_k \\ 0 \end{bmatrix}, \tag{5.86}$$

it is simple to prove that

$$T_1 \frac{m_0}{\mu_1\mu_2} = \hat{\mathcal{X}}_1\hat{\mathcal{X}}_2 + \hat{\mathcal{Y}}_1\hat{\mathcal{Y}}_2 \cos J. \tag{5.87}$$

Then we put

$$\mathcal{Z}_k = \hat{\mathcal{X}}_k + \sqrt{-1}\hat{\mathcal{Y}}_k = \frac{\sqrt{-1}n_k a_k}{\sqrt{1-e_k^2}} \left[\exp(\sqrt{-1}w_k^\star) + z_k)\right] \tag{5.88}$$

where z_k is defined in (5.45) ($z_k = e_k \exp\left(\sqrt{-1}\varpi_k^\star\right)$), and we obtain

$$T_1 = -\frac{\mu_1\mu_2}{2m_0}\left\{\mathcal{Z}_1\bar{\mathcal{Z}}_2 + \bar{\mathcal{Z}}_1\mathcal{Z}_2 - (\mathcal{Z}_1 - \bar{\mathcal{Z}}_1)(\mathcal{Z}_2 - \bar{\mathcal{Z}}_2)\left(\frac{\cos J - 1}{2}\right)\right\}. \tag{5.89}$$

The transformations to planar Poincaré variables is then straightforward.

5.3.6 Final form of the classical Hamiltonian

After some algebraic computations, the classical Hamiltonian in the planar Poincaré variables is given by

$$H(\mathbf{\Lambda}, \boldsymbol{\lambda}^\star, \boldsymbol{\eta}^\star, \boldsymbol{\xi}^\star; C) = H_0(\Lambda_1, \Lambda_2) + \varepsilon H_1(\Lambda_1, \Lambda_2, \lambda_1^\star, \lambda_2^\star, \eta_1^\star, \eta_2^\star, \xi_1^\star, \xi_2^\star; C) \tag{5.90}$$

where C is the magnitude of the total angular momentum, H_0 is given in (5.3) and where H_1 can be expanded in Fourier series of the angles $\boldsymbol{\lambda}^\star$ and in power series of $\boldsymbol{\eta}^\star$ and $\boldsymbol{\xi}^\star$, as in (5.8).

5.4 Development of the relativistic Hamiltonian in the reduced problem

In the relativistic case, in order to transform the simplified relativistic Hamiltonian H_{Rel} to the required form (5.3), we simply have to express the relativistic part H_2 of the Hamiltonian with respect to the planar Poincaré coordinates.

Because we are concerned with the secular variations of orbital elements, we can further simplify the problem using the averaging principle (see section 4.4) and replacing the relativistic perturbation H_2 with $\langle H_2 \rangle$, which is the average of H_2 with respect to the fast angles (the mean longitudes or the mean anomalies) over the periods. In this way, the averaging makes it possible to reduce the number of the degrees of freedom, because $\langle H_2 \rangle$ does not depend on the fast angles and so the evolution of the conjugate momenta to the fast angles become independent by the relativistic part, i.e. the semi-major axis a_i are constant. At the same time, assuming that no strong mean motion resonances are present and the system is far enough from collisions, we can still obtain qualitative information on the long-term changes of the slowly varying orbital elements (i.e., on the slow angles and their conjugate momenta).

Thus, in the absence of strong mean motion resonances, the fast angles M_i $(= l_i)$ can be eliminated in the Hamiltonian H_2 by the following averaging formula:

$$\langle H_2 \rangle = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} H_2 \mathrm{d}M_1 \mathrm{d}M_2 \tag{5.91}$$

and, using the averaging principle, the relativistic Hamiltonian (5.3) becomes

$$H_{\mathrm{Rel}} = H_0 + \varepsilon H_1 + \frac{1}{c^2} \langle H_2 \rangle. \tag{5.92}$$

Then we should average out each component of the Hamiltonian H_2 over the mean anomalies. Using the fact that $\mathbf{P}_i = \mu_i \mathbf{v}_i$, the mean relativistic Hamiltonian H_2 can be written as[7]:

$$\langle H_2 \rangle = \sum_{i=1}^{2} \mu_i \left(-\gamma_{1,i} \langle v_i^4 \rangle - \gamma_{2,i} \left\langle \frac{v_i^2}{r_i} \right\rangle - \gamma_{3,i} \left\langle \frac{(\mathbf{r}_i \cdot \mathbf{v}_i)^2}{r_i^3} \right\rangle + \gamma_{4,i} \left\langle \frac{1}{r_i^2} \right\rangle \right). \tag{5.93}$$

where $r_i = \parallel \mathbf{r}_i \parallel$ and $v_i = \parallel \mathbf{v}_i \parallel = \parallel \dot{\mathbf{r}}_i \parallel$, with the accuracy of $O(c^{-2})$.

To average out the whole relativistic Hamiltonian, we can calculate the integrals using a convenient change of variables, written in the general form of:

$$\langle \mathcal{X} \rangle = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \mathcal{X} \mathrm{d}M_1 \mathrm{d}M_2 = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} (\mathcal{X} \mathcal{J}) \mathrm{d}\nu_1 \mathrm{d}\nu_2 \tag{5.94}$$

where $\mathcal{J}$ is a scaling function defined by

$$\mathcal{J} = \mathcal{J}_1 \mathcal{J}_2 = \frac{\mathrm{d}M_1}{\mathrm{d}\nu_1} \frac{\mathrm{d}M_2}{\mathrm{d}\nu_2} = \frac{(1-e_1)^{3/2}}{(1+e_1 \cos \nu_1)^2} \frac{(1-e_2)^{3/2}}{(1+e_2 \cos \nu_2)^2} \tag{5.95}$$

where ν_i is the true anomaly of the $i-$th planet.

For the calculation of the integrals, the following relationships are useful

$$\begin{aligned} r_i &= \frac{a_i(1-e_i^2)}{1+e_i \cos \nu_i}, \\ v_i &= \frac{n_i a_i}{(1-e_i)^{1/2}} \sqrt{1+e_i^2+2e_i \cos \nu_i}, \\ \mathbf{r}_i \cdot \mathbf{v}_i &= n_i a_i^2 \sqrt{1-e_i^2} \frac{e_i \sin \nu_i}{1+e \cos \nu_i} \end{aligned} \tag{5.96}$$

where $n_i = \sqrt{\mathcal{G}(m_0+m_i)} a_i^{-3/2} = \beta_i^{1/2} a_i^{-3/2}$. The components of the mean relativistic Hamiltonian

[7]We remember that the quantities β_i, μ_i, $\gamma_{1,i}$, $\gamma_{2,i}$, $\gamma_{3,i}$ and $\gamma_{4,i}$ are defined in (3.65).

can be written explicitly as follows[8]:

$$\begin{aligned}
\langle v_i^4 \rangle &= \frac{n_i^4 a_i^4}{\sqrt{1-e_i^2}}\left[4 - 3\sqrt{1-e_i^2}\right], \\
\left\langle \frac{v_i^2}{r_i} \right\rangle &= n_i^2 a_i, \\
\left\langle \frac{(\mathbf{r}_i \cdot \mathbf{v}_i)^2}{r_i^3} \right\rangle &= \frac{n_i^2 a_i}{\sqrt{1-e_i^2}}\left[1 - \sqrt{1-e_i^2}\right], \\
\left\langle \frac{1}{r_i^2} \right\rangle &= \frac{1}{a_i^2\sqrt{1-e_i^2}}.
\end{aligned} \tag{5.97}$$

Finally, the mean relativistic Hamiltonian is:

$$\frac{1}{c^2}\langle H_2 \rangle = \sum_{i=1}^{2} \mu_i \left(-\frac{3\beta_i^2}{c^2 a_i^2 \sqrt{1-e_i^2}} + \frac{\beta_i^2(15-\upsilon_i)}{8a_i^2 c^2} \right). \tag{5.98}$$

Because the canonical Delaunay elements are defined through equation (5.10), we obtained that the mean relativistic Hamiltonian is

$$\frac{1}{c^2}\langle H_2 \rangle = \sum_{i=1}^{2} \mu_i^5 \beta_i^4 \left(-\frac{3}{c^2 L_i^3 G_i} + \frac{15-\upsilon_i}{8c^2 L_i^4} \right). \tag{5.99}$$

We recall that in the secular relativistic Hamiltonian $\langle H_2 \rangle$, only the star-planets interactions are considered.

To write the mean relativistic Hamiltonian in the planar Poincaré variables (5.20), we use the relationship between the eccentricity e_i and the variables $\Lambda_i, \eta_i^\star$ and $\xi_i^\star$, given by

$$e_i = \sqrt{1 - \left(1 - \frac{(\eta_i^\star)^2 + (\xi_i^\star)^2}{2\Lambda_i}\right)^2}. \tag{5.100}$$

Thus we obtain

$$\begin{aligned}
\langle H_2(\boldsymbol{\Lambda}, \boldsymbol{\eta}^\star, \boldsymbol{\xi}^\star) \rangle &= \sum_{i=1}^{2} \frac{\mu_i^5 \beta_i^4}{8\Lambda_i^4}\left[15 - \upsilon_i - 24\left(1 - \frac{(\eta_i^\star)^2 + (\xi_i^\star)^2}{2\Lambda_i}\right)^{-1}\right] = \\
&= \sum_{i=1}^{2} \frac{\mu_i^5 \beta_i^4}{8\Lambda_i^4}\left[15 - \upsilon_i - 24\sum_{j=0}^{\infty} \frac{1}{2^j \Lambda_i^j}\left((\eta_i^\star)^2 + (\xi_i^\star)^2\right)^j\right] = \\
&= \sum_{i=1}^{2} \frac{\mu_i^5 \beta_i^4}{8\Lambda_i^4}\left[15 - \upsilon_i - 24\sum_{j=0}^{\infty} \frac{1}{2^j \Lambda_i^j}\left(\sum_{s=0}^{j} \binom{j}{s} (\eta_i^\star)^{2s} (\xi_i^\star)^{2(j-s)}\right)\right],
\end{aligned} \tag{5.101}$$

[8] It may be useful the following formulæ:

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} \frac{\sin^4 f}{(1+e\cos f)^2}\mathrm{d}f &= \frac{3(2-e^2-2\sqrt{1-e^2})}{2e^4}, \\
\frac{1}{2\pi}\int_0^{2\pi} \frac{\sin^2 f}{1+e\cos f}\mathrm{d}f &= \frac{1-\sqrt{1-e^2}}{e^4}.
\end{aligned}$$

.

where we have expanded in power series of $\boldsymbol{\eta}^\star, \boldsymbol{\xi}^\star$ around the origin.

If we want to truncate the development to the degree d in the eccentricities, we find:

$$\langle H_2(\boldsymbol{\Lambda}, \boldsymbol{\eta}^\star, \boldsymbol{\xi}^\star)\rangle = \sum_{i=1}^{2} \frac{\mu_i^5 \beta_i^4}{8\Lambda_i^4} \left[15 - \upsilon_i - 24 \sum_{j=0}^{\lfloor d/2 \rfloor} \frac{1}{2^j \Lambda_i^j} \left(\sum_{s=0}^{j} \binom{j}{s} (\eta_i^\star)^{2s} (\xi_i^\star)^{2(j-s)} \right) \right]. \tag{5.102}$$

5.4.1 Relativistic precession of pericenter

With our simplifications, if we don't consider the Newtonian interaction between the two planets, we obtain that the secular Hamiltonian related to the relativistic correction is

$$\langle H(\mathbf{L}, \mathbf{G})\rangle = -\sum_{i=1}^{2} \frac{\beta_i^2 \mu_i^3}{2L_i^2} + \frac{1}{c^2} \left(-\frac{3\mu_i^5 \beta_i^4}{L_i^3 G_i} + \frac{\mu_i^5 \beta_i^4 (15 - \upsilon_i)}{8L_i^4} \right). \tag{5.103}$$

In this Hamiltonian all slow angle are cyclic; hence $\mathbf{L}$ and $\mathbf{G}$ would be integrals of motion in the absence of mutual planetary interactions. However, they are no longer constant when we consider the Newtonian contributions.

The equations of motion of the mean orbital coordinates thus become

$$\begin{aligned} \frac{\mathrm{d}l_k}{\mathrm{d}t} &= \frac{\beta_k^2 \mu_k^3}{L_k^3} + \frac{1}{c^2} \left(\frac{9\mu_i^5 \beta_i^4}{L_i^4 G_i} - \frac{\mu_i^5 \beta_i^4 (15 - \upsilon_i)}{2L_i^5} \right). \\ \frac{\mathrm{d}g_k}{\mathrm{d}t} &= \frac{1}{c^2} \frac{3\mu_i^5 \beta_i^4}{L_i^3 G_i^2}. \\ \frac{\mathrm{d}g_k}{\mathrm{d}t} &= 0, \end{aligned} \tag{5.104}$$

for $k = 1, 2$, whereas all mean momenta are constants of motions, i.e. $L_k(t) = L_k(0), G_k(t) = G_k(0)$ and $\mathcal{H}_k(t) = \mathcal{H}_k(0)$.

We can recognize in equations (5.104) the classical relativistic effect of the pericenter precession. Indeed, in Keplerian orbital elements, the equation (5.104) becomes:

$$\omega_{g_k} = \frac{\mathrm{d}g_k}{\mathrm{d}t} = \frac{1}{c^2} \frac{3\mathcal{G}^{3/2}(m_0 + m_k)^{3/2}}{a_k^{5/2}(1 - e_k^2)}, \tag{5.105}$$

which, over an averaged orbit of the k-th body of period

$$T_k = 2\pi \sqrt{\frac{a_k^3}{\mathcal{G}(m_0 + m_k)}}, \tag{5.106}$$

amount to

$$\Delta\omega_k = \frac{6\pi\mathcal{G}(m_0 + m_k)}{c^2 a_k (1 - e_k^2)}. \tag{5.107}$$

which is the classical formula for the relativistic pericenter precession for the two-body problem (see (2.41)). We derive it here to keep the paper self-consistent.

5.4.2 Final form of the relativistic Hamiltonian

After some algebraic computations, the relativistic Hamiltonian in the planar Poincaré variables is given by

$$\begin{aligned} H_{\mathrm{rel}}(\boldsymbol{\Lambda}, \boldsymbol{\lambda}^{\star}, \boldsymbol{\eta}^{\star}, \boldsymbol{\xi}^{\star}; C) = & \, H_0(\Lambda_1, \Lambda_2) + \varepsilon H_1(\Lambda_1, \Lambda_2, \lambda_1^{\star}, \lambda_2^{\star}, \eta_1^{\star}, \eta_2^{\star}, \xi_1^{\star}, \xi_2^{\star}; C) + \\ & + \frac{1}{c^2} \langle H_2(\Lambda_1, \Lambda_2, \eta_1^{\star}, \eta_2^{\star}, \xi_1^{\star}, \xi_2^{\star}; C) \rangle \end{aligned} \tag{5.108}$$

where C is the magnitude of the total angular momentum, H_0 and H_1 are defined in (5.90) and $\langle H_2 \rangle$ is defined in (5.102).

Chapter 6

Secular evolution in action-angle coordinates

We have seen that the classical Hamiltonian in the planar Poincaré variables is given by[1]

$$H_{\rm New}(\boldsymbol{\Lambda},\boldsymbol{\lambda},\boldsymbol{\eta},\boldsymbol{\xi};C)=H_0(\Lambda_1,\Lambda_2)+\varepsilon H_1(\boldsymbol{\Lambda},\boldsymbol{\lambda},\boldsymbol{\eta},\boldsymbol{\xi};C) \tag{6.1}$$

and the relativistic Hamiltonian is given by

$$H_{\rm Rel}(\boldsymbol{\Lambda},\boldsymbol{\lambda},\boldsymbol{\eta},\boldsymbol{\xi};C)=H_0(\Lambda_1,\Lambda_2)+\varepsilon H_1(\boldsymbol{\Lambda},\boldsymbol{\lambda},\boldsymbol{\eta},\boldsymbol{\xi};C)+\frac{1}{c^2}\langle H_2(\boldsymbol{\Lambda},\boldsymbol{\eta},\boldsymbol{\xi};C)\rangle, \tag{6.2}$$

where C is the magnitude of the total angular momentum, H_0 is the part that describes the Keplerian motion

$$H_0(\Lambda_1,\Lambda_2)=-\sum_{j=1}^{2}\frac{\mathcal{G}^2(m_0+m_j)^2\mu_j^3}{2\Lambda_j^2}, \tag{6.3}$$

and H_1 and $\langle H_2\rangle$ are the perturbation parts, which can be expanded in Fourier series of the angles $\boldsymbol{\lambda}$ and in power series of $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$.

In this chapter, we want to study the equations of motion of Hamiltonians (6.1) and (6.2) using the tools provided by Hamiltonian's theory and by perturbation theory. In particular, the aim is to reconstruct the evolution of the eccentricities of the planets by using analytical techniques, extending the Laplace-Lagrange theory.

To do this, first we use the averaging principle to simplify the problem. We remember that all extrasolar systems which we study are non-resonant and so, using the averaging principle, we can still obtain qualitative information on the long-term changes of the slowly varying orbital elements. Then we study the secular Hamiltonians using the perturbation method based on the Lie series and on the Birkhoff's normal form, which were explained in detail in chapter 4.

Previous works of *Libert & Henrard (2005, 2006)* for coplanar systems show that this analytical model gives an accurate description of the behavior of planetary systems which are not close to a mean-motion resonance. Moreover, they have shown that an expansion up to order 12 in the eccentricities is usually required for reproducing the secular behavior of extrasolar planetary systems. This expansion has also been used by *Beaugé et al. (2006)* to successfully reproduce the motions

[1] For the following, we will omit all the superscripts which we used in chapter 5 to distinguish the spatial Poincaré variables from the planar ones.

of irregular satellites with eccentricities up to 0.7. *Veras & Armitage (2007)* have highlighted the limitations of lower order expansions; using only a fourth-order expansion in the eccentricities, they did not reproduce, even qualitatively, the secular dynamics of extrasolar planetary systems. All the previous results have been obtained considering a secular Hamiltonian at order one in the masses. Thus, following the works of *Libert and Henrand*, we decide to expand the two Hamiltonians up to order 12 in the eccentricities.

