An elementary introduction to celestial mechanics

1. Lagrangian and Hamiltonian equations.

In this note we review, starting from elementary notions in classical mechanics, several problems in celestial mechanics, showing how they may be solved at first order in perturbation theory, obtaining quantitative results already in reasonably good agreement with the observed data. In this and in the following sections 2-7 we summarize very briefly a few prerequisites.
We suppose the reader familiar with the notion of *ideal constraint*, and with the description of the mechanical problems by means of generalized coordinates, using the classical *Lagrange equations*

1.1 $$\frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \tag{1.1}$$

where

1.2 $$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) = T(\mathbf{q}, \dot{\mathbf{q}}) - V(\mathbf{q}, t) \tag{1.2}$$

It is useful to allow the t–dependence because in several problems such dependence appears explicitly (and periodic in time). The equivalent *Hamilton equations* are

1.3 $$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} \qquad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}} \tag{1.3}$$

where, by definition

1.4 $$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \tag{1.4}$$

and

1.5 $$H(\mathbf{p}, \mathbf{q}) = T(\mathbf{p}, \mathbf{q}) + V(\mathbf{q}, t) \tag{1.5}$$

The notion of *canonical transformation* $\mathbf{q}, \mathbf{p} \longleftrightarrow \mathbf{Q}, \mathbf{P}$ is also needed. A canonical transformation can be characterized by the property

1.6 $$\mathbf{p} \cdot d\mathbf{q} + \mathbf{Q} \cdot d\mathbf{P} = dF \tag{1.6}$$

which is implied by the existence of a "generating function" $F(\mathbf{q}, \mathbf{P}, t)$ such that

1.7 $$\mathbf{p} = \frac{\partial F(\mathbf{q}, \mathbf{P}, t)}{\partial \mathbf{q}}, \qquad \mathbf{Q} = \frac{\partial F(\mathbf{q}, \mathbf{P}, t)}{\partial \mathbf{P}}, \qquad H'(\mathbf{Q}, \mathbf{P}, t) = H(\mathbf{q}, \mathbf{p}) + \frac{\partial F(\mathbf{q}, \mathbf{P}, t)}{\partial t} \tag{1.7}$$

which is a sufficient condition for the canonicity of the map.

2. Central motion.

One of the most well known integrable systems is the *two body problem.* The system consists of two point masses with masses m_1 and m_2, respectively, interacting through a conservative force with potential energy depending only on their distance

2.1 $$V(\mathbf{x}_1, \mathbf{x}_2) = V(|\mathbf{x}_1 - \mathbf{x}_2|) \tag{2.1}$$

and we shall assume that $V(\rho)$ is defined for $\rho > 0$ and that it is such that

2.2 $$\lim_{\rho \to 0} \rho^2 V(\rho) = 0 \tag{2.2}$$

2.3 $$\inf_{\rho \geq \varepsilon} V(\rho) = -V_\varepsilon > -\infty \qquad \forall \varepsilon > 0 \tag{2.3}$$

The forces associated with the potential (2.1) satisfy the third law of dynamics: hence the center of mass undergoes a uniform rectilinear motion, and it may be supposed at rest in the origin of a suitable inertial frame $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$. To determine the position of the two points it will be enough to give the vector $\underline{\rho}(t) = \mathbf{x}_1(t) - \mathbf{x}_2(t)$ because it will be

2.4 $$\mathbf{x}_1(t) = \frac{m_2}{m_1 + m_2}\,\underline{\rho}(t), \qquad \mathbf{x}_2(t) = \frac{-m_1}{m_1 + m_2}\,\underline{\rho}(t) \tag{2.4}$$

Moreover, as a consequence of the third law of dynamics, the angular momentum $\mathbf{K}$ will be conserved, and it may be assumed to be parallel to the unit vector $\mathbf{i}_3$