To validate our results, we will compare our analytical integration with the direct numerical integration.

Finally, in the case of a coplanar three-body system, we look for a criterion to determine a priori when the relativistic corrections are important, i.e. to determine a priori when the difference between the classical case and the relativistic case are not negligible. To do this, we analyze the quadratic part of the classical and of the relativistic secular Hamiltonian.

6.1 The invariance of the semi-major axes according to Lagrange and Poisson

In the case of extrasolar systems that we have considered, we have seen that the values of the semi-major axes oscillate around a mean value. In effect, this result can be generalized to all non-resonant planetary systems.

To do this, we look for a solution of the equations of motion of Hamiltonians (6.1) and (6.2) using the Lindstedt method, i.e. we look for a solution which is a development of powers of ε of the form

$$\begin{aligned}\Lambda_j &= \Lambda_{j,0}(t) + \varepsilon\Lambda_{j,1}(t) + \varepsilon^2\Lambda_{j,2}(t) + ...\\ \lambda_j &= \lambda_{j,0}(t) + \varepsilon\lambda_{j,1}(t) + \varepsilon^2\lambda_{j,2}(t) + ...\\ \xi_j &= \xi_{j,0}(t) + \varepsilon\xi_{j,1}(t) + \varepsilon^2\xi_{j,2}(t) + ...\\ \eta_j &= \eta_{j,0}(t) + \varepsilon\eta_{j,1}(t) + \varepsilon^2\eta_{j,2}(t) + ...\end{aligned} \tag{6.4}$$

for $j = 1, 2$. In particular, the solution of the unperturbed part is elementary because the quantities $\mathbf{\Lambda}_0$, $\boldsymbol{\eta}_0$, $\boldsymbol{\xi}_0$ are constants determined by the initial data, while the mean longitudes $\boldsymbol{\lambda}_0(t)$ evolve linearly in time

$$\lambda_{j,0}(t) = \phi_j + \nu_j t, \qquad \nu_j = \frac{\mathcal{G}^2(m_0 + m_j)^2\mu_j^3}{\Lambda_{j,0}^3}, \tag{6.5}$$

where ν_j are the *mean motion frequencies* (for simplicity we assume that the initial time is $t_0 = 0$).

We assume for the following that the mean motion frequencies are non-resonant, i.e. that $\mathbf{k}\cdot\boldsymbol{\nu}$ is equal to zero if and only if $\mathbf{k} = \mathbf{0}$.

It can be proven that the coefficients $\mathbf{\Lambda}_r$, $\boldsymbol{\lambda}_r$, $\boldsymbol{\xi}_r$, $\boldsymbol{\eta}_r$ of ε^r are sum of terms of the form

$$A_{\mathbf{k}}(\mathbf{\Lambda}_0, \boldsymbol{\phi}_0, \boldsymbol{\eta}_0, \boldsymbol{\xi}_0)\, t^s \exp(i\,(\mathbf{k}\cdot\boldsymbol{\nu})\, t) \tag{6.6}$$

where $A_{\mathbf{k}}(\mathbf{\Lambda}_0, \boldsymbol{\phi}_0, \boldsymbol{\eta}_0, \boldsymbol{\xi}_0)$ are functions of the initial data. These terms can be classified as

- *periodic* if $s = 0$ and $\mathbf{k} \neq \mathbf{0}$;
- *pure secular terms* if $s \neq 0$ and $\mathbf{k} = \mathbf{0}$;

- *mixed secular terms* if $s \neq 0$ and $\mathbf{k} \neq \mathbf{0}$.

Moreover, it can be proven that

- as regard the pure secular terms, we have $s < r$ for the function $\boldsymbol{\Lambda}_r(t)$ and $s \leq r$ for the functions $\boldsymbol{\lambda}_r, \boldsymbol{\xi}_r, \boldsymbol{\eta}_r$;
- as regard the mixed secular terms, we have $s < r$.

It is simple to prove this result in the case of the semi-major axis at the first order. Indeed, in this case, we have

$$\dot{\Lambda}_{j,1}(t) = -\sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{Z}^2_+} \sum_{\mathbf{k}\in\mathbb{Z}^2\backslash\mathbf{0}} \sqrt{-1}\, k_j\, c_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{k}}(\boldsymbol{\Lambda}_0)\boldsymbol{\xi}_0^{\boldsymbol{\alpha}/2}\boldsymbol{\eta}_0^{\boldsymbol{\beta}/2}\exp[\sqrt{-1}(\mathbf{k}\cdot\boldsymbol{\lambda}_0)] \tag{6.7}$$

and so

$$\begin{aligned} \Lambda_{j,1}(t) &= -\sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{Z}^2_+} \sum_{\mathbf{k}\in\mathbb{Z}^2\backslash\mathbf{0}} \frac{k_j c_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{k}}(\boldsymbol{\Lambda}_0)\exp(\sqrt{-1}\mathbf{k}\cdot\boldsymbol{\phi}_0)}{\mathbf{k}\cdot\boldsymbol{\nu}}\boldsymbol{\xi}_0^{\boldsymbol{\alpha}/2}\boldsymbol{\eta}_0^{\boldsymbol{\beta}/2}\exp[\sqrt{-1}(\mathbf{k}\cdot\boldsymbol{\nu})t] = \\ &= \sum_{\mathbf{k}\in\mathbb{Z}^2\backslash\mathbf{0}} C_{\mathbf{k}}(\boldsymbol{\Lambda}_0,\boldsymbol{\xi}_0,\boldsymbol{\eta}_0,\boldsymbol{\nu},\boldsymbol{\phi})\,{}^{\sin}_{\cos}((\mathbf{k}\cdot\boldsymbol{\nu})t). \end{aligned} \tag{6.8}$$

Then, at the first development of ε, the perturbation does not introduce systematic variations of the semi-major axes, but only quasi-periodic variations, i.e. the time always appears as the argument of a trigonometric function. It can be proven that the series (6.8) converges uniformly in ε if the frequencies $\boldsymbol{\nu}$ are strongly non-resonant, i.e. if they satisfy for example the Diophantine condition (4.87).

The formal statement of the theorem is the following:

Theorem 3 (Lagrange)**.** *If there are no resonances between the mean motion frequencies of the planets, then the semi-major axes are not subject to secular variations in the approximation of order 1 of the masses.*

The Lagrange's theorem is generalized by Poisson's theorem:

Theorem 4 (Poisson)**.** *If there are no resonances between the mean motion frequencies of the planets, then the functions that represent the movement of the semi-major axes do not contain pure secular terms at least up to the second order in the masses.*

Then, at the second development of ε, the values of the semi-major axes oscillate around a mean value, but the amplitude of these oscillations may depend linearly on the time. In particular, it is important to note that the average values of the semi-major axes remain constant at least up to the second order in the masses.

Thus, based on these theorems and since we consider only non-resonant extrasolar systems, we decide to simplify the problem by averaging the two Hamiltonians with respect to the fast angles M_i $(= l_i)$, for $i = 1, 2$. This approximation corresponds to fixing the value of $\boldsymbol{\Lambda}$, that remains constant under the flow, and thereby the semi-major axes. The averaged Hamiltonian, depending only on the secular variables, reduces the problem to a system with 2 degrees of freedom. Thus, assuming that no strong mean motion resonances are present and that the system is far enough from collisions, the averaging makes it possible to reduce the number of the degrees of freedom, and to obtain qualitative information on the long-term changes of the eccentricities.

6.2 Secular Hamiltonians

6.2.1 Expansion of the Hamiltonians

In order to construct the first basic approximation of the normal form, we first expand the Hamiltonians (6.1) and (6.2) in Taylor-Fourier series.

To do this, we pick a fixed value $\mathbf{\Lambda}^*$ for the actions $\mathbf{\Lambda}$ and we perform a translation $\mathcal{T}_{\mathbf{\Lambda}^*}$ defined as

$$L_j = \Lambda_j - \Lambda_j^*, \qquad j = 1, 2. \tag{6.9}$$

Obviously this is a canonical transformation that leaves the coordinates $\boldsymbol{\lambda}, \boldsymbol{\eta}$ and $\boldsymbol{\xi}$ unchanged.

In particular, taking into account the result of Lagrange, we can choose $\mathbf{a}^*$ as the average values of the semi-major axes during the evolution and then we can determine $\mathbf{\Lambda}^*$ via the formula (5.1).

Then, we expand the transformed Hamiltonians $H_{\rm New}^{(\mathcal{T})} = H_{\rm New} \circ \mathcal{T}_{\mathbf{\Lambda}^*}$ and $H_{\rm Rel}^{(\mathcal{T})} = H_{\rm Rel} \circ \mathcal{T}_{\mathbf{\Lambda}^*}$ in power series of $\mathbf{L}, \boldsymbol{\eta}, \boldsymbol{\xi}$ around the origin.
For example, the expansion of the term H_0 is simple:

$$\begin{aligned} -\frac{\beta_j^2 \mu_j^3}{2\Lambda_j^2} &= -\frac{\beta_j^2 \mu_j^3}{2(\Lambda_j^* + L_j)^2} = -\frac{\beta_j^2 \mu_j^3}{2\left(\Lambda_j^*\right)^2} \left(1 + \frac{L_j}{\Lambda_j^*}\right)^{-2} = \\ &= -\frac{\mathcal{G} m_0 m_j}{(a_j^*)^2} \sum_{k=0}^{\infty} (-1)^k \frac{k+1}{\left(\Lambda_j^*\right)^k} (L_j)^k , \end{aligned} \tag{6.10}$$

where a_j^* is the corresponding value of Λ_j^*.

Thus, forgetting an unessential constant, we rearrange the Hamiltonians of the system as

$$H_{\rm New}^{(\mathcal{T})}(\mathbf{L}, \boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\xi}) = \boldsymbol{\nu}^* \cdot \mathbf{L} + \sum_{j_1=2}^{\infty} h_{j_1,0}^{(\rm Kep)}(\mathbf{L}) + \varepsilon \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} h_{j_1,j_2}^{(\rm New)}(\mathbf{L}, \boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\xi}) \tag{6.11}$$

and

$$H_{\rm Rel}^{(\mathcal{T})}(\mathbf{L}, \boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\xi}) = \boldsymbol{\nu}^* \cdot \mathbf{L} + \sum_{j_1=2}^{\infty} h_{j_1,0}^{(\rm Kep)}(\mathbf{L}) + \varepsilon \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} h_{j_1,j_2}^{(\rm Rel)}(\mathbf{L}, \boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\xi}), \tag{6.12}$$

where

- $\boldsymbol{\nu}^* = [\nu_1^*, \nu_2^*]$ are the mean motion frequencies, defined as

$$\nu_j^* = \frac{\partial H_0}{\partial \Lambda_j}\bigg|_{\Lambda_j = \Lambda_j^*} = \frac{\beta_j^2 \mu_j^3}{(\Lambda_j^*)^3}; \tag{6.13}$$

- the terms $h_{j_1,0}^{(\rm Kep)}$ of the Keplerian part are homogeneous polynomials of degree j_1 in the actions $\mathbf{L}$, the explicit expression of which can be determined in a straightforward manner;

- the functions $h_{j_1,j_2}^{(\rm New)}$ and $h_{j_1,j_2}^{(\rm Rel)}$ are homogeneous polynomials of degree j_1 in the actions $\mathbf{L}$ and of degree j_2 in the secular variables $(\boldsymbol{\eta}, \boldsymbol{\xi})$; the coefficients of such homogeneous polynomials do depend analytically and periodically on the angles $\boldsymbol{\lambda}$;

- if we expand $H_2^{(\mathcal{T})} = H_2 \circ \mathcal{T}_{\mathbf{\Lambda}^*}$ in power series of $\mathbf{L}, \boldsymbol{\eta}, \boldsymbol{\xi}$ around the origin:

$$\frac{1}{c^2} H_2^{(\mathcal{T})}(\mathbf{L}, \boldsymbol{\eta}, \boldsymbol{\xi}) = \frac{1}{c^2} \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} h_{j_1,j_2}^{(\mathrm{PNew})}(\mathbf{L}, \boldsymbol{\eta}, \boldsymbol{\xi}), \tag{6.14}$$

where the functions $h_{j_1,j_2}^{(\mathrm{PNew})}$ are homogeneous polynomials of degree j_1 in the actions $\mathbf{L}$ and of degree j_2 in the secular variables $(\boldsymbol{\eta}, \boldsymbol{\xi})$, the relationship between $h_{j_1,j_2}^{(\mathrm{New})}$ and $h_{j_1,j_2}^{(\mathrm{Rel})}$ is given by

$$h_{j_1,j_2}^{(\mathrm{Rel})} = h_{j_1,j_2}^{(\mathrm{New})} + \frac{1}{\varepsilon \cdot c^2} h_{j_1,j_2}^{(\mathrm{PNew})}. \tag{6.15}$$

The functions $h_{j_1,j_2}^{(\mathrm{New})}$ and $h_{j_1,j_2}^{(\mathrm{Rel})}$ are also expanded in Fourier series of the angles $\boldsymbol{\lambda}$. Let us also recall that the coefficients of the Fourier expansion decay exponentially with $|\mathbf{k}| = |k_1| + |k_2|$.

In these Hamiltonians, it is simple to see that $\boldsymbol{\lambda}$ play the role of the fast angles, because $\boldsymbol{\lambda} = O(1)$, while the semi-major axis, the argument of perihelion and the longitudes of node are slow angles, because $\mathbf{\Lambda}$, $\boldsymbol{\eta}$, $\boldsymbol{\xi} = O(\varepsilon)$.

All the expansions were carried out using a specially devised algebraic manipulator developed by Marco Sansottera (for details, see *Giorgilli & Sansottera (2011)*). In our computations we truncate the expansion as follows. The Keplerian part is expanded up to the quadratic terms. The terms $h_{j_1,j_2}^{(\mathrm{New})}$ and $h_{j_1,j_2}^{(\mathrm{Rel})}$ include the linear terms in the fast actions $\mathbf{L}$, all terms up to degree 12 in the secular variables $(\boldsymbol{\xi}, \boldsymbol{\eta})$ and all terms up to the trigonometric degree 12 with respect to the angles $\boldsymbol{\lambda}$. The choice of the limits in the expansion is uniform for all the systems that will be considered.

6.2.2 Averaging over the mean motions

Now we perform an average over the fast angles $\boldsymbol{\lambda}$ of the Hamiltonians $H_{\mathrm{New}}^{(\mathcal{T})}$ (6.11) and $H_{\mathrm{Rel}}^{(\mathcal{T})}$ (6.12). More precisely, we calculate the averaged Hamiltonians

$$H^{(\mathrm{sec})}(\mathbf{L}, \boldsymbol{\xi}, \boldsymbol{\eta}) = \langle H^{(\mathcal{T})}(\mathbf{L}, \boldsymbol{\lambda}, \boldsymbol{\xi}, \boldsymbol{\eta}) \rangle_{\boldsymbol{\lambda}}, \tag{6.16}$$

where $\langle \mathcal{X} \rangle$ is defined in (5.91), namely we average $H^{(\mathcal{T})}$ by removing all the Fourier harmonics depending on the angles.

It is important to notice that the average over the fast angles $\boldsymbol{\lambda}$ is equivalent to distribute the mass of each planet around its orbit and to replace the attraction of each planet by the attraction of the ring so obtained.

Thus, the averaged Hamiltonians are

$$H_{\mathrm{New}}(\mathbf{L}, \boldsymbol{\eta}, \boldsymbol{\xi}) = \boldsymbol{\nu}^* \cdot \mathbf{L} + \sum_{j_1=2}^{\infty} h_{j_1,0}^{(\mathrm{Kep})}(\mathbf{L}) + \varepsilon \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \langle h_{j_1,j_2}^{(\mathrm{New})}(\mathbf{L}, \boldsymbol{\eta}, \boldsymbol{\xi}) \rangle \tag{6.17}$$

and

$$H_{\mathrm{Rel}}(\mathbf{L}, \boldsymbol{\eta}, \boldsymbol{\xi}) = \boldsymbol{\nu}^* \cdot \mathbf{L} + \sum_{j_1=2}^{\infty} h_{j_1,0}^{(\mathrm{Kep})}(\mathbf{L}) + \varepsilon \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \langle h_{j_1,j_2}^{(\mathrm{Rel})}(\mathbf{L}, \boldsymbol{\eta}, \boldsymbol{\xi}) \rangle. \tag{6.18}$$

We recall again that if no strong mean motion resonances are present and if the system is far enough from collisions, the secular dynamics is an accurate description of the real dynamics.

By definition, the planetary secular normal form does not depend on the mean longitudes of the planets λ_1, λ_2. As a consequence, the actions L_1, L_2 are constants of motion and so we set

$$L_1 = 0, \qquad L_2 = 0. \tag{6.19}$$

The secular system is therefore completely described by the canonical action-angle variables $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ and the two Hamiltonians take the form

$$\begin{aligned} H_{\text{New}}(\boldsymbol{\eta}, \boldsymbol{\xi}; C, \boldsymbol{\Lambda}^*) &= \varepsilon \sum_{j_2=0}^{\infty} \langle h_{0,j_2}^{(\text{New})}(\boldsymbol{\eta}, \boldsymbol{\xi}; C, \boldsymbol{\Lambda}^*) \rangle, \\ \langle h_{0,j_2}^{(\text{New})}(\boldsymbol{\eta}, \boldsymbol{\xi}; C, \boldsymbol{\Lambda}^*) \rangle &= \sum_{\substack{\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{Z}_+^{2n} \\ \alpha_1+\alpha_2+\beta_1+\beta_2 = j_2}} c_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{0}}^{(\text{New})}(\boldsymbol{\Lambda}^*, C) \boldsymbol{\xi}^{\boldsymbol{\alpha}} \boldsymbol{\eta}^{\boldsymbol{\beta}}, \end{aligned} \tag{6.20}$$

and

$$\begin{aligned} H_{\text{Rel}}(\boldsymbol{\eta}, \boldsymbol{\xi}; C, \boldsymbol{\Lambda}^*) &= \varepsilon \sum_{j_2=0}^{\infty} \langle h_{0,j_2}^{(\text{Rel})}(\boldsymbol{\eta}, \boldsymbol{\xi}; C, \boldsymbol{\Lambda}^*) \rangle \\ \langle h_{0,j_2}^{(\text{Rel})}(\boldsymbol{\eta}, \boldsymbol{\xi}; C, \boldsymbol{\Lambda}^*) \rangle &= \sum_{\substack{\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{Z}_+^{2n} \\ \alpha_1+\alpha_2+\beta_1+\beta_2 = j_2}} c_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{0}}^{(\text{Rel})}(\boldsymbol{\Lambda}^*, C) \boldsymbol{\xi}^{\boldsymbol{\alpha}} \boldsymbol{\eta}^{\boldsymbol{\beta}}, \end{aligned} \tag{6.21}$$

where $\boldsymbol{\xi}^{\boldsymbol{\alpha}} \boldsymbol{\eta}^{\boldsymbol{\beta}} = \xi_1^{\alpha_1} \xi_2^{\alpha_2} \eta_1^{\beta_1} \eta_2^{\beta_2}$ and where the quantities C and $\boldsymbol{\Lambda}^*$ play the role of constant parameters.

The Hamiltonians so constructed are the secular ones, describing the slow motion of the eccentricities and pericenters.

6.3 Secular dynamics of the planets

In (6.20) and in (6.21), the small parameter ε (approximately the mass of the largest planet relative to that of the star) multiplies the entire Hamiltonians. Therefore, it no longer plays the role of a perturbation parameter separating an integrable part from its perturbation, but simply shows that the motion described by the Hamiltonians (6.20) and (6.21) is slow.

In order to study these Hamiltonians with the tools described in chapter 4, we first have to find an integrable approximation and a new perturbation parameter - say ϵ - such that, in suitable action-angle variables $\mathbf{p}, \mathbf{q}$, the Hamiltonians (6.20) and (6.21) can be written as $H_0(\mathbf{p}) + \epsilon H_\epsilon(\mathbf{p}, \mathbf{q}, \epsilon)$, with H_ϵ of order ϵ with respect to $\text{grad}_{\mathbf{p}}(H_0)$.

6.3.1 Averaged Hamiltonians in diagonal form

As we have seen, the expansion of the Hamiltonians contain only specific combinations of terms, in view of the D'Alembert rules. Thus, we use the D'Alembert rules to better characterize the developments (6.20) and (6.21).

Using the property 3 of the paragraph 5.1.1 (i.e., $k_1 + k_2$ and $\alpha_1 + \alpha_2 + \beta_1 + \beta_2$ must have the same parity), we have that $\sum_{j=1}^{2} (\alpha_j + \beta_j)$ must be even, because $\mathbf{k} = \mathbf{0}$. This implies that all the terms $\langle h_{0,j_2} \rangle$ with odd j_2 vanishes. Moreover, we can neglect the term $\langle h_{0,0} \rangle$ because it is constant. Then, it is convenient to introduce the following quantities

$$H_j^{(\text{New})}(\boldsymbol{\eta}, \boldsymbol{\xi}) = \varepsilon \langle h_{0,2j+2}^{(\text{New})}(\boldsymbol{\eta}, \boldsymbol{\xi}; C, \boldsymbol{\Lambda}^*) \rangle, \qquad H_j^{(\text{Rel})}(\boldsymbol{\eta}, \boldsymbol{\xi}) = \varepsilon \langle h_{0,2j+2}^{(\text{Rel})}(\boldsymbol{\eta}, \boldsymbol{\xi}; C, \boldsymbol{\Lambda}^*) \rangle, \tag{6.22}$$

so that the Hamiltonians become

$$\begin{aligned} H_{\text{New}}(\boldsymbol{\eta},\boldsymbol{\xi};C,\boldsymbol{\Lambda}^*) &= H_0^{(\text{New})}(\boldsymbol{\eta},\boldsymbol{\xi}) + \sum_{j=1}^{\infty} H_j^{(\text{New})}(\boldsymbol{\eta},\boldsymbol{\xi}), \\ H_{\text{Rel}}(\boldsymbol{\eta},\boldsymbol{\xi};C,\boldsymbol{\Lambda}^*) &= H_0^{(\text{Rel})}(\boldsymbol{\eta},\boldsymbol{\xi}) + \sum_{j=1}^{\infty} H_j^{(\text{Rel})}(\boldsymbol{\eta},\boldsymbol{\xi}). \end{aligned} \tag{6.23}$$

In particular, $H_j^{(\text{New})}$ and $H_j^{(\text{Rel})}$ are homogeneous polynomials of degree $2j+2$ in $(\boldsymbol{\eta},\boldsymbol{\xi})$, for each $j \in \mathbb{N}_0$.

Still because of D'Alembert rules, we want to show that $H_0^{(\text{New})}$ and $H_0^{(\text{Rel})}$ have the form

$$\begin{aligned} H_0^{(\text{New})}(\boldsymbol{\eta},\boldsymbol{\xi}) &= \frac{1}{2}\boldsymbol{\eta}\cdot\mathcal{A}_{(\text{New})}\boldsymbol{\eta} + \frac{1}{2}\boldsymbol{\xi}\cdot\mathcal{A}_{(\text{New})}\boldsymbol{\xi} \\ H_0^{(\text{Rel})}(\boldsymbol{\eta},\boldsymbol{\xi}) &= \frac{1}{2}\boldsymbol{\eta}\cdot\mathcal{A}_{(\text{Rel})}\boldsymbol{\eta} + \frac{1}{2}\boldsymbol{\xi}\cdot\mathcal{A}_{(\text{Rel})}\boldsymbol{\xi} \end{aligned} \tag{6.24}$$

where $\mathcal{A}_{(\text{New})}$ and $\mathcal{A}_{(\text{Rel})}$ are real symmetric 2×2 matrices.

For the following, we use the index 1 to indicate the innermost planet and the index 2 to indicate the outer planet.

Using formula (5.101), it is simple to prove that the relationship between $\mathcal{A}_{(\text{Rel})}$ and $\mathcal{A}_{(\text{New})}$ is given by:

$$\frac{1}{2}\mathcal{A}_{(\text{Rel})} = \frac{1}{2}\mathcal{A}_{(\text{New})} - \frac{1}{c^2}\cdot\begin{bmatrix} \frac{3\mathcal{G}^{3/2}(m_0+m_1)^{3/2}}{2(a_1^*)^{5/2}} & 0 \\ 0 & \frac{3\mathcal{G}^{3/2}(m_0+m_2)^{3/2}}{2(a_2^*)^{5/2}} \end{bmatrix}. \tag{6.25}$$

Moreover, there is a canonical transformation $(\boldsymbol{\eta},\boldsymbol{\xi}) \to (\mathbf{x},\mathbf{y})$ such that $H_0^{(\text{New})}$ and $H_0^{(\text{Rel})}$ can be rewritten as

$$\begin{aligned} H_0^{(\text{New})}(\mathbf{x},\mathbf{y}) &= \frac{1}{2}\sum_{i=1}^{2}\omega_i^{(\text{New})}(x_i^2+y_i^2), \\ H_0^{(\text{Rel})}(\mathbf{x},\mathbf{y}) &= \frac{1}{2}\sum_{i=1}^{2}\omega_i^{(\text{Rel})}(x_i^2+y_i^2). \end{aligned} \tag{6.26}$$

To prove (6.24) and (6.26), we do not distinguish the classical case from the relativistic one, and we omit for the moment the subscript and the superscript "(New)" and "(Rel)".

Using D'Alembert rules, the general form of H_0 in the planar modified Delaunay variables must be:

$$H_0(\mathbf{P},\mathbf{p}) = \sum_{\substack{\boldsymbol{\alpha}\in\mathbb{Z}_+^2 \\ |\boldsymbol{\alpha}|=2}} \sum_{\substack{\mathbf{m}\in\mathbb{Z}^2 \\ |\mathbf{m}|=0,2}} a_{\boldsymbol{\alpha},\mathbf{m}}(\mathbf{P})^{\boldsymbol{\alpha}/2}\cos(\mathbf{m}\cdot\mathbf{p}), \tag{6.27}$$

where $a_{\boldsymbol{\alpha},\mathbf{m}}$ are suitable coefficients. Taking into account the limitations of $\mathbf{m}$ and $\boldsymbol{\alpha}$, we can rewrite (6.27) in the following form

$$H_0(\mathbf{P},\mathbf{p}) = \sum_{1\le j\le k\le 2} b_{j,k}\sqrt{P_jP_k}\cos(p_j-p_k) + d_{j,k}\sqrt{P_jP_k}\cos(p_j+p_k) \tag{6.28}$$

where the coefficients $b_{j,k}$ and $d_{j,k}$ depending only on the constants $\boldsymbol{\Lambda}^*$ and on C, namely on average values of the semi-major axes of the planets and on the total angular momentum. Note that when

$k = j$ the terms in the sum become $b_{j,j}P_j$ and $d_{j,j}P_j\cos(2p_j)$. Now, using the property 2 of the paragraph 5.1.1 (i.e., $m_1 + m_2 = 0$), we have that all coefficients $d_{j,k}$ in (6.28) must be equal to zero, i.e. the Hamiltonian H_0 must be in the form

$$H_0(\mathbf{P}, \mathbf{p}) = \sum_{j=1}^{2} b_{j,j}P_j + b_{1,2}\sqrt{P_1P_2}\cos(p_1 - p_2). \tag{6.29}$$

Now, using the trigonometric formulæ, we can notice that

$$\begin{gathered} 2P_j = \eta_j^2 + \xi_j^2, \\ 2\sqrt{P_1P_2}\cos(p_1 - p_2) = \sqrt{2P_1}\cos p_1\sqrt{2P_2}\cos p_2 + \sqrt{2P_1}\sin p_1\sqrt{2P_2}\sin p_2 = \eta_1\eta_2 + \xi_1\xi_2. \end{gathered} \tag{6.30}$$

Therefore in H_0 can not appear terms with the product of one of the variables ξ with one of η; moreover, the coefficients of the monomials ξ_j^2 and η_j^2 must be the same, as well as the coefficients of the monomials $\xi_1\xi_2$ and $\eta_1\eta_2$ must be the same.

Thus the Hamiltonian H_0 in (6.23) can be rewritten in matrix form as

$$H_0(\boldsymbol{\eta}, \boldsymbol{\xi}) = \frac{1}{2}\boldsymbol{\eta}\cdot\mathcal{A}\boldsymbol{\eta} + \frac{1}{2}\boldsymbol{\xi}\cdot\mathcal{A}\boldsymbol{\xi} \tag{6.31}$$

where $\mathcal{A}$ is a real symmetric 2×2 matrix.

Because $\mathcal{A}$ is symmetric, it admits 2 real eigenvalues which we denote with ω_1 and ω_2. Moreover, $\mathcal{A}$ is diagonalizable by a rotation of the vectors, i.e. there exists an orthogonal matrix R such that

$$R^T\mathcal{A}R = \Omega \tag{6.32}$$

where $R^TR = I$ (I is the identity matrix) and

$$\Omega = \begin{bmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{bmatrix}. \tag{6.33}$$

Now using the canonical transformation (see Lemma 4.1)

$$\boldsymbol{\eta} = R\mathbf{x}, \qquad \boldsymbol{\xi} = R\mathbf{y}, \tag{6.34}$$

where we have used the property $R = (R^T)^{-1}$, the Hamiltonian (6.31) can be rewritten as

$$H_0(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\sum_{i=1}^{2}\omega_i(x_i^2 + y_i^2) \tag{6.35}$$

which is trivially integrable, because it is the Hamiltonian of a system of 2 harmonic oscillators.

We can rewrite the Hamiltonians (6.23) in the new canonical variables $(\mathbf{x}, \mathbf{y})$ defined in (6.34) as

$$\begin{aligned} H_{\mathrm{New}}(\mathbf{x}, \mathbf{y}; C, \boldsymbol{\Lambda}^*) &= H_0^{(\mathrm{New})}(\mathbf{x}, \mathbf{y}) + \sum_{j=1}^{\infty} H_j^{(\mathrm{New})}(\mathbf{x}, \mathbf{y}), \\ H_{\mathrm{Rel}}(\mathbf{x}, \mathbf{y}; C, \boldsymbol{\Lambda}^*) &= H_0^{(\mathrm{Rel})}(\mathbf{x}, \mathbf{y}) + \sum_{j=1}^{\infty} H_j^{(\mathrm{Rel})}(\mathbf{x}, \mathbf{y}) \end{aligned} \tag{6.36}$$

where

$$H_0^{(\mathrm{New})}(\mathbf{x},\mathbf{y}) = \frac{1}{2}\sum_{i=1}^{2}\omega_i^{(\mathrm{New})}(x_i^2+y_i^2), \qquad H_0^{(\mathrm{Rel})}(\mathbf{x},\mathbf{y}) = \frac{1}{2}\sum_{i=1}^{2}\omega_i^{(\mathrm{Rel})}(x_i^2+y_i^2). \tag{6.37}$$

The origin $(\boldsymbol{\xi},\boldsymbol{\eta}) = (\mathbf{0},\mathbf{0})$ is an elliptic equilibrium point[2], and, in the original variables $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, the solutions of the equations of motion given by the Hamiltonians $H_0^{(\mathrm{New})}$ and $H_0^{(\mathrm{Rel}}$ are

$$\xi_j(t) = \sum_{k=1}^{2} M_{j,k}\cos(\omega_k t+\phi_k), \qquad \eta_j(t) = \sum_{k=1}^{2} N_{j,k}\cos(\omega_k t+\psi_k) \tag{6.38}$$

for $j=1,2$, where $M_{j,k}, N_{j,k}, \phi_k$ and ψ_k depend at this level - i.e. considering only the H_0 part of the secular Hamiltonian - only on the semi-major axis of the planets, i.e. only on the initial data $\boldsymbol{\xi}(0)$ and $\boldsymbol{\eta}(0)$. This is usually known as the *Lagrange-Laplace solution* for the secular planetary motion.

6.3.2 Use of action-angle variables

In order to study the Hamiltonians (6.36), we have to introduce suitable action-angle variables $\boldsymbol{\Phi},\boldsymbol{\varphi}$ and a new perturbation parameter - say ϵ - such that the Hamiltonians (6.36) can be written in the form (4.40).

The lack of a perturbation parameter is just a trivial matter, because the perturbation parameter[3] ϵ is easily replaced by the distance from the origin. Indeed, if we consider the dynamics inside a polydisk $\Delta_{\rho R}$ with center of the origin of $\mathbb{R}^{2n}$ defined as

$$\Delta_{\rho R} = \{(\mathbf{x},\mathbf{y})\in\mathbb{R}^{2n}\,|\,x_j^2+y_j^2\le\rho^2R^2,\ j=1,...,n\} \tag{6.39}$$

where R is a positive number and $\rho>0$ being a parameter, then the homogeneous polynomial $H_s(\mathbf{x},\mathbf{y})$ is of order $O(\rho^{2s+2})$, so ρ plays the role of perturbation parameter.

To introduce the power expansion in ϵ, we do a scaling transformation

$$x_j = \sqrt{\epsilon}x'_j, \qquad y_j = \sqrt{\epsilon}y'_j \tag{6.40}$$

which is not canonical, but preserves the canonical form of the equations if the new Hamiltonian is defined as

$$H'(\mathbf{x}',\mathbf{y}';C,\boldsymbol{\Lambda}^*) = \frac{1}{\epsilon}H(\mathbf{x},\mathbf{y};C,\boldsymbol{\Lambda}^*)\Big|_{\mathbf{x}=\sqrt{\epsilon}\mathbf{x}',\,\mathbf{y}=\sqrt{\epsilon}\mathbf{y}'}. \tag{6.41}$$

The Hamiltonians (6.36) become

$$\begin{aligned} H'(\mathbf{x}',\mathbf{y}';C,\boldsymbol{\Lambda}^*) &= H_0'^{(\mathrm{New})}(\mathbf{x}',\mathbf{y}') + \sum_{j=1}^{\infty}\epsilon^j H_j'^{(\mathrm{New})}(\mathbf{x}',\mathbf{y}'),\\ H'_{\mathrm{rel}}(\mathbf{x}',\mathbf{y}';C,\boldsymbol{\Lambda}^*) &= H_0'^{(\mathrm{Rel})}(\mathbf{x}',\mathbf{y}') + \sum_{j=1}^{\infty}\epsilon^j H_j'^{(\mathrm{Rel})}(\mathbf{x}',\mathbf{y}'). \end{aligned} \tag{6.42}$$

[2]A point of equilibrium is of elliptic type when the eigenvalues of the system of differential equations linearized in the neighborhood of equilibrium are all pure imaginary.

[3]While in the original planetary problem the natural perturbation parameter ε is the largest planetary mass relative to that of the star, in the secular problem the natural perturbation parameter ϵ becomes the square of the largest value assumed by the planetary eccentricities during the secular evolution. Indeed, it is simple to prove that if the eccentricities are small, the size of terms $H_j^{(\mathrm{New})}$ and $H_j^{(\mathrm{Rel})}$ with $j\ge1$ (which play the role of a perturbation in (6.36)), relative respectively to $H_0^{(\mathrm{New})}$ and $H_0^{(\mathrm{Rel})}$ is of order $[\max_k\ e_k]^{2j}$.

where

$$H_j'^{\,(\mathrm{New})}(\mathbf{x}',\mathbf{y}') = H_j^{(\mathrm{New})}(\mathbf{x},\mathbf{y})\Big|_{\mathbf{x}=\mathbf{x}',\mathbf{y}=\mathbf{y}'}, \qquad H_j'^{\,(\mathrm{Rel})}(\mathbf{x}',\mathbf{y}') = H_j^{(\mathrm{Rel})}(\mathbf{x},\mathbf{y})\Big|_{\mathbf{x}=\mathbf{x}',\mathbf{y}=\mathbf{y}'}. \tag{6.43}$$

The scaling transformation (6.40) introduces in the Hamiltonians the power expansion in ϵ. This means that the natural reordering of the power series as homogeneous polynomials corresponds exactly to the use of a parameter, and, for this reason, in the following we continue to use the Hamiltonians in the form (6.36).