2.5 $$\mathbf{K} = A\mathbf{i}_3 \tag{2.5}$$

Using (2.4), (2.5) and the definition of $\underline{\rho}(t)$ we obtain

2.6 $$\mathbf{K} = A\mathbf{i}_3 = m_1\mathbf{x}_1 \wedge \dot{\mathbf{x}}_1 + m_2\mathbf{x}_2 \wedge \dot{\mathbf{x}}_2 = \frac{m_1 m_2}{m_1 + m_2}\,\underline{\rho} \wedge \underline{\dot{\rho}} \tag{2.6}$$

Hence $\underline{\rho}$ and $\underline{\dot{\rho}}$ have to lie in the plane $(\mathbf{i}_1, \mathbf{i}_2)$, and the motion may be parameterized by the polar coordinates (ρ, ϑ) in such plane. The resulting Lagrangian is

2.7 $$\mathcal{L}(\dot{\rho}, \dot{\vartheta}, \rho, \vartheta) = \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}(\dot{\rho}^2 + \rho^2\dot{\vartheta}^2) - V(\rho) \tag{2.7}$$

which is the Lagrangian for a single point mass P of mass $m = \frac{m_1 m_2}{m_1 + m_2}$ moving on the plane and attracted by a force directed to the origin O and with potential energy depending only on $|P - O|$. The motion of such a system is called *central motion.*
The Lagrangian (2.7) does not depend explicitly on time nor on the coordinate ϑ. The conserved quantities so obtained are obviously the energy

2.8 $$E = \frac{1}{2}m(\dot{\rho}^2 + \rho^2\dot{\vartheta}^2) + V(\rho) \tag{2.8}$$

and the angular momentum along the $\mathbf{i}_3$ axis

2.9 $$A = \rho^2\dot{\vartheta} \tag{2.9}$$

We shall suppose $A \neq 0$, and this implies that

2.10 $$\rho(t) \geq \rho_0 \quad \forall t \in \mathbf{R} \tag{2.10}$$

To check (2.10) we put (2.9) in (2.8) obtaining

2.11 $$E = \frac{1}{2}m(\dot{\rho}^2 + \frac{A^2}{\rho^2}) + V(\rho) \tag{2.11}$$

Given E the assumption (2.2) implies the existence of ρ_0 such that for $\rho \leq \rho_0$

2.12 $$E - \frac{1}{2}m(\dot{\rho}^2 + \frac{A^2}{\rho^2}) - V(\rho) < 0 \tag{2.12}$$

from which (2.10) follows.
The energy conservation relation (2.11) shows that the motion of ρ is a one dimensional motion with potential energy

2.13 $$V_A(\rho) = \frac{mA^2}{2\rho^2} + V(\rho) \tag{2.13}$$

which, therefore, can be integrated by the *quadrature*

2.14
$$t - t_0 = \int_{\rho_0}^{\rho} \frac{d\rho'}{\sqrt{\frac{2}{m}(E - V(\rho')) - \frac{A^2}{\rho'^2}}} \tag{2.14}$$

Combining (2.14) with (2.9) we find

2.15
$$d\vartheta = \frac{A}{\rho^2}dt = \frac{A\rho^{-2}d\rho}{\sqrt{\frac{2}{m}(E - V(\rho)) - \frac{A^2}{\rho^2}}} \;\to\; \vartheta - \vartheta_0 = \int_{\rho_0}^{\rho} \frac{A{\rho'}^{-2}d\rho'}{\sqrt{\frac{2}{m}(E - V(\rho')) - \frac{A^2}{\rho'^2}}} \tag{2.15}$$

The integrals (2.14) and (2.15) give the equations of motions. In particular one can show that the motion can be represented, with a suitable transformation of variables, by two point masses rotating on two unit circles at constant speeds, depending on E and A: this type of motions is called *quasiperiodic* with two periods. See Section 9 below.

3. Kepler's laws

For the purpose of our applications to celestial mechanics we are particularly interested in the motion described by the integrals (2.14), (2.15) when the potential energy is gravitational, *i.e.*

3.1
$$V(\rho) = -m\frac{g}{\rho} \tag{3.1}$$

where g is a constant essentially equal to the product of the gravitational constant k times the mass of a celestial body (*e.g.* the Sun or the Earth or the Moon).