Now the classical way to proceed is to introduce the canonical transformation to action-angle variables

$$x_j = \sqrt{2\Phi_j}\cos\varphi_j, \qquad y_j = \sqrt{2\Phi_j}\sin\varphi_j \tag{6.44}$$

for $j = 1, 2$. The Hamiltonians (6.36) become

$$\begin{aligned} H(\mathbf{\Phi},\boldsymbol{\varphi};C,\mathbf{\Lambda}^*) &= \boldsymbol{\omega}^{(\mathrm{New})}\cdot\mathbf{\Phi} + \sum_{j=1}^{\infty}\varepsilon^j H_j^{(\mathrm{New})}(\mathbf{\Phi},\boldsymbol{\varphi}), \\ H_{\mathrm{rel}}(\mathbf{\Phi},\boldsymbol{\varphi};C,\mathbf{\Lambda}^*) &= \boldsymbol{\omega}^{(\mathrm{Rel})}\cdot\mathbf{\Phi} + \sum_{j=1}^{\infty}\varepsilon^j H_j^{(\mathrm{Rel})}(\mathbf{\Phi},\boldsymbol{\varphi}) \end{aligned} \tag{6.45}$$

where

$$\begin{aligned} H_j^{(\mathrm{New})}(\mathbf{\Phi},\boldsymbol{\varphi}) &= \sum_{\substack{\boldsymbol{\alpha}\in\mathbb{Z}_+^2 \\ |\alpha|=j}}\ \sum_{\mathbf{k}\in\mathcal{K}\subset\mathbb{Z}^2} c_{\boldsymbol{\alpha},\mathbf{k}}^{(\mathrm{New})}\Phi_1^{\alpha_1}\Phi_2^{\alpha_2}\exp(\sqrt{-1}(\mathbf{k}\cdot\boldsymbol{\varphi})), \\ H_j^{(\mathrm{Rel})}(\mathbf{\Phi},\boldsymbol{\varphi}) &= \sum_{\substack{\boldsymbol{\alpha}\in\mathbb{Z}_+^2 \\ |\alpha|=j}}\ \sum_{\mathbf{k}\in\mathcal{K}\subset\mathbb{Z}^2} c_{\boldsymbol{\alpha},\mathbf{k}}^{(\mathrm{Rel})}\Phi_1^{\alpha_1}\Phi_2^{\alpha_2}\exp(\sqrt{-1}(\mathbf{k}\cdot\boldsymbol{\varphi})), \end{aligned} \tag{6.46}$$

where $c_{\boldsymbol{\alpha},\mathbf{k}}^{(\mathrm{New})}$ and $c_{\boldsymbol{\alpha},\mathbf{k}}^{(\mathrm{Rel})}$ are suitable coefficients, $|\boldsymbol{\alpha}| = \alpha_1 + \alpha_2$ and where $\mathcal{K}$ is a finite subset of $\mathbb{Z}^2$ defined as

$$\mathcal{K} = \{\mathbf{k}\in\mathbb{Z}^2 \quad | \quad k_l \in \{-2\alpha_l, -2\alpha_l + 2, ..., 2\alpha_l - 2, 2\alpha_l\} \quad \text{for} \quad l = 1, 2\}. \tag{6.47}$$

In the new action-angle variables $\mathbf{\Phi}, \boldsymbol{\varphi}$, the secular Hamiltonians (6.45) have the form required to be studied with the tools discussed in chapter 4.

6.4 Secular evolution in action-angle coordinates

The secular Hamiltonians (6.45) have the form of a perturbed system of harmonic oscillators, and thus we can construct a Birkhoff normal form, by means of Lie series. Finally, an analytical integration of the action-angle equations will allow us to check the accuracy of our secular approximation, by comparing it to the direct numerical integration done in chapter 3.

Because the classical and relativistic Hamiltonian (6.45) have the same shape, in the following we do not distinguish the two cases and will consider the following Hamiltonian

$$H(\mathbf{\Phi},\boldsymbol{\varphi};C,\mathbf{\Lambda}^*) = \boldsymbol{\omega}\cdot\mathbf{\Phi} + \sum_{j=1}^{\infty} H_j(\mathbf{\Phi},\boldsymbol{\varphi}), \tag{6.48}$$

where

$$H_j(\boldsymbol{\Phi}, \boldsymbol{\varphi}) = \sum_{\substack{\boldsymbol{\alpha}\in\mathbb{Z}_+^2 \\ |\alpha|=j}} \sum_{\mathbf{k}\in\mathcal{K}\subset\mathbb{Z}^2} c_{\boldsymbol{\alpha},\mathbf{k}} \Phi_1^{\alpha_1} \Phi_2^{\alpha_2} \exp(\sqrt{-1}(\mathbf{k}\cdot\boldsymbol{\varphi})) \tag{6.49}$$

and where $c_{\boldsymbol{\alpha},\mathbf{k}}$ and $\boldsymbol{\omega}$ depend on the case and where $\mathcal{K}$ is defined in (6.47).

6.4.1 Birkhoff's normal form

The construction of the Birkhoff normal form via Lie series is explained in detail in chapter 4, thus we recall only some facts adapted to the present context.

Let r an integer greater than 1 such that the non-resonance condition

$$\mathbf{k}\cdot\boldsymbol{\omega} \neq 0 \tag{6.50}$$

is fulfilled for any $\mathbf{k}\in\mathbb{Z}^2$ such that $0 \leq |\mathbf{k}| \leq r+1$, where $|\mathbf{k}| = |k_1| + |k_2|$.

The aim is to give to the Hamiltonians the normal form at order $r \geq 1$

$$H^r(\boldsymbol{\Phi}^r, \boldsymbol{\varphi}^r; C, \boldsymbol{\Lambda}^*) = \boldsymbol{\omega}\cdot\boldsymbol{\Phi}^r + \sum_{j=1}^{r} \bar{H}_j^{(j-1)}(\boldsymbol{\Phi}^r) + \sum_{j=r+1}^{\infty} H_j^{(r)}(\boldsymbol{\Phi}^r, \boldsymbol{\varphi}^r), \tag{6.51}$$

where $\bar{H}_1^{(0)} \equiv \bar{H}_1$ and

- $\bar{H}_s^{(s-1)}$, for $s = 1, ..., r$, is the average of $H_s^{(s-1)}$ over the angles $\boldsymbol{\varphi}$ and it is a homogeneous polynomial of degree s in $\boldsymbol{\Phi}^r$;
- the un-normalized remainder terms $H_s^{(r)}$, where $s > r$, are homogeneous polynomials of degree $s+1$ in $\boldsymbol{\Phi}^r$ and they depend on the angles $\boldsymbol{\varphi}^r$.

To give to the Hamiltonians the normal form (6.51) at order r, we proceed by induction. Assuming that the Hamiltonian is in normal form up to a given order s, with $s < r$, which is trivially true for $s = 0$. Using the algorithm of Lie series transform, we can calculate the new Hamiltonian as

$$H^{s+1} = \exp(L_{\chi_{s+1}})H^s \tag{6.52}$$

and the new coordinates $(\boldsymbol{\Phi}^{s+1}, \boldsymbol{\varphi}^{s+1})$ are given by

$$\begin{aligned} \boldsymbol{\Phi}^{s+1} &= \exp(L_{\chi_{s+1}})\boldsymbol{\Phi}^s, \\ \boldsymbol{\varphi}^{s+1} &= \exp(L_{\chi_{s+1}})\boldsymbol{\varphi}^s. \end{aligned} \tag{6.53}$$

The generating function $\chi_{s+1}(\boldsymbol{\Phi}^{s+1}, \boldsymbol{\varphi}^{s+1})$ is determined by solving the following equation

$$\left\{\boldsymbol{\omega}\cdot\boldsymbol{\Phi}^{s+1}, \chi_{s+1}\left(\boldsymbol{\Phi}^{s+1}, \boldsymbol{\varphi}^{s+1}\right)\right\} + H_{s+1}^{(s)}\left(\boldsymbol{\Phi}^{s+1}, \boldsymbol{\varphi}^{s+1}\right) = \bar{H}_{s+1}^{(s)}\left(\boldsymbol{\Phi}^{s+1}\right). \tag{6.54}$$

If the condition (6.50) is satisfied, χ_{s+1} is well defined and analytic because the Fourier expansion of $H_{s+1}^{(s)}$ contains only a finite number of terms.

Let us remark that the Birkhoff normal form is not always convergent at high order, especially when the eccentricities are significant or the system is too close to a mean-motion resonance.

Assume that the non-resonance condition (6.50) is satisfied up to an order $r+1$ large enough and suppose we have constructed the normal form (6.51) to order r. Thus the remainder $R^{(r)}$

$$R^{(r)}(\mathbf{\Phi}^r, \boldsymbol{\varphi}^r) = \sum_{j=r+1}^{\infty} H_j^{(r)}(\mathbf{\Phi}^r, \boldsymbol{\varphi}^r) \tag{6.55}$$

is "small enough" and we can neglect it.

In this way, we obtain the optimal normal form at order r

$$H^r(\mathbf{\Phi}^r; C, \mathbf{\Lambda}^*) = \boldsymbol{\omega} \cdot \mathbf{\Phi}^r + \sum_{j=1}^{r} \bar{H}_j^{(j-1)}(\mathbf{\Phi}^r), \tag{6.56}$$

which is trivially integrable. Indeed the equations of motion for the truncated Hamiltonian are

$$\dot{\Phi}_j^r = 0, \qquad \dot{\varphi}_j^r = \frac{\partial H^r}{\partial \Phi_j^r} = \omega_j + \sum_{l=1}^{r} \frac{\partial \bar{H}_l^{(l-1)}}{\partial \Phi_j^r}, \tag{6.57}$$

for $j = 1, 2$.

6.4.2 Analytical integration

Using the equations in (6.57), we can compute the long-term evolution on the secular invariant torus, namely

$$\Phi_j^r(t) = \Phi_j^r(0), \qquad \varphi_j^r(t) = \varphi_j^r(0) + t\,\frac{\partial H^r}{\partial \Phi_j^r}(\mathbf{\Phi}^r(0)), \tag{6.58}$$

for $j = 1, 2$.

Now we have to come back to the original variables $(\mathbf{\Phi}, \boldsymbol{\varphi})$. As we have seen, the inverse of the transformation (6.53) is

$$\begin{aligned} \mathbf{\Phi}^s(t) &= \exp(-L_{\chi_{s+1}})\mathbf{\Phi}^{s+1}(t), \\ \boldsymbol{\varphi}^s(t) &= \exp(-L_{\chi_{s+1}})\boldsymbol{\varphi}^{s+1}(t), \end{aligned} \tag{6.59}$$

where $\chi_{s+1} = \chi_{s+1}(t)$ is a function of time, and so the relationship between $(\mathbf{\Phi}(t), \boldsymbol{\varphi}(t))$ and $(\mathbf{\Phi}^r(t), \boldsymbol{\varphi}^r(t))$ is given by

$$\begin{aligned} \mathbf{\Phi}(t) &= \exp(-L_{\chi_1(t)}) \circ \ldots \circ \exp(-L_{\chi_r(t)})\mathbf{\Phi}^r(t), \\ \boldsymbol{\varphi}(t) &= \exp(-L_{\chi_1(t)}) \circ \ldots \circ \exp(-L_{\chi_r(t)})\boldsymbol{\varphi}^r(t). \end{aligned} \tag{6.60}$$

6.5 Application to some extrasolar systems

Using the method described previously, we can calculate the evolution of $(\mathbf{\Phi}(t), \boldsymbol{\varphi}(t))$ both in the classical case than in the relativistic one.

Now we have to come back to the original orbital elements, and in particular we are interested in the evolution of the eccentricities. It is simple to prove that the relationship between the eccentricity e and $(\mathbf{\Phi}, \boldsymbol{\varphi})$ is given by

$$e_i(t) = \sqrt{1 - \left(1 - \frac{\eta_i^2(t) + \xi_i^2(t)}{2\Lambda_i^*}\right)^2}, \tag{6.61}$$

for $i = 1, 2$, where

$$\begin{aligned} \boldsymbol{\eta}(t) &= R\mathbf{x}(t), \qquad x_i = \sqrt{2\Phi_i}\cos\varphi_i, \\ \boldsymbol{\xi}(t) &= R\mathbf{y}(t), \qquad y_i = \sqrt{2\Phi_i}\sin\varphi_i, \end{aligned} \tag{6.62}$$

and where the matrix R is defined in (6.34).

Thus, the analytical integration via normal form actually reduces to a transformation of the initial conditions to secular action-angles coordinates, the computation of the flow at time t in these coordinates, followed by a transformation back to the original orbital elements.

We apply this method to the extrasolar systems that we have studied in chapter 3. In particular, after having constructed the secular approximation, we decide to construct a Birkhoff normal form up to order $r = 5$, which corresponds to taking into account the secular variables up to order 12 [4]. For simplicity we set the average values of the semi-major axes (that we remember are fixed) equal to their initial values.

The results are shown in Figure 6.1: the red curve represents the evolution of the eccentricity in the classical case, while the blue curve represents the evolution of the eccentricity in the relativistic case. Thanks to the speed semi-analytical integration, we are able to get the system dynamics on longer time than in the numerical case.

As expected, the differences between the classical case and the relativistic case are similar to those found in the case of numerical integration.

6.5.1 Comparison between semi-analytical integration and direct numerical integration

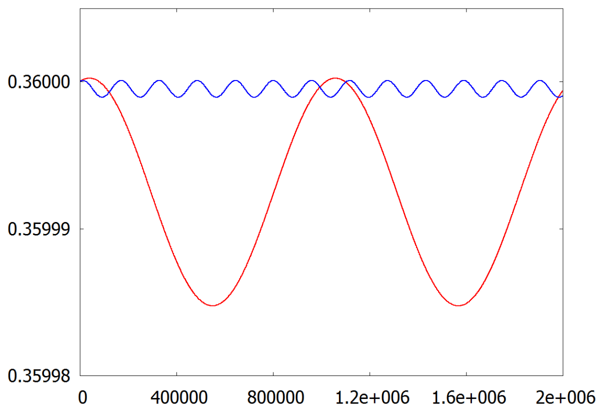
To validate our results, we will compare our semi-analytical integration with the direct numerical integration. The results are shown in Figure 6.2: the red curve represents the evolution of the eccentricity in the classical case obtained via numerical integration, the blue curve the relativistic case via numerical integration, the green curve the classical case via semi-analytical integration and, finally, the yellow curve the relativistic case via semi-analytical integration.

As one can see, the results obtained with the two methods are in agreement. In particular, although the truncations of series expansions involved in the construction of the secular system are made without estimates of the remainders, the dynamics of the full system is well represented by the dynamics of the secular system.

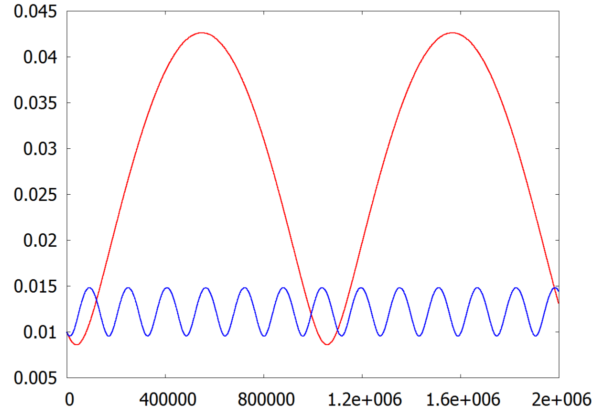
The main difference between the two methods is in the running time of the programs to achieve the results: the semi-analytic integration in fact turns out to be much faster than numerical integration. The execution times of the programs used for the numerical and the semi-analytical integration of the equations of motion are given in Table 6.1 (note that the final integration times and the integration step size are different for the various extrasolar systems). On the other hand, in the case of semi-analytical integration, the use of the averaging principle to simplify the problem is justified only if we consider non-resonant systems.

Finally, it would certainly be desirable to compare the dynamics obtained with the numerical method and that obtained by the semi-analytic method on longer time, but this would require a considerable amount of CPU time which we have not.

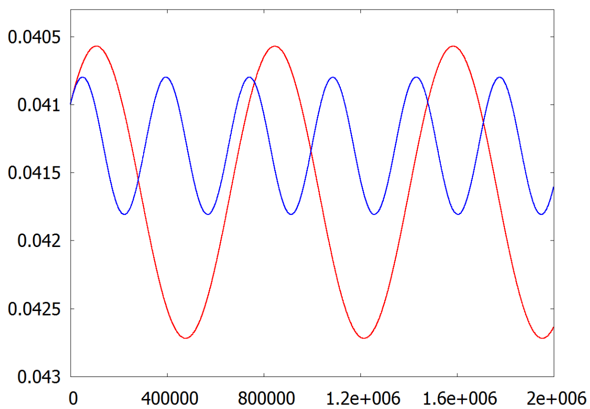
[4] We remember that *Libert* & *Henrard* have shown that an expansion up to order 12 in the eccentricities is usually required for reproducing the secular behavior of extrasolar planetary systems which are not close to a mean-motion resonance.