This choice gives for the effective potential $V_A(\rho)$ a shape shown in Fig 3.1

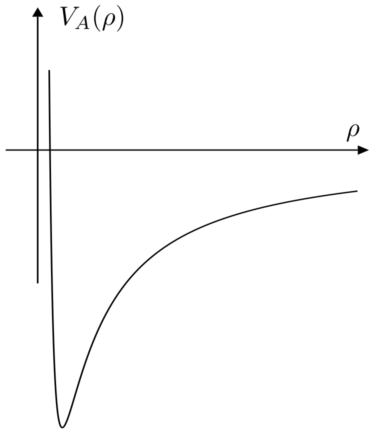

Fig. 3.1: The effective gravitational potential. The minimum of V_A is $-mg/2A^2$.

and we shall study in particular the case $E < 0$ so that it is possible to find ρ_- and ρ_+ such that $\rho_- \le \rho(t) \le \rho_+$. This means that we are imposing $0 > E > -mg/2A^2$.

Under the above conditions the problem is known as Kepler's problem, and it is governed by Kepler's laws of motions

(a) The trajectories of the motion are ellipses with focus in O.
(b) The motion on the ellipses has constant areal velocity around the focus O.
(c) The ratio between the square of the revolution period T and the cube of the length of the ellipse's major axis is a constant depending solely on g.

Moreover the focal distances ρ_- and ρ_+ are such that

3.2
$$\rho_+ + \rho_- = mg/(-E) \tag{3.2}$$

3.3
$$\rho_+\rho_- = \frac{mA^2}{-2E} \tag{3.3}$$

and the period of revolution T is

3.4
$$T = \frac{\pi}{\sqrt{2g}}(\rho_+ + \rho_-)^{3/2} \tag{3.4}$$

To check Kepler's laws we write (2.11) as

3.5
$$\frac{1}{2}(\dot\rho^2 + \frac{A^2}{\rho^2}) - \frac{g}{\rho} = \frac{E}{m} \tag{3.5}$$

The value of $\frac{1}{\rho_-}$ and$\frac{1}{\rho_+}$ are therefore the solutions of the following polynomial equation in $\frac{1}{\rho}$

3.6
$$\frac{E}{m} - \frac{A^2}{2\rho^2} + \frac{g}{\rho} = \frac{A^2}{2}\Big(\frac{1}{\rho_-} - \frac{1}{\rho}\Big)\Big(\frac{1}{\rho} - \frac{1}{\rho_+}\Big) = 0 \tag{3.6}$$

This gives immediately (3.2) and (3.3), since

3.7
$$\frac{1}{\rho_+} + \frac{1}{\rho_-} = \frac{2g}{A^2} \qquad \frac{1}{\rho_+}\frac{1}{\rho_-} = \frac{-2E}{mA^2} \tag{3.7}$$

Starting from (3.6) and setting $\vartheta(\rho_-) = \pi$ we can study the motion between ρ_- and ρ_+ by writing (2.15) as

3.8
$$\vartheta - \pi = \int_{\rho_-}^{\rho} \frac{d\rho'}{\rho'^2\sqrt{\Big(\frac{1}{\rho_-} - \frac{1}{\rho'}\Big)\Big(\frac{1}{\rho'} - \frac{1}{\rho_+}\Big)}} \tag{3.8}$$

This is an elementary integral: performing the change of variable $y = 1/\rho$ one finds

3.9
$$\frac{1}{\rho} = \frac{1}{2}\Big[\Big(\frac{1}{\rho_+} + \frac{1}{\rho_-}\Big) + \Big(\frac{1}{\rho_-} - \frac{1}{\rho_+}\Big)\cos(\vartheta - \pi)\Big] \tag{3.9}$$

where the origin of the angle ϑ is chosen so that $\vartheta = 0$ corresponds to the point furthest away from the origin, $\rho = \rho_+$.