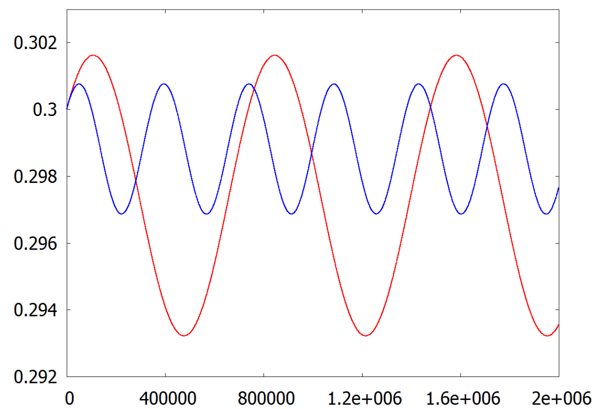
(a) Eccentricity planet HD 190360 b



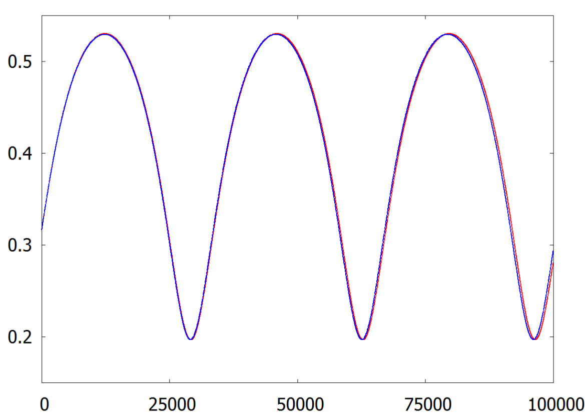
(b) Eccentricity planet HD 190360 c



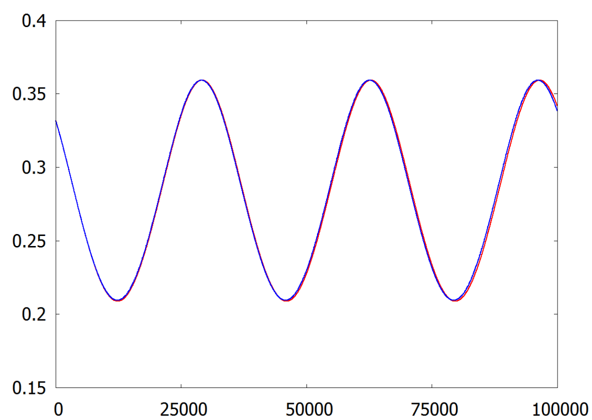
(c) Eccentricity planet HD 11964 b



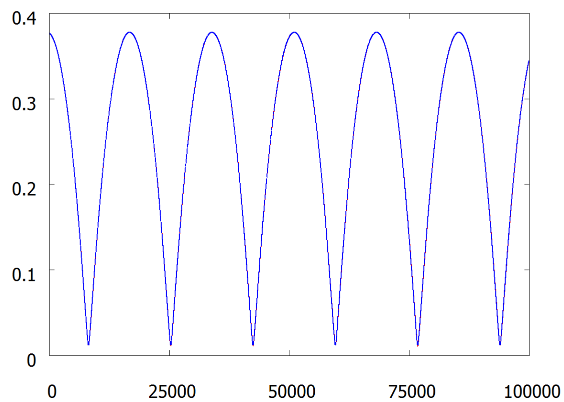
(d) Eccentricity planet HD 11964 c



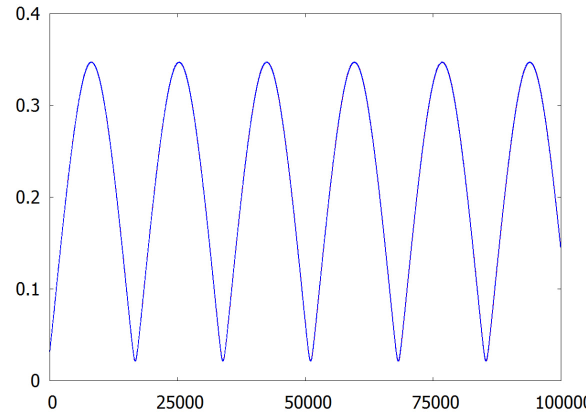
(e) Eccentricity planet HD 169830 b



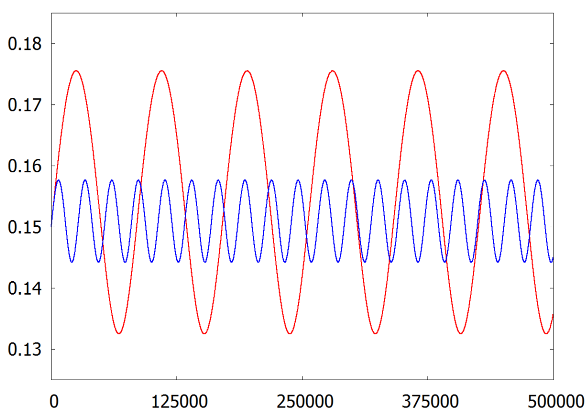
(f) Eccentricity planet HD 169830 c



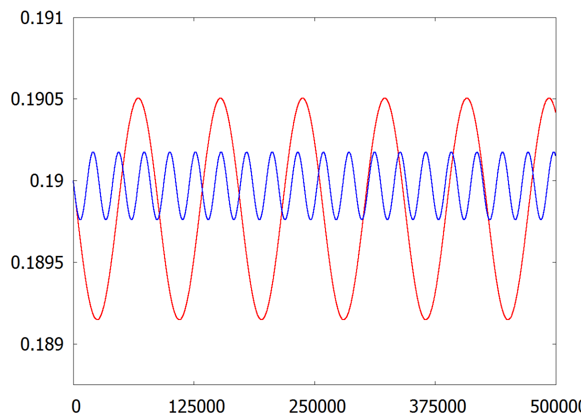
(g) Eccentricity planet HD 12661 b



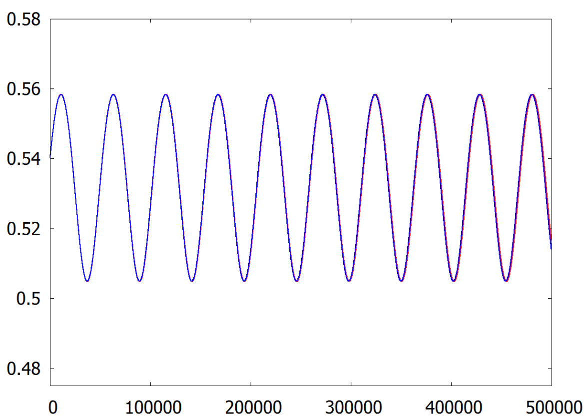
(h) Eccentricity planet HD 12661 c



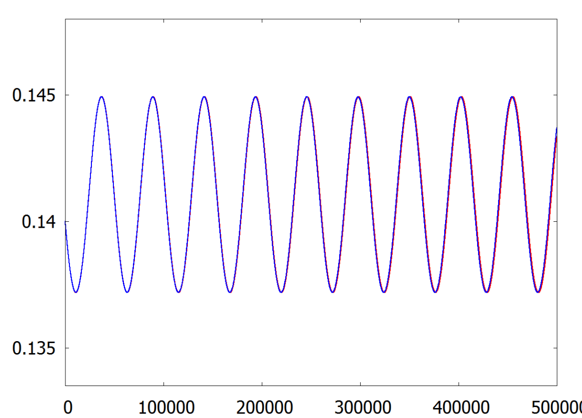
(i) Eccentricity planet BD 082823 b



(j) Eccentricity planet BD 082823 c

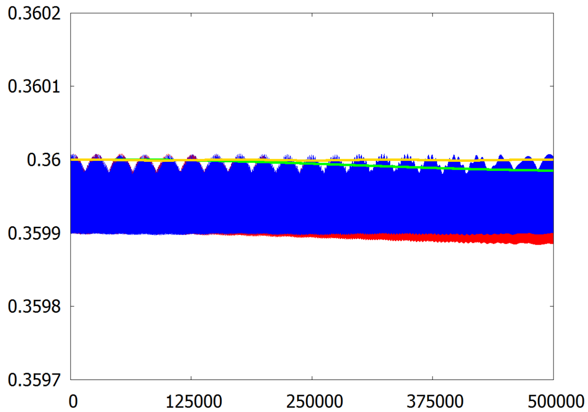


(k) Eccentricity planet HIP 5158 b

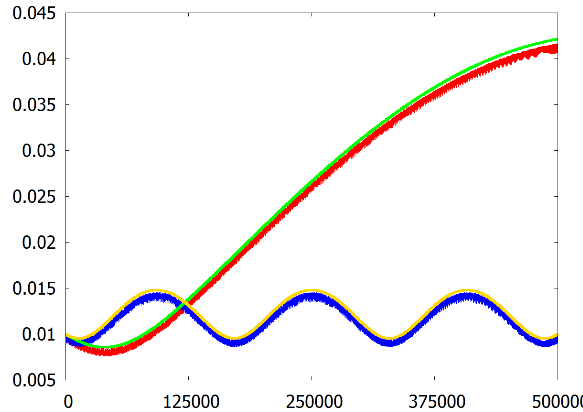


(l) Eccentricity planet HIP 5158 c

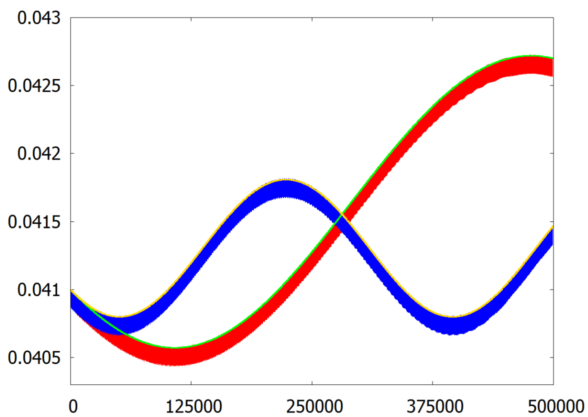
Figure 6.1: Comparison of the evolution of the eccentricity in the classical case and in the relativistic one (semi-analytical integration).



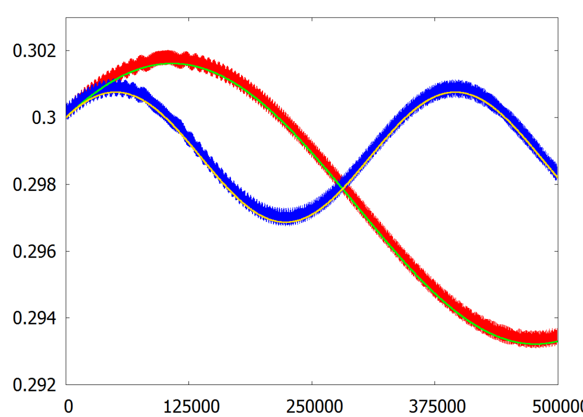
(a) Eccentricity planet HD 190360 b



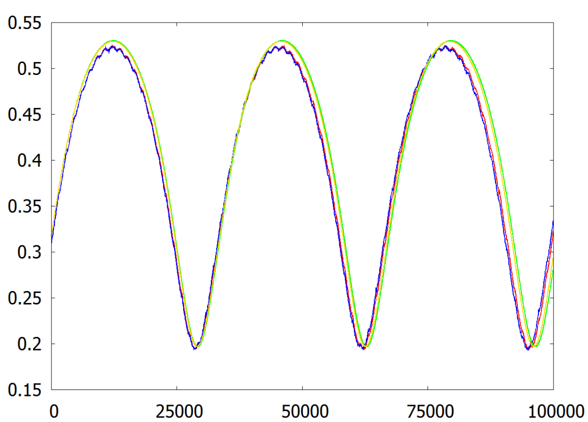
(b) Eccentricity planet HD 190360 c



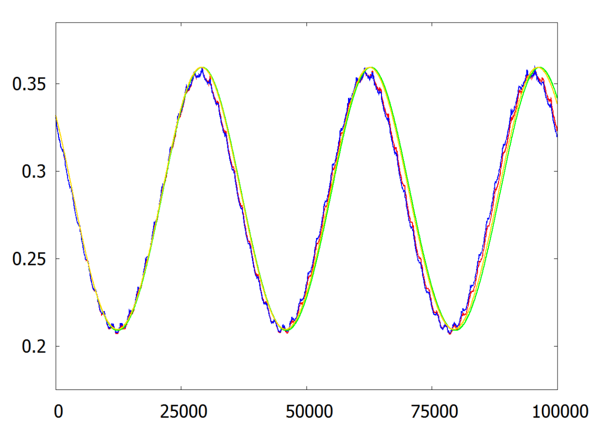
(c) Eccentricity planet HD 11964 b



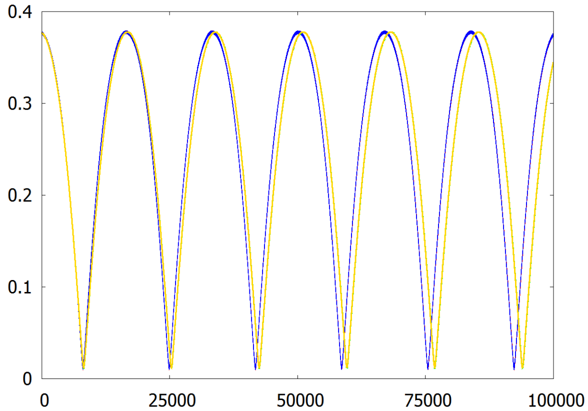
(d) Eccentricity planet HD 11964 c



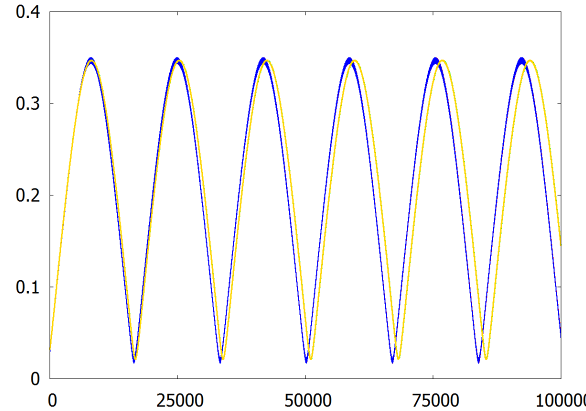
(e) Eccentricity planet HD 169830 b



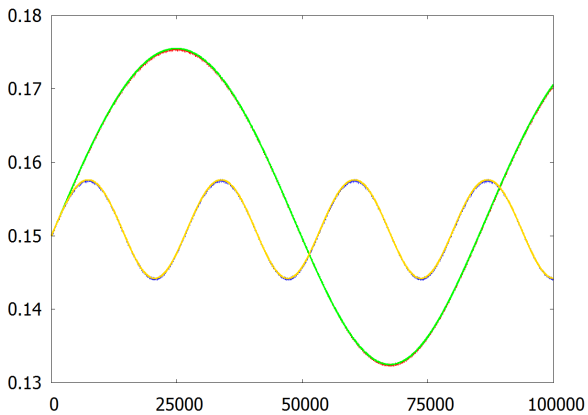
(f) Eccentricity planet HD 169830 c



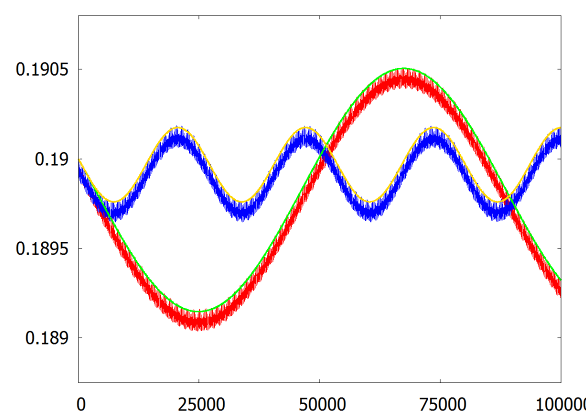
(g) Eccentricity planet HD 12661 b



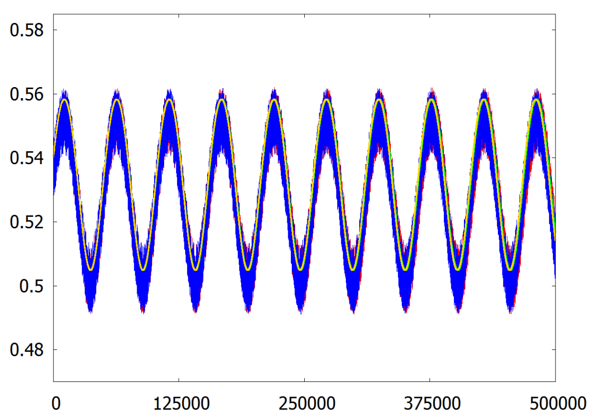
(h) Eccentricity planet HD 12661 c



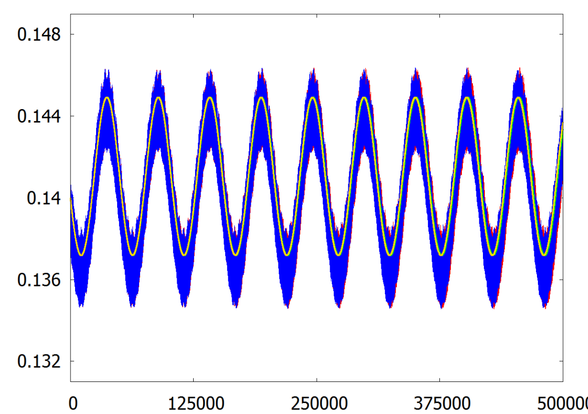
(i) Eccentricity planet BD 082823 b



(j) Eccentricity planet BD 082823 c



(k) Eccentricity planet HIP 5158 b



(l) Eccentricity planet HIP 5158 c

Figure 6.2: Comparison of the evolution of the eccentricity obtained in different way, i.e. via numerical integration and via semi-analytical integration. See text for more details.

Table 6.1: The following tables list the execution times of the programs used for the numerical and the semi-analytical integration of the equations of motion (see text for more details).

- **HD 190360**

  Analytical Integration (Classical and Relativistic case): 59.2 sec
  Numerical Integration:

Classical case	Relativistic case	Simplified relativistic case
1 h 22 min 37.0 sec	10 h 31 min 27.0 sec	6 h 43 min 20.8 sec

- **HD 11964**

  Analytical Integration (Classical and Relativistic case): 59.2 sec
  Numerical Integration:

Classical case	Relativistic case	Simplified relativistic case
41 min 20.4 sec	5 h 16 min 2.0 sec	3 h 21 min 59.8 sec

- **HD 169830**

  Analytical Integration (Classical and Relativistic case): 36.7 sec
  Numerical Integration:

Classical case	Relativistic case	Simplified relativistic case
3 min 19.8 sec	25 min 15.8 sec	16 min 12.9 sec

- **HD 12661**

  Analytical Integration (Classical and Relativistic case): 36.7 sec
  Numerical Integration:

Classical case	Relativistic case	Simplified relativistic case
3 min 19.8 sec	25 min 15.8 sec	16 min 12.9 sec

- **BD 082823**

  Analytical Integration (Classical and Relativistic case): 41.0 sec
  Numerical Integration:

Classical case	Relativistic case	Simplified relativistic case
1 h 22 min 36.6 sec	10 h 33 min 16.2 sec	6 h 43 min 7.9 sec

- **HIP 5158**

  Analytical Integration (Classical and Relativistic case): 41.0 sec
  Numerical Integration:

Classical case	Relativistic case	Simplified relativistic case
16 min 39.0 sec	2 h 6 min 19.0 sec	1 h 21 min 4.5 sec

6.6 A priori criterion to determine the importance of the relativistic correction

We would like to find a criterion to determine a priori when the relativistic corrections are important, i.e. to determine a priori when the difference between the classical case and the relativistic case are not negligible.

To do this, the idea is to analyze the quadratic part of the classical secular Hamiltonian and of the relativistic secular one. Indeed, if there are significant differences among the coefficients in cases where relativistic effects are important and, vice-versa, if there are minimal differences among the coefficients in cases where relativistic effects are negligible, we can find the searched criteria relying only on the quadratic part of the secular Hamiltonians.

In particular, we remember that if $H_0^{(\mathrm{New})}$ and $H_0^{(\mathrm{Rel})}$ are the quadratic parts of the secular Hamiltonians (6.23)-(6.24):

$$\begin{aligned} H_0^{(\mathrm{New})}(\boldsymbol{\eta},\boldsymbol{\xi}) &= \boldsymbol{\eta}\cdot A_{(\mathrm{New})}\boldsymbol{\eta} + \boldsymbol{\xi}\cdot A_{(\mathrm{New})}\boldsymbol{\xi} \\ H_0^{(\mathrm{Rel})}(\boldsymbol{\eta},\boldsymbol{\xi}) &= \boldsymbol{\eta}\cdot A_{(\mathrm{Rel})}\boldsymbol{\eta} + \boldsymbol{\xi}\cdot A_{(\mathrm{Rel})}\boldsymbol{\xi}, \end{aligned} \tag{6.63}$$

then $A_{(\mathrm{New})}$ and $A_{(\mathrm{Rel})}$ are real symmetric 2×2 matrices such that:

$$A_{(\mathrm{Rel})} = A_{(\mathrm{New})} - \frac{1}{c^2}\cdot\begin{bmatrix} \frac{3\mathcal{G}^{3/2}(m_0+m_1)^{3/2}}{2(a_1^*)^{5/2}} & 0 \\ 0 & \frac{3\mathcal{G}^{3/2}(m_0+m_2)^{3/2}}{2(a_2^*)^{5/2}} \end{bmatrix}, \tag{6.64}$$

where we use the index 1 to indicate the innermost planet and the index 2 to indicate the outer planet. It is easy to see the that the coefficients of $\xi_1\xi_2$ and $\eta_1\eta_2$ are the same both in the classical and in the relativistic case. Moreover, it is simple to prove that the unit of measurement of $A_{(\mathrm{Rel})}$ and $A_{(\mathrm{New})}$ is yr^{-1} (where yr denotes the year).

If we denote by $A^{(1)}$ the coefficient of ξ_1^2 and η_1^2 and by $A^{(2)}$ the coefficient of ξ_2^2 and η_2^2, the relative difference between the coefficients in the classical and in the relativistic case is given by

$$\frac{A^{(i)}_{(\mathrm{Rel})} - A^{(i)}_{(\mathrm{New})}}{A^{(i)}_{(\mathrm{Rel})}}, \qquad i = 1, 2. \tag{6.65}$$

The data are shown in the following tables.

As expected, in cases HD 190360, HD 11964 and BD 082823 the relative difference between the relativistic and the classical coefficients is remarkable (e.g. the relative difference is greater than 0.45 for the coefficient of ξ_1^2 and η_1^2, and it is greater than 0.035 for the coefficient of ξ_2^2 and η_2^2), while in the other cases the difference is negligible (e.g. the relative difference is less than 0.002).

Table 6.2 We report here the expansion of the secular classical and relativistic Hamiltonians of the extrasolar systems up to degree 2 in $(\boldsymbol{\xi}, \boldsymbol{\eta})$. Moreover, we report the relative difference between the coefficients in the classical and in the relativistic case, given by formula (6.65).

- **HD 190360**

ξ_1	ξ_2	η_1	η_2	Classical case (yr^{-1})	Relativistic case (yr^{-1})
2	0	0	0	$-2.5241044053489410 \cdot 10^{-6}$	$-1.9358830497117772 \cdot 10^{-5}$
1	1	0	0	$1.7066351622958102 \cdot 10^{-8}$	$1.7066351622958102 \cdot 10^{-8}$
0	2	0	0	$-1.7320564857123492 \cdot 10^{-8}$	$-2.0570535193981765 \cdot 10^{-8}$
0	0	2	0	$-2.5241044053489410 \cdot 10^{-6}$	$1.9358830497117772 \cdot 10^{-5}$
0	0	1	1	$1.7066351622958102 \cdot 10^{-8}$	$1.7066351622958102 \cdot 10^{-8}$
0	0	0	2	$-1.7320564857123492 \cdot 10^{-8}$	$-2.0570535193981765 \cdot 10^{-8}$

Relative difference:

- coefficient of ξ_1^2, η_1^2: 0.869614830000979
- coefficient of ξ_2^2, η_2^2: 0.157991530420127

- **HD 11964**

ξ_1	ξ_2	η_1	η_2	Classical case (yr^{-1})	Relativistic case (yr^{-1})
2	0	0	0	$-4.6272830094286514 \cdot 10^{-6}$	$-9.0514410721029731 \cdot 10^{-6}$
1	1	0	0	$1.5492955932942727 \cdot 10^{-7}$	$1.5492955932942727 \cdot 10^{-7}$
0	2	0	0	$-1.5824755540437131 \cdot 10^{-7}$	$-1.6450650937015877 \cdot 10^{-7}$
0	0	2	0	$-4.6272830094286514 \cdot 10^{-6}$	$-9.0514410721029731 \cdot 10^{-6}$
0	0	1	1	$1.5492955932942727 \cdot 10^{-7}$	$1.5492955932942727 \cdot 10^{-7}$
0	0	0	2	$-1.5824755540437131 \cdot 10^{-7}$	$-1.6450650937015877 \cdot 10^{-7}$

Relative difference:

- coefficient of ξ_1^2, η_1^2: 0.488779413955399
- coefficient of ξ_2^2, η_2^2: 0.038046846837556

- **HD 169830**

ξ_1	ξ_2	η_1	η_2	Classical case (yr^{-1})	Relativistic case (yr^{-1})
2	0	0	0	$-1.3244339842965693 \cdot 10^{-4}$	$-1.3270515925390962 \cdot 10^{-4}$
1	1	0	0	$4.3051477118903886 \cdot 10^{-5}$	$4.3051477118903886 \cdot 10^{-5}$
0	2	0	0	$-4.4802688973893151 \cdot 10^{-5}$	$-4.4808982225461097 \cdot 10^{-5}$
0	0	2	0	$-1.3244339842965693 \cdot 10^{-4}$	$-1.3270515925390962 \cdot 10^{-4}$
0	0	1	1	$4.3051477118903886 \cdot 10^{-5}$	$4.3051477118903886 \cdot 10^{-5}$
0	0	0	2	$-4.4802688973893151 \cdot 10^{-5}$	$-4.4808982225461097 \cdot 10^{-5}$

Relative difference:

- coefficient of ξ_1^2, η_1^2: 0.001972499228548
- coefficient of ξ_2^2, η_2^2: 0.000140446206439

- **HD 12661**

ξ_1	ξ_2	η_1	η_2	Classical case (yr^{-1})	Relativistic case (yr^{-1})
2	0	0	0	$-1.8980723258516056 \cdot 10^{-4}$	$-1.8997180723091949 \cdot 10^{-4}$
1	1	0	0	$1.3858019683402154 \cdot 10^{-4}$	$1.3858019683402154 \cdot 10^{-4}$
0	2	0	0	$-1.5827736051885459 \cdot 10^{-4}$	$-1.5828720141461170 \cdot 10^{-4}$
0	0	2	0	$-1.8980723258516056 \cdot 10^{-4}$	$-1.8997180723091949 \cdot 10^{-4}$
0	0	1	1	$1.3858019683402154 \cdot 10^{-4}$	$1.3858019683402154 \cdot 10^{-4}$
0	0	0	2	$-1.5827736051885459 \cdot 10^{-4}$	$-1.5828720141461170 \cdot 10^{-4}$

Relative difference:

- coefficient of ξ_1^2, η_1^2: 0.000866310891905
- coefficient of ξ_2^2, η_2^2: 0.000062171139986

- **BD 082823**

ξ_1	ξ_2	η_1	η_2	Classical case (yr^{-1})	Relativistic case (yr^{-1})
2	0	0	0	$-3.6840491507349428 \cdot 10^{-5}$	$-1.1665139285063237 \cdot 10^{-4}$
1	1	0	0	$1.4992857272466206 \cdot 10^{-6}$	$1.4992857272466206 \cdot 10^{-6}$
0	2	0	0	$-1.4419273424417406 \cdot 10^{-6}$	$-1.5973446899364618 \cdot 10^{-6}$
0	0	2	0	$-3.6840491507349428 \cdot 10^{-5}$	$-1.1665139285063237 \cdot 10^{-4}$
0	0	1	1	$1.4992857272466206 \cdot 10^{-6}$	$1.4992857272466206 \cdot 10^{-6}$
0	0	0	2	$-1.4419273424417406 \cdot 10^{-6}$	$-1.5973446899364618 \cdot 10^{-6}$

Relative difference:

- coefficient of ξ_1^2, η_1^2: 0.684183012246392
- coefficient of ξ_2^2, η_2^2: 0.097297313769455

- **HIP 5158**

ξ_1	ξ_2	η_1	η_2	Classical case (yr^{-1})	Relativistic case (yr^{-1})
2	0	0	0	$-7.2321643316479597 \cdot 10^{-5}$	$-7.2407635908971517 \cdot 10^{-5}$
1	1	0	0	$3.7796386950599819 \cdot 10^{-6}$	$3.7796386950599819 \cdot 10^{-6}$
0	2	0	0	$-2.3736218727755973 \cdot 10^{-6}$	$-2.3740222285574030 \cdot 10^{-6}$
0	0	2	0	$-7.2321643316479597 \cdot 10^{-5}$	$-7.2407635908971517 \cdot 10^{-5}$
0	0	1	1	$3.7796386950599819 \cdot 10^{-6}$	$3.7796386950599819 \cdot 10^{-6}$
0	0	0	2	$-2.3736218727755973 \cdot 10^{-6}$	$-2.3740222285574030 \cdot 10^{-6}$

Relative difference:

- coefficient of ξ_1^2, η_1^2: 0.001187617734130
- coefficient of ξ_2^2, η_2^2: 0.000168640283562

Thus we can use the quadratic part of the secular Hamiltonians, and in particular the relative difference (6.65), to find the searched criteria.

To do this, we have to express the coefficients of $A_{(\mathrm{New})}$ in terms of the semi-major axes a_1^* and a_2^* (where $a_1^* < a_2^*$), and in terms of the masses of the two planets m_1 and m_2 and of the central star m_0. In particular, for simplicity we consider only coplanar systems, i.e. we consider only systems in which the inclinations of the two planets are equal ($i_1 = i_2 = 0$).

The coefficients of $A_{(\mathrm{Rel})}$ are then given by (6.64).

To express the coefficients of $A_{(\mathrm{New})}$ in terms a_1^*, a_2^*, m_0, m_1 and m_2, we have to calculate the following quantity:

$$H_{\mathrm{sec}} = \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} H_1 \mathrm{d}M_1 \mathrm{d}M_2, \tag{6.66}$$

where M_1 and M_2 are the mean anomalies, and H_1 is the perturbation part (3.23)

$$H_1 = \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{m_0} - \frac{\mathcal{G} m_1 m_2}{\parallel \mathbf{r}_1 - \mathbf{r}_2 \parallel}. \tag{6.67}$$

The development of the perturbation part H_1 as a function of M_1 and M_2 is provided in chapter 5.

We start with the averaging of the scalar product $\mathbf{p}_1 \cdot \mathbf{p}_2$. Using the fact that $\mathbf{p}_i$ has the form

$$\mathbf{p}_i = \left(\frac{1}{m_i} + \frac{1}{m_0 + m_{3-i}}\right)^{-1} \dot{\mathbf{r}}_i - \frac{m_i m_{3-i}}{m_0 + m_1 + m_2}\dot{\mathbf{r}}_{3-i} \tag{6.68}$$

for $i = 1, 2$, where $\dot{\mathbf{r}}_i$ is the astrocentric velocity of the i-th planet, the product $\mathbf{p}_1 \cdot \mathbf{p}_2$ can be reduced to a sum of terms of the form $\dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_j$. In particular, it is simple to prove that the average of expressions of the form $\dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_j$ is (see *Brouwer & Clemence (1961)* for more details):

$$\frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_j \mathrm{d}M_1 \mathrm{d}M_2 = \delta_{i,j} a_i^2 n_i^2, \qquad i, j = 1, 2, \tag{6.69}$$

where n_i denote mean motions of the planets and $\delta_{i,j}$ stands for the Kronecker delta. Thus, the scalar product $\mathbf{p}_1 \cdot \mathbf{p}_2$ depends on L_i only, which are integrals of the secular mode. We have therefore shown that the scalar product $\mathbf{p}_1 \cdot \mathbf{p}_2$ does not contribute to the secular dynamics of the system.

Now we have to calculate the more difficult part of the problem (see also *Migaszewski and Goździewski (2008a)* for more details), i.e.

$$\frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} -\frac{\mathcal{G} m_1 m_2}{\parallel \mathbf{r}_1 - \mathbf{r}_2 \parallel} \mathrm{d}M_1 \mathrm{d}M_2. \tag{6.70}$$

The development of the inverse of the distance of the two planets as a function of M_1 and M_2 is provided in section 5.3.3 (we consider only systems in which the inclinations of the two planets are equal). For our purposes, it is sufficient to develop each term up to order 2 in eccentricities e_1 and e_2. In addition, we decide to stop the developments up to order 6 in α, where α is the ratio of the semi-major axis:

$$\alpha = \frac{a_1^*}{a_2^*}, \qquad \alpha < 1. \tag{6.71}$$

Basically, the problem has been reduced to the calculation of definite integrals from products of trigonometric functions $\sin(x)$ and $\cos(x)$ in some natural powers. To integrate (6.70) term by term, we used the mathematical software MATHEMATICA.

The coefficient $A^{(i)}_{\text{New}}$ of ξ_i^2 and η_i^2 is given by

$$A^{(i)}_{\text{New}} = -\frac{m_1 m_2 \sqrt{\mathcal{G}(m_0+m_i)}}{m_0 m_i a_2^* \sqrt{a_i^*}} \left[\frac{3}{8}\alpha^2 + \frac{45}{64}\alpha^4 + \frac{525}{512}\alpha^6\right] + o(\alpha^6), \tag{6.72}$$

and the coefficient $A^{(i)}_{\text{Rel}}$ is given by

$$A^{(i)}_{\text{Rel}} = A^{(i)}_{\text{New}} - \frac{1}{c^2}\frac{3\mathcal{G}^{3/2}(m_0+m_i)^{3/2}}{2(a_i^*)^{5/2}}, \tag{6.73}$$

for $i = 1, 2$, where α is defined in (6.71). Moreover, the coefficient of $\xi_1\xi_2$ and $\eta_1\eta_2$ is given by

$$\frac{\sqrt{\mathcal{G}} m_1 m_2}{m_0}\left(\frac{(m_0+m_1)(m_0+m_2)}{a_1^*(a_2^*)^5}\right)^{1/4}\left[\frac{15}{16}\alpha^3 + \frac{105}{128}\alpha^5\right] + o(\alpha^6). \tag{6.74}$$

We test the above formulæ on the extrasolar systems that we have considered. In particular, we calculate the relative error between the coefficients $\tilde{A}^{(i)}$ given by formulæ (6.72)-(6.73) and the coefficients $A^{(i)}$ given by Table 6.2:

$$\frac{A^{(i)}_{(\text{New})} - \tilde{A}^{(i)}_{(\text{New})}}{A^{(i)}_{(\text{New})}}, \qquad \frac{A^{(i)}_{(\text{Rel})} - \tilde{A}^{(i)}_{(\text{Rel})}}{A^{(i)}_{(\text{Rel})}}, \qquad i = 1, 2. \tag{6.75}$$

The results are reported in Table 6.3. In all cases we have considered, the relative error is very small both in the classical case than in the relativistic one.