The motion between ρ_+ and ρ_- can be studied likewise, and it still verifies (3.9). From elementary geometry it is well known that (3.9) is the polar equation of an ellipse with focus at the origin, focal distances ρ_+ and ρ_-, major axis along the x axis, semiaxes a, b given by $a = \frac{\rho_+ + \rho_-}{2}$ and $b = \sqrt{\rho_+\rho_-}$, and eccentricity $e = \frac{\rho_+ - \rho_-}{\rho_+ + \rho_-}$.

This proves the Kepler's law (a). To prove (b) one simply remarks that the areal velocity, *i.e.* the area swept by the radius ρ per unit time, is $\frac{1}{2}\rho^2\dot\vartheta = \frac{A}{2}$, and it is therefore constant for any central motion.

The Kepler's law c) is a consequence of (3.4), and to prove the latter one can compute explicitly the period using (2.14) or, more simply, dividing the area of the ellipse by the areal velocity. In this way one obtains

3.10
$$T = \pi\frac{\rho_+ + \rho_-}{2}\sqrt{\rho_+\rho_-}\frac{2}{A} = \frac{\pi}{\sqrt{2g}}(\rho_+ + \rho_-)^{3/2} = \frac{2\pi a^{3/2}}{\sqrt{g}} \tag{3.10}$$

where we have used (3.2) and (3.3). By using the same equalities and the above expression of the eccentricity e we obtain also

3.11
$$A = \sqrt{ga(1 - e^2)} \tag{3.11}$$

4. Action-angle variables for the two body problem

The action–angle variables are canonical coordinates which are particularly adapted to study problems which are close to integrable ones: in our case we shall employ them to study various precession problems in systems that are close to either the two body problem or to the rigid body with a fixed point.
We introduce here the action-angle variables for the two body problem. The Hamiltonian of the system (see (1.4), (2.8)) is

4.1
$$H(p_\rho, p_\vartheta, \rho, \vartheta) = \frac{p_\rho^2}{2} + \frac{p_\vartheta^2}{2\rho^2} - \frac{g}{\rho} \tag{4.1}$$

where the overall factor $1/m$ in the original Lagrangian has been set equal to 1.
To introduce the action-angle variables, we write first the following expressions for the action variables

4.2
$$L = \frac{1}{2\pi}\oint p_\rho d\rho \qquad G = \frac{1}{2\pi}\oint p_\vartheta d\vartheta \tag{4.2}$$

where the integrals are computed on the solutions, in the sense that

4.3
$$p_\rho = \sqrt{2E - \tfrac{p_\vartheta^2}{\rho^2} + \tfrac{2g}{\rho}} \qquad p_\vartheta = \text{const} \tag{4.3}$$

Performing explicitly the integrals we find

4.4
$$L = \frac{1}{2\pi}\oint \sqrt{2E - \tfrac{p_\vartheta^2}{\rho^2} + \tfrac{2g}{\rho}}\,d\rho = \frac{g}{\sqrt{-2E}} \qquad G = p_\vartheta \tag{4.4}$$

The corresponding conjugate "angle variables" can be found by using the generating function S

4.5
$$S_(\rho, \vartheta, L, G) = S_\rho(\rho, L) + S_\vartheta(\vartheta, G) \tag{4.5}$$

where

4.6
$$S_\rho = \int p_\rho(L,G)d\rho \qquad S_\vartheta = \int p_\vartheta(L,G)d\vartheta \tag{4.6}$$

and

4.7
$$p_\rho = \sqrt{\tfrac{-g^2}{L^2} - \tfrac{G^2}{\rho^2} + \tfrac{2g}{\rho}} \qquad p_\vartheta = G \tag{4.7}$$

The angle variable are obtained by differentiation of S. It is convenient to define the integrals in (4.6) in such a way that the integration constant does not depend on L, G. Integrating locally (4.6) (e.g. for motions with increasing ρ) we get

4.8
$$\begin{aligned}\lambda &= \frac{\partial S}{\partial L} = \frac{\partial S_\rho}{\partial L} = \frac{\partial}{\partial L}\int_{\rho_-}^{\rho} \sqrt{\tfrac{-g^2}{L^2} - \tfrac{G^2}{r^2} + \tfrac{2g}{r}}\,dr = \\ &= \frac{g^2}{L^3}\int_{\rho_-}^{\rho} \frac{1}{\sqrt{\tfrac{-g^2}{L^2} - \tfrac{G^2}{r^2} + \tfrac{2g}{r}}}dr = \frac{g^2}{L^3}(t - t_-);\end{aligned} \tag{4.8}$$

4.9
$$\begin{aligned}\gamma &= \frac{\partial S}{\partial G} = \frac{\partial}{\partial G}\Big[\int_{\rho_-}^{\rho} \sqrt{\tfrac{-g^2}{L^2} - \tfrac{G^2}{r^2} + \tfrac{2g}{r}}\,dr + \int_0^{\vartheta} G d\vartheta'\Big] = \\ &= \int_{\rho_-}^{\rho} \frac{G}{r^2\sqrt{\tfrac{-g^2}{L^2} + 2\tfrac{g}{r} - \tfrac{G^2}{r^2}}}\,dr + \vartheta = \vartheta_-\end{aligned} \tag{4.9}$$

where we have used (2.14)-(2.15) and t_- is the instant in which the system has $\rho = \rho_-$ and ϑ_- is the angle between the major semiaxis of the Keplerian ellipse and a fixed "nodal" axis from which the angle ϑ is computed on the plane of the motion.
The set of canonical variables (L, G, λ, γ) have the following properties:

4.10 $$L = \frac{g}{\sqrt{-2E}} = \sqrt{ga} \tag{4.10}$$

where a is the major semiaxis of the orbit and, in the last equality, we have used (3.2). From (4.10) and (4.8) we obtain

4.11 $$\lambda - \lambda_- = t\frac{\sqrt{g}}{a^{3/2}} = 2\pi\frac{t}{T} \tag{4.11}$$

as it has to be since (L, λ) are conjugates action-angle variables. The angle λ is usually called *average anomaly*. Moreover the conjugates action-angle variables (G, γ) are such that

4.12 $$G = A = L(1 - e^2)^{1/2} \tag{4.12}$$

where we have used (3.11), and γ, which is constant over the solution of motion, is the angle between the major axis of the ellipse and a nodal axis.
The canonical variables (L, G, λ, γ) are called *Delaunay variables*. In terms of such variables the Hamiltonian of the two bodies problem becomes simply

4.13 $$H = -\frac{g^2}{2L^2} \tag{4.13}$$

5. Spherical trigonometry

We collect here a few classical spherical trigonometry results needed in the following sections. Calling A, B, C the three sides and α, β, γ the three angles of the spherical triangle in Fig. 5.1:

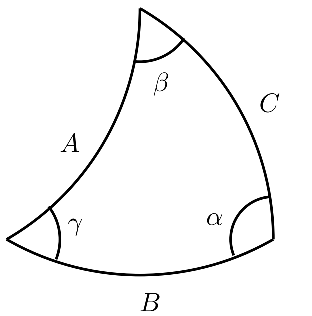

Fig. 5.1: Spherical triangle with the sides formed by the arcs A, B, C opposite to the angles α, β, γ.

the following are the key relations of spherical trigonometry:

5.1 $$\cos C = \cos A \cos B + \sin A \sin B \cos\gamma \tag{5.1}$$

5.2 $$\cos\gamma = -\cos\alpha\cos\beta + \sin\alpha\sin\beta\cos C \tag{5.2}$$

5.3 $$\frac{\sin\alpha}{\sin A} = \frac{\sin\beta}{\sin B} = \frac{\sin\gamma}{\sin C} \tag{5.3}$$