Using (6.72) and (6.73), the relative difference (6.65) between the coefficients in the classical and in the relativistic case is given by

$$\frac{A^{(i)}_{(\text{Rel})} - A^{(i)}_{(\text{New})}}{A^{(i)}_{(\text{Rel})}} = \frac{3\mathcal{G}a_2^* m_0 m_i (m_0+m_i)}{2c^2(a_i^*)^2 m_1 m_2 \mathcal{D}(\alpha) + 3\mathcal{G}a_2^* m_0 m_i (m_0+m_i)}, \tag{6.76}$$

for $i = 1, 2$, where

$$\mathcal{D}(\alpha) = \frac{3}{8}\alpha^2 + \frac{45}{64}\alpha^4 + \frac{525}{512}\alpha^6. \tag{6.77}$$

Comparing the data in Table 6.2 and the corresponding plots in Figure 6.1, we think that, in the case of a coplanar three-body system, the relativistic correction on the i-th planet may be important if the relative difference is greater than δ, i.e. if

$$\frac{3\mathcal{G}a_2^* m_0 m_i (m_0+m_i)}{2c^2(a_i^*)^2 m_1 m_2 \mathcal{D}(\alpha) + 3\mathcal{G}a_2^* m_0 m_i (m_0+m_i)} > \delta, \tag{6.78}$$

where

- if $A^{(i)}_{(\text{New})} < -1\times 10^{-5}\, yr^{-1}$, then $\delta = 0.0025$;
- if $-1\times 10^{-5}\, yr^{-1} \le A^{(i)}_{(\text{New})} < -5\times 10^{-7}\, yr^{-1}$, then $\delta = 0.01$;
- if $-5\times 10^{-7} \le A^{(i)}_{(\text{New})}\, yr^{-1}$, then $\delta = 0.025$.

It is also simple to see that the condition (6.78) is equivalent to the condition

$$\frac{c^2(a_i^*)^2 m_1 m_2 \mathcal{D}(\alpha)}{\mathcal{G} a_2^* m_0 m_i (m_0 + m_i)} < \frac{3}{2}\left(\frac{1}{\delta} - 1\right). \tag{6.79}$$

The criterion introduced above is clearly heuristic and quite rough, nevertheless we think it is useful to discriminate the cases in which the relativistic corrections are important from those in which they are not.

Table 6.3 We report here the relative error (6.75) between the coefficient $A^{(i)}$ of ξ_i^2 and η_i^2 given in Table 6.1, and the coefficient $\tilde{A}^{(i)}$ given by formula (6.72)-(6.73).

- **HD 190360**

	Classical case	Relativistic case
coefficient of: ξ_1^2, η_1^2	0.000682507716323	0.000695726790969
coefficient of: ξ_2^2, η_2^2	0.000682510876349	0.000684913514844

- **HD 11964**

	Classical case	Relativistic case
coefficient of: ξ_1^2, η_1^2	0.000683021126193	0.000690200202867
coefficient of: ξ_2^2, η_2^2	0.000683022224624	0.000683581131333

- **HD 169830**

	Classical case	Relativistic case
coefficient of: ξ_1^2, η_1^2	0.001132737282618	0.001131879216172
coefficient of: ξ_2^2, η_2^2	0.001132739159952	0.001132678064295

- **HD 12661**

	Classical case	Relativistic case
coefficient of: ξ_1^2, η_1^2	0.004566038936478	0.004562687772306
coefficient of: ξ_2^2, η_2^2	0.004566037394066	0.004565796896265

- **BD 082823**

	Classical case	Relativistic case
coefficient of: ξ_1^2, η_1^2	0.000683617301988	0.000693258483257
coefficient of: ξ_2^2, η_2^2	0.000683618178821	0.000684989416831

- **HIP 5158**

	Classical case	Relativistic case
coefficient of: ξ_1^2, η_1^2	0.000690986292650	0.000690994290910
coefficient of: ξ_2^2, η_2^2	0.000691025215660	0.000691026364531

6.7 Conclusions

Except for very precise simulations, in the study of the dynamics of Solar system, we just deal with Newtonian mechanics and the relativistic effects are in general not taken into account in orbit computations. Our planetary system is in fact mainly composed by an inner set of terrestrial low mass planets and an outer set of giant very massive planets. The large distances between the planets and the Sun and the fact that the masses of the planets are relatively small implies that the effects of relativity are so small as not to worry about it. Moreover, the perturbations due to the larger asteroids are almost always much more significant than the relativistic corrections.

When the relativistic effects are taken into account, only the effects due to the Sun are considered and, also in this case, the secular relativistic effects generated by the Sun are appreciable only for the argument of the perihelion and mean anomaly of the inner Solar system. Thus, in the case of the Solar system, Newtonian theory provides very reliable results.

Conversely, these facts are not true in general in the case of extrasolar systems. As we have seen, in general the relativistic corrections due to the star are important and, in some cases, are indispensable (specially when semi-major axes are of the order of 10^{-1} AU or less) in orbit computations of extrasolar planets.

After the discovery in 1995 of the first extrasolar planet in orbit around a main-sequence star, the number of known-extrasolar planets did not cease to grow. The planets so far discovered are big and most of them have orbits close to the central stars. By virtue of their small semi-major axes and high eccentricities, extrasolar planetary systems with multiple planets allow for General Relativity to exhibit much more pronounced effects than in the case of the Solar system. As a result, these systems provide a new test of General Relativity.

In this thesis, we have analyzed the long-term evolution of several exoplanetary systems both in the classical case than in the relativistic one. We have tried to evaluate how General Relativity theory affects the orbital dynamics of the extrasolar system, in the limit of point masses, by means of a series of numerical tests. In particular, we have limited ourselves to consider coplanar and non-resonant extrasolar systems, consist of a central star and two planets orbiting around it.

To describe the approximate dynamics of a system of "point-like masses" due to their mutual gravitational interactions, including general relativistic effects, we use the *Einstein-Infeld-Hoffmann* Hamiltonian (up to the $1/c^2$ approximation). However, it is important to remember that, in general, the corrections provided by the EIH model to the classic model still do not cover all physics governing the dynamics of such systems.

Thanks to numerical integration of the Hamilton's equations, we note that, in some of the cases studied, the difference between predictions of General Relativity and by the classical model are significant, while in some other cases both theories give practically the same outcome. As expected, looking at the element of the examined systems, we may conclude that the corrections become very important for systems with the innermost planet close to the star, with other body relatively distant. The results show in fact that the relativistic effects can accumulate over time to induce substantial changes in the dynamics. In particular, we find that for the systems HD 190360, HD 11964 and BD 082823 the relativistic corrections are important, while for the systems HD 169830, HD 12661 and HIP 5158 the relativistic corrections are so small to be negligible.

Moreover, we note that, where the relativistic effects are important, they seem to provide "stability" to the system. These results are in agreement with those obtained by Laskar (*Laskar (2008)*) in the case of the Solar system. Quoting Laskar: (in the case of the inner Solar system) *"The difference of behavior of the secular system with and without general relativity (GR) is impressive.*

[...] The contribution of GR is thus essential in order to ensure the relative stability of Mercury."

Thus, following the results here presented, it seems that relativistic corrections are not unavoidable for all dynamical studies, but they are necessary for the precise dynamical modeling of close extrasolar planets. Quoting Sitarski (1983): *"...it seems that in all the modern investigations it is the very time to replace Newtonian equations of motions by those resulting from general relativity theory"*.

The major defect of the numerical integration is that it is CPU consuming and that the time required to integrate a system is very long. For this reason, we have therefore looked for a "semi-analytical" integrations of the Hamiltonian equations, using the tools provided by Hamiltonian system and by perturbation theory.

On the other hand, the relativistic Hamiltonian turns out to be uncomfortable to be treated using the canonical perturbation theory. Thus, we have looked for a simplification of the relativistic Hamiltonian, skipping the relativistic correction due to the mutual interactions of the two planetary masses. In particular, we assume that the mutual interactions between the star and the two planets are of relativistic type (i.e. we consider the relativistic corrections to the Newtonian gravity) and that the mutual interaction between the two planets is only of Newtonian type (i.e. we skip the relativistic corrections caused by the two planetary masses). As we have seen, this assumption leads to a substantial simplification to the relativistic Hamiltonian and, at the same time, the dynamic obtained by this simplified Hamiltonian is very similar to that described in the real one, at least numerically in the systems that we have considered (only in some particular cases the relativistic effects generated by the planets seem to have some importance).

Then, starting from the classical and the (simplified) relativistic secular Hamiltonians, we have computed a high-order Birkhoff normal form via Lie series, introducing action-angle coordinates for the secular variables. This enabled us to compute analytically the evolution on the secular invariant torus and to obtain the long-term evolution of the eccentricities.

To obtain the secular Hamiltonians, we have simple performed an average over the fast angles of the Hamiltonians, which corresponds to fixing the values of the semi-major axes. It is important to remember that, to obtain qualitative information on the long-term changes of the slowly varying orbital elements using the averaging principle, the extrasolar systems must be non-resonant.

As a result, for all the systems that are not too close to a mean-motion resonance, we have shown an excellent agreement with the direct numerical integration of the full three-body problem.

Furthermore, evaluating the difference between the quadratic part of the secular classical Hamiltonian and the secular relativistic Hamiltonian, we have set up a simple (and rough) criterion to discriminate between the cases in which the relativistic corrections are important from those in which they are not. In particular, in the case of a coplanar three-body system, we think that the relativistic corrections on the i-th planet may be important if

$$\frac{c^2 a_i^2 m_1 m_2 \mathcal{D}(\alpha)}{\mathcal{G} a_2 m_0 m_i (m_0 + m_i)} < \frac{3}{2}\left(\frac{1}{\delta} - 1\right),$$

where $\mathcal{D}(\alpha)$ is given in (6.77) and δ is given in section 6.6.

Let us remark that these results could be extended to the spatial case with minor changes. Indeed, after the reduction of the angular momentum, the starting Hamiltonian would have exactly the same form as $H^{(\mathcal{T})}_{\mathrm{New}}$ and $H^{(\mathcal{T})}_{\mathrm{Rel}}$, defined in (6.11)-(6.12).

Finally, a natural extension to the present work would be the study of the secular relativistic evolution of systems that are close to or in a mean-motion resonance. The effects of mean-motion resonances are in fact of great impact on the long-term behavior of the exoplanetary system. Indeed the terms of the perturbation associated to mean-motion resonances have small frequencies and thus influence the secular behavior of the system. For this reason, a good description of the secular dynamics of an exoplanetary system should include a careful treatment of the influence of mean-motion resonances on the long-term evolution. Therefore, it is necessary to replace the classical circular approximation with a torus which is invariant up to order two in the masses, i.e. to replace the first order averaged Hamiltonian with the one at order two in the masses. For more details of the benefit of a second order approach, see for example *Laskar (1988)*, *Sansottera, Locatelli and Giorgilli (2013)* and *Libert and Sansottera (2013)*.

Moreover, having such a good analytical description of the orbits both in the classical than in the relativistic case, we can also study the effective stability of extrasolar planetary systems in the framework of the KAM and Nekhoroshev theories.

Appendix A

Summary of differential geometry

A *manifold* M is a topological space which satisfies the Hausdorf criterion, i.e. such that each point of M has an open neighborhood which has a continuous 1-1 map onto an open set of R^n, where n is the dimension of the manifold.

A *chart* is a homeomorphism h from $D \subset M$ to $U \subset \mathbb{R}^n$ which assigns a n-uple of coordinates (x^i) to a point $p \in D$:

$$\begin{aligned} h\colon D &\to U, \\ p &\mapsto h(p) = (x^1, ..., x^n). \end{aligned} \tag{A.1}$$

An *atlas* is a collection of charts whose domains cover the entire manifold.

The *tangent space* V_p of a manifold M in a point p is the vector space of all derivations. A *derivation* v is a map from the space $\mathcal{F}_p$ of C^∞ functions in p into the real numbers

$$\begin{aligned} v\colon \mathcal{F}_p &\to \mathbb{R}, \\ f &\mapsto v(f). \end{aligned} \tag{A.2}$$

which is linear and which satisfies the Leibniz rule. A coordinate basis of the tangent vector space is given by the partial derivate ∂_i, i.e. a tangent vector v can be expanded in this basis $v = v^i\partial_i$ and $v(f) = v^i\partial_i(f)$ for each $f \in \mathcal{F}$.

A dual vector w is a linear mapping assigning a real number to vector

$$\begin{aligned} w\colon V &\to \mathbb{R}, \\ v &\mapsto w(v). \end{aligned} \tag{A.3}$$

The space of dual vectors to a tangent vector space V is the dual space V^*. Specifically, the differential of a function $f \in \mathcal{F}$ is a dual vector defined by

$$\begin{aligned} \mathrm{d}f\colon V &\to \mathbb{R}, \\ v &\mapsto \mathrm{d}f(v) = v(f). \end{aligned} \tag{A.4}$$

Accordingly, the differential of the coordinate function x^i form a basis $\{\mathrm{d}x^i\}$ of the dual space, which is orthonormal to the coordinate basis $\{\delta_j\}$ of the tangent space, i.e. $\mathrm{d}x^i(\delta_j) = \delta_j(\mathrm{d}x^i) = \delta^i{}_j$.

A *tensor* T of rank (r, s) on M at p is a multi-linear mapping of r dual vector and s vectors into the real number:

$$T\colon \overbrace{V_p^* \times ... \times V_p^*}^{r} \times \overbrace{V_p \times ... \times V_p}^{s} \to \mathbb{R}. \tag{A.5}$$

We find that vectors are $(1,0)$ tensors, duals are $(0,1)$ tensors, and scalars are $(0,0)$ tensors. A basis for tensors of arbitrary rank (r,s) is obtained by the tensor product of suitably many elements of the bases $\{\delta_i\}$ of the tangent space and $\{dx^j\}$ of the dual space:

$$T = T^{i_1,\dots,i_r}{}_{j_1,\dots,j_s}\delta_{i_1} \otimes \dots \otimes \delta_{i_r} \otimes \mathrm{d}x^{j_1} \otimes \dots \otimes \mathrm{d}x^{j_s}. \tag{A.6}$$

The *metric tensor* $g = g_{ik}\mathrm{d}x^i \otimes \mathrm{d}x^j$ is a symmetric, non degenerate tensor of rank $(0,2)$, i.e it satisfies $g(x,y) = g(y,x)$ and $g(x,y)$ is equal zero for all x only if $y = 0$. Since the metric is non-degenerate and linear, it must be one-to-one and onto, and hence invertible. We can define the *inverse metric* as a $(2,0)$ tensor g^{ik} such that $g^{ik}g_{kj} = \delta^i{}_j$. The metric tensor and inverse metric tensor is also used to raise (e.g. $v^i = g^{ij}v_j$) and lower (e.g. $v_i = g_{ij}v^j$) the indices of arbitrary tensors. Once we have endowed a manifold with a metric, we can define an associated *scalar product* between two vectors on the tangent space as $x \cdot y = g(x,y) = g_{ik}x^iy^k$.

A *pseudo-Riemannian manifold* (M,g) is a differentiable manifold M equipped with a non-degenerate, smooth, symmetric metric tensor g which needs not be positive-definite (unlike a Riemannian manifold), but must be non-degenerate. The signature (p,q) of a pseudo-Riemannian metric is the number (counted with multiplicity) of positive (p), negative (q) and zero eigenvalues of the real symmetric matrix g_{ik}. A *Lorentzian manifold* is an important special case of a pseudo-Riemannian manifold in which the signature of the metric is $(1, n-1)$.

A basic principle of general relativity is that space-time can be modeled as a 4-dimensional Lorentzian manifold of signature $(3,1)$ or, equivalently, $(1,3)$, where the Riemann tensor g represents the curvature of the space-time due to the presence of matter (energy).

A *curve* γ is defined as a map from some interval $I \subset \mathbb{R}$ to the manifold

$$\begin{aligned} \gamma\colon I \subset \mathbb{R} &\to M, \\ t &\mapsto \gamma(t). \end{aligned} \tag{A.7}$$

Its tangent vector is $\dot{\gamma}(t)$.

The *covariant derivative* (or a *linearly connection*) ∇ on a manifold M maps a pair of vectors to a vector

$$\begin{aligned} \nabla\colon V \times V &\to V, \\ (x,y) &\mapsto \nabla_x y, \end{aligned} \tag{A.8}$$

which is linear and which satisfied the following conditions for all $f \in \mathcal{F}$:

- the covariant derivative of a function f is its differential: $\nabla_v f = v^i \nabla_i f = v(f) = \mathrm{d}f(v)$;
- $\nabla_{fx}y = f\nabla_x y$ and $\nabla_x(fy) = f\nabla_x y + y\mathrm{d}f(x)$.

Due to the linearity, it is completely specified by the covariant derivatives of the basis vectors

$$\begin{aligned} \nabla_{\delta_i}\delta_j &= \Gamma^k{}_{ij}\delta_k, \\ \nabla_x y &= x^k(\delta_k(y^i) + \Gamma^i{}_{kj}v^j)\delta_i. \end{aligned} \tag{A.9}$$

The functions $\Gamma^k{}_{ij}$ are called *connection coefficients* or *Christoffel symbols* (they are not tensors).

A vector v is said to be *parallel transported* along γ if $\nabla_{\dot{\gamma}}v = 0$. A *geodesic curve* is defined as a curve whose tangent vector is parallel transported along γ, i.e. $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$ (for geodesic equation see (2.10)).

The covariant derivative can also be extended for tensors. A covariant derivative on a manifold M is a map ∇ from smooth (n,m) tensor $T^{i_1,\dots,i_n}{}_{j_1,\dots j_m}$ to smooth $(n,m{+}1)$ tensor $\nabla_{j_{m+1}}T^{i_1,\dots,i_n}{}_{j_1,\dots j_m}$ that is linear and satisfies the following conditions:

- it obeys the Leibniz rule for derivatives;
- it commutes with contraction.

The *torsion* of a connection is defined by

$$\begin{aligned} T\colon V \times V &\to V, \\ (x,y) \mapsto T(x,y) &= \nabla_x y - \nabla_y x - [x,y], \end{aligned} \tag{A.10}$$

where $[x,y]$ is the Lie derivative. The torsion vanishes if and only if the connection is symmetric. It can be proven that the Einstein equivalence principle implies that $T=0$.

Let (M,g) be a pseudo-Riemannian manifold. Given two vectors u,w, we require that the scalar product $g(u,w)$ is unchanged if we deliver u and w along any curve, i.e. $\nabla_v g(u,w)=0$ for all v. The covariant derivative ∇ is called a *Levi-Civita connection* if it preserves the metric (i.e. $\nabla g=0$) and it is torsion-free (i.e. $T=0$). In this case, the Christoffel symbols are completely determined by the metric:

$$\begin{aligned} \Gamma^l{}_{ij} &= \Gamma^l{}_{ji}, \\ \Gamma^l{}_{ij} &= \frac{1}{2} g^{lk}\left(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}\right). \end{aligned} \tag{A.11}$$

The *curvature* is the amount by which a geometric object deviates from being flat. The curvature is described by the general *Riemann curvature tensor*, which measures the extent to which the metric tensor is not locally isometric to a Euclidean space. The Riemann tensor on M is the $(1,3)$ tensor $R^i{}_{jkl}$ such that for any cobasis dual field ω_j on M $R^i{}_{jkl}\omega_i = -[\nabla_k,\nabla_l]\omega_j$. The components of the Riemann tensor are given by

$$R^i{}_{jkl} = \partial_k\Gamma^i{}_{lj} - \partial_l\Gamma^i{}_{kj} + \Gamma^m{}_{lj}\Gamma^i{}_{km} - \Gamma^m{}_{kj}\Gamma^i{}_{lm}. \tag{A.12}$$

The Riemann tensor obeys three important symmetries

$$\begin{aligned} R_{ijkl} =& - R_{jikl} = -R_{ijlk}, \\ R_{ijkl} =& R_{klij}, \end{aligned} \tag{A.13}$$

which reduces its 256 components in four dimensions to 21. In addition, the *Bianchi identity* holds:

$$\begin{aligned} R_{i[jkl]} &= 0, \\ R^m{}_{l[ki;j]} &= 0. \end{aligned} \tag{A.14}$$

The first Bianchi identity reduces the number of independent components of the Riemann tensor to 20. The contraction of the Riemann tensor over its first and third indices is the *Ricci tensor*, thus its components are $R_{ik} = R^m{}_{imk} = R_{ik}$. A further contraction yields the *Ricci scalar* $R = R^i{}_i$.

The *Einstein tensor* is the combination

$$G_{ik} = R_{ik} - \frac{1}{2} R g_{ik}. \tag{A.15}$$

Contracting the second Bianchi identity, we find the contracted Bianchi identity $\nabla_i G^{ik} = 0$.

Energy-momentum tensor

The *energy-momentum tensor* (or *stress-energy tensor*) is a tensor T of type $(2,0)$ that describes the density and flux of energy and momentum in space-time. In general relativity, the stress-energy tensor is symmetric, i.e $T^{ik} = T^{ki}$ and it satisfies the *equation of continuity*

$$\nabla_i T^{ik} = 0. \tag{A.16}$$

The component T^{00} is the density of relativistic mass, i.e. the energy density divided by the speed of light squared. The component $T^{0\alpha}$ is the density of the αth component of linear momentum, i.e. it describes the flux of relativistic mass (i.e. the amount of energy) through unit surface perpendicular to the x^α axis in unit time. The components $T^{\alpha\beta}$ represent the *stress tensor* (denoted by $\sigma_{\alpha\beta}$), i.e. the flux of αth component of linear momentum passing per unit time through unit surface perpendicular to the x^β axis. In particular, $T^{\alpha\alpha}$(not summed) represents normal stress, which is called *pressure* when it is independent of direction, while the remaining components $T^{\alpha\beta}$ with $\alpha \neq \beta$ represent *shear stress.*

An example is the stress-energy tensor of a system of particles. We describe their mass distribution in the space using a "mass density" in the form:

$$\mu = \sum_a m_a \delta(\mathbf{r} - \mathbf{r}_a), \tag{A.17}$$

where $\mathbf{r}_a$ is the radius-vector of the particles, m_a is the mass of the particle a, δ is the Dirac delta function and the summation extends over all the particles of the system.
It can be proven that the energy momentum tensor of the system of non-interacting particles is

$$T^{ik} = \mu c \frac{dx^i}{ds}\frac{dx^k}{dt} = \mu c u^i u^k \frac{ds}{dt}, \tag{A.18}$$

which can be rewritten as

$$T^{ik} = \sum_a m_a c u^i u^k \delta(\mathbf{r} - \mathbf{r}_a). \tag{A.19}$$

In general relativity the stress-energy tensor serves a role similar to that of mass distribution in Newtonian physics; it tells space how to deform, creating what we observe as gravity.

Appendix B

Hansen coefficients

We have seen in section 5.3.1 that the development of $\left(\frac{r}{a}\right)^n \theta^m$ is

$$\left(\frac{r}{a}\right)^n \theta^m = \sum_{k=-\infty}^{+\infty} X_k^{n,m}(e) \exp(\sqrt{-1}(k-m)M) \tag{B.1}$$

where $X_k^{n,m}(e)$ are the Hansen coefficients.

These coefficients are defined in the following way:

- if $k \neq 0$

$$X_k^{n,m}(e) = \beta^{|m-k|}(1+\beta^2)^{-n-1} \sum_{s=0}^{\infty} P^{(n)}_{s+\max(0,k-m)}(m,\kappa) P^{(n)}_{s+\max(0,k-m)}(-m,-\kappa)\beta^{2s} \tag{B.2}$$

  where

$$\beta = \frac{e}{1+\sqrt{1-e^2}}, \qquad \kappa = \frac{k}{1+\beta^2},$$

$$P_s^{(n)}(m,\kappa) = \sum_{r=0}^{r^\star} \frac{(-n+m-1)_r}{(1)_r} \frac{\kappa^{s-r}}{(1)_{s-r}}$$

$$(a)_0 = 1, \qquad (a)_k = a(a+1)...(a+k-1) = (a+k-1)(a)_{k-1}$$

$$r^\star = \begin{cases} \min\{s, n-m+1\} & \text{if } n-m+1 \geq 0 \\ s & \text{if } n-m+1 < 0 \end{cases}$$

- if $k = 0$

  - if $n \geq -1$

$$X_0^{n,m}(e) = (-1)^{|m|} \frac{(n+2)_{|m|}}{(1)_{|m|}} \beta^{|m|}(1+\beta^2)^{-n-1} F(-n-1, -n+|m|-1, 1+|m|; \beta^2) \tag{B.3}$$

  - if $n < -1$

$$\begin{aligned} X_0^{n,m}(e) = (-1)^{|m|} \frac{(n+2)_{|m|}}{(1)_{|m|}} \left(\frac{e}{2}\right)^{|m|} (1-e^2)^{\frac{n+3}{2}} \times \\ \times F\left(\frac{n+|m|+2}{2}, \frac{n+|m|+3}{2}, 1+|m|; e^2\right), \end{aligned} \tag{B.4}$$

where F is the hypergeometric function defined as

$$F(a,b,c;x) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{x^k}{k!}, \tag{B.5}$$

which converges if c is not a negative integer for all $|x| < 1$. Here, $(a)_k$ is the Pochhammer symbol, which is defined by

$$(a)_0 = 1, \qquad (a)_k = (a+k-1)(a)_{k-1}.$$

Expansion by Means of Lie Transform

Because the relation between the mean anomaly and the true anomaly is not trivial, we present another algorithm by J. Henrard which can be used to transform an expansion in the true anomaly, into one in the mean anomaly.

The idea is to transform an analytic function

$$F(\nu, e) = \sum_{i=0}^{\infty} \frac{F_i^0(\nu)}{i!} e^i \tag{B.6}$$

into $G(M,e) = F(\nu(M,e), e)$. The algorithm by Henrard uses the Lie transforms method, and the generator is

$$W(\nu, e) = \frac{\partial \nu}{\partial e} = \frac{4\sin\nu + e\sin(2\nu)}{2(1-e^2)} = 2\sin(\nu)\sum_{i=0}^{\infty} e^{2i} + \frac{1}{2}\sin(2\nu)\sum_{i=0}^{\infty} e^{2i+1} = \sum_{i=0}^{\infty} \frac{W_i(\nu)}{i!} e^i. \tag{B.7}$$

The transformed function

$$G(M, e) = \sum_{i=0}^{\infty} \left. \frac{F_0^i(\nu)}{i!} \right|_{\nu=M} e^i \tag{B.8}$$

is computed by the recursive formula

$$F_{i-j}^{j}(\nu) = F_{i+1-j}^{j-1}(\nu) + \sum_{k=0}^{i-j} \binom{i-j}{k} \frac{\partial F_{i-j-k}^{j-1}}{\partial \nu} W_k(\nu). \tag{B.9}$$

For example, this algorithm can be used to transform $r = r(\nu, e)$ into $r = r(M, e)$.

Bibliography

[1] Adams F.C. & Laughlin G.: *Effects of secular interactions in extrasolar planetary systems.* Astrophys. J, **649**, 992-1003 (2006).

[2] Arnold V.I., Kozlov V.V. and Neishtadt A.I.: *Mathematical aspects of classical and celestial mechanics*, Dyn. Systems III, V.I. Arnold, ed. Springer, Berlin (1988).

[3] Arnold V.I.: *Mathematical Methods of Classical Mechanics*, Second Edition, Springer, Graduate Texts in Mathematics, 60 (1989).

[4] Beaugé C., Nesvorńy D., Dones L.: *A high-order analytical model for the secular dynamics of irregular satellites*, AJ, **131**, 2299-2313 (2006).

[5] Benettin G.: *The elements of Hamiltonian perturbation theory*, Course notes, Department of Mathematics, University of Padova (Italy).

[6] Benitez F., Gallardo T.: *The Relativistic Factor in the Orbital Dynamics of Point Masses*, Celest Mech Dyn Astr, **101**, 289-307 (2008).

[7] Birkhoff G.D.: *Dynamical systems*, AMS Ccolloquium Publications, **IX** (1927).

[8] Biscani F. and Carloni S.: *A first-order secular theory for the post-Newtonian two-body problem with spin*, MNRAS, **000**, 1-17 (2013).

[9] Brouwer D., Clemence G. M.: *Methods of Celestial Mechanics.*, Academic Press, New York (1961).

[10] Brumberg V.: *Essential Relativistic Celestial Mechanics.* Adam Hilger, London (1991).

[11] Brumberg V.: *On derivation of EIH (Einstein-Infeld-Hoffman) equations of motion from the linearized metric of general relativity theory.* Celestial Mechanics, **99**, 245-252 (2008).

[12] Deprit A.: *Canonical transformations depending on a small parameter*, Celestial Mechanics, **1**, 12-30 (1969).

[13] Duriez L.: *Cours de Mécanique céleste classique*, Courses notes, Astronomy Laboratory of the University of Lille 1 and IMCCE of the Paris Observatory.

[14] Duriez L.: *Le Problème des Deux Corps revisité*, Les Méthodes modernes de la Mécanique Céleste, Editions Frontiéres, 9-34 (1989).

[15] Duriez L.: *Le Développment de la Fonction Perturbatrice*, Les Méthodes modernes de la Mécanique Céleste, Editions Frontiéres, 36-61 (1989).

[16] Einsten A.: *Explanation of the perihelion motion of Mercury from the general theory of relativity*, Translated from the "Königlich Preußische Akademie der Wissenschaften", **354**, 831-839 (1915).

[17] Einstein A.: *The foundation of the general theory of relativity*, Translated from the "Annalen der Physik", **354**, 769-822 (1916).

[18] Einstein A., Infeld L., Hoffmann B.: *The Gravitational Equations and the Problem of Motion*, The Annals of Mathematics, **39**, 65-100 (1938).

[19] Einstein A. and Infeld L.: *On the motion of particles in general relativity theory*, Canadian Journal of Mathematics, **I**, 209-241 (1949).

[20] Giorgilli A.: *Hamiltonian systems*, Course notes, Department of Mathematics, University of Milano (Italy).

[21] Giorgilli A.: *Meccanica celeste*, Course notes, Department of Mathematics, University of Milano (Italy).

[22] Giorgilli A., Delshams A., Fontich E., Galgani L., and Simó L.: *Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem.* J. Diff. Eqs., **77**, 167-198 (1989).

[23] Giorgilli A. and Locatelli U.: *Introduction to canonical perturbation theory for nearly integrable systems*, Chaotic worlds. In: Proceedings of the Nato Advanced Study Institute, (2003).

[24] Giorgilli A., Locatelli U. and Sansottera M.: *Kolmogorov and Nekhoroshev theory for the problem of three bodies*, CeMDA, **104**, 159-173 (2009).

[25] Giorgilli A. and Sansottera M.: *Methods of algebraic manipulation in perturbation theory*, Workshop Series of the Asociacion Argentina de Astronomia, **3**, 147-183 (2011).

[26] Heimberger J., Soffel M., Ruder H.: *Relativistic effects in the motion of artificial satellites - The oblateness of the central body II*, Celestial Mechanics and Dynamical Astronomy, **47**, 205 (1990).

[27] Iorio L.: *On the possibility of measuring the solar oblateness and some relativistic effects from planetary ranging.* A&A **433**, 385-393 (2005).

[28] Landau L. D. and Lifshitz E.M.: *Mechanics - Volume 1 of Course of Theoretical Physics*, Second Edition, Pergamon Press (1969).

[29] Landau L. D. and Lifshitz E.M.: *The Classical Theory of Fields - Volume 2 of Course of Theoretical Physics*, Third English Edition, Pergamon Press (1971).

[30] Laplace P.S.: *A Treatise in Celestial Mechanics*, Volume IV, Book X, Chapter VII (1805), reprinted by Chelsea, New York (1966).

[31] Laskar, J.: *Secular evolution of the solar system over 10 million years.*, A&A **198**, 341-362 (1988).

[32] Laskar J.: *Systèmes de variables et éléments*, Les Méthodes modernes de la Mécanique Céleste, Editions Frontiéres, 63-87 (1989).

[33] Laskar J.: *Manipulation des Séries*, Les Méthodes modernes de la Mécanique Céleste, Editions Frontiéres, 90-106 (1989).

[34] Laskar J. and Roboutel P.: *Stability of the planetary three body-problem - I. Expansion of the Planetary Hamiltonian*, CeMDA, **62**, 193L (1995).

[35] Laskar J.: *Large Scale Chaos and Marginal Stability in the Solar System.*, CM&DA **64**, 115-162 (1996).

[36] Laskar J.: *Chaotic diffusion in the Solar System.*, Icarus, **196**, 1-15 (2008).

[37] Lense J. and Thirring H.: *The Influence of the Self-Rotation of Central Bodies on the Movements of the Planets and Moon According to Einstein's Theory of Gravitation*, Translated from the "Physikalische Zeitschrift", **19**, 156-163 (1918).

[38] Libert A.-S. and Henrard J.: *Analytical approach to the secular behaviour of exoplanetary systems*, CeMDA, **93**, 187-200 (2005).

[39] Libert A.-S. and Henrard J.: *Secular apsidal configuration of non-resonant exoplanetary systems*, Icarus, **183**, 186-192 (2006).

[40] Libert A.-S. and Henrard J.: *Analytical study of the proximity of exoplanetary systems to mean-motion resonances*, A&A, **461**, 759-763 (2007).

[41] Libert A.-S. and Sansottera M.: *On the extension of the Laplace-Lagrange secular theory to order two in the masses for extrasolar systems*, DOI: 10.1007/s10569-013-9501-z (2013).

[42] Locatelli U. and Giorgilli A.: *Construction of Kolmogorov's normal form for a planetary system*, Regular and Chaotic Dynamics 10, 153-171 (2005) .

[43] Migaszewski C. and Goździewski K.: *Secular dynamics of coplanar, non-resonant planetary system under the general relativity and quadrupole moment perturbations*, MNRAS, **000**, 0-17 (2008).

[44] Migaszewski C. and Goździewski K.: *A secular theory of coplanar, non-resonant planetary system*, MNRAS, **388**, 789-802 (2008a).

[45] Morbidelli A.: *Modern Celestial Mechanics - Aspects of Solar System Dynamics*, (2011).

[46] Moser J.: *Lectures on Hamiltonian systems*, Mem. Amer. Math. Soc. **81** (1968).

[47] Poincaré H.: *Leçons de Mécanique Céleste, tomes I-II*, Gauthier-Villars (1905).

[48] Richardson D.L. and Kelly T.J.: *Two-body motion in the post-Newtonian approximation*, Celestial Mechanics, **43**, 193 (1988).

[49] Roboutel P.: *Stability of the Planetary Three-Body Problem - II. KAM Theory and Existence of Quasiperiodic Motions*, CeMDA, **62**, 219-261 (1995).

[50] Sansottera M., Locatelli U., Giorgilli A.: *A Semi-Analytic Algorithm for Constructing Lower Dimensional Elliptic Tori in Planetary Systems*, CeMDA, **111**, 337-361 (2011).

[51] Sansottera M., Locatelli U. and Giorgilli A.: *On the stability of the secular evolution of the planar Sun-Jupiter-Saturn-Uranus system*, Mathematics and Computers in Simulation, Volume 88, 1-14 (2013).

[52] Siegel C.L., Moser J.K.: *Lectures on Celestial Mechanics*, Springer, Berlin(1971).

[53] Sitarski G.: *Effects of general relativity in the motion of minor planets and comets.* Acta Astronomia **33**, 295-304 (1983).

[54] Tisserand F.: *Traité de mécanique céleste*, Gauthier-Villars, (1888).

[55] Veras D. and Armitage P.J.: *Extrasolar Planetary Dynamics with a Generalized Planar Laplace-Lagrange Secular Theory*, ApJ, **661**, 1311-1322 (2007).

[56] Will C.M.: *Theory and Experiment in Gravitational Physics Revised Edition*, Cambridge University Press (1993).

[57] Whittaker E.T., Watson G.N.: *A course of modern analysis*, fourth edition(1927), reprinted by Cambridge University Press (1992).