5.4 $$\sin C\cos\beta = \cos B\sin A - \sin B\cos A\cos\gamma \tag{5.4}$$

5.5 $$\cos A \cos\gamma = \sin A \cot B - \sin\gamma \cot\beta \tag{5.5}$$

5.6 $$dA = \cos\beta dC + \cos\gamma dB + \sin B \sin\gamma d\alpha \tag{5.6}$$

To check the spherical identities (5.1)-(5.6) simply draw the spherical triangle in Fig. 5.1 by locating the vertex 2 with the angle γ on the z axis, the vertex 1 with the β angle on the xz plane: so that the three vertices are expressed in Cartesian coordinates as $\mathbf{r}_1 = (\sin A, 0, \cos A)$, $\mathbf{r}_2 = (0, 0, 1)$ and $\mathbf{r}_3 = (\sin B \cos\gamma, \sin B \sin\gamma, \cos B)$. Then

(i) to check (5.1) observe that $\mathbf{r}_1 \cdot \mathbf{r}_3 = \cos C$;

(ii) to check (5.2) apply (5.1) to the spherical triangle formed on the sphere by the perpendicular to the planes containing the arcs A, B, C;

(iii) to check (5.3) note that $\mathbf{r}_1 \cdot \mathbf{r}_2 \wedge \mathbf{r}_3 = \sin A \sin B \sin\gamma$ has to be symmetric in the interchange of the role of $(A, \alpha), (B, \beta), (C, \gamma)$;

(iv) to check (5.4) remark that $\mathbf{r}_1 \wedge \mathbf{r}_3 \cdot \underline{j} = -\sin C \cos\beta$;

(v) the identity (5.5) is a consequence of (5.1) and (5.4);

(vi) finally (5.6) is obtained by differentiating the expression of $\cos A$ obtained from (5.1) with the substitution $(A, \alpha) \to (C, \gamma)$ and then using (5.3) and (5.5)

6. Kinematic description of the rigid body and action angle coordinates.

A *rigid body* is a system of n material points $P_1, ..., P_n$ with masses $m_1, ..., m_n$ such that there exists a reference frame $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ with origin O and coordinate unit vectors $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ in which the coordinates of $P_1, ..., P_n$ are constant. This implies that internal constraints are present in the system and we assume that the constraints are ideal.
The motion of a rigid system is therefore identified with that of the frame $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$, called *comoving frame*, with respect to a frame $(\Omega; \bar{\mathbf{i}}, \bar{\mathbf{j}}, \bar{\mathbf{k}})$, called *fixed frame*.
The comoving frame can be chosen conveniently. In what follows we shall choose $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ in such a way that the coordinates (x_i, y_i, z_i), $i = 1, ..., n$ of the vectors OP_i representing the points verify the relations

6.1 $$\sum_i m_i x_i y_i = \sum_i m_i x_i z_i = \sum_i m_i y_i z_i = 0 \tag{6.1}$$

This is always possible, since the matrix $I_{\alpha\beta} = \sum_i m_i[(OP_i)^2\delta_{\alpha\beta} - (OP_i)_\alpha(OP_i)_\beta]$, whose non diagonal elements are exactly the sums in (6.1), is a symmetric positive definite matrix, and therefore can be diagonalized by a suitable rotation of the axes of the comoving frame.
In order to describe the rigid motion we start from the fundamental relation

6.2 $$\mathbf{v}_P = \mathbf{v}_O + \omega \wedge OP \tag{6.2}$$

where ω is the *angular velocity* of the frame $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ with respect to $(\Omega; \bar{\mathbf{i}}, \bar{\mathbf{j}}, \bar{\mathbf{k}})$.
By means of (6.2) we can derive the following results:
(i) Calling $\mathbf{K}_O$ the angular momentum of the rigid body with respect to O, and setting $I_1 = \sum_i m_i(y_i^2 + z_i^2)$, $I_2 = \sum_i m_i(x_i^2 + z_i^2)$, $I_3 = \sum_i m_i(x_i^2 + y_i^2)$, we have the following equality

6.3 $$\mathbf{K}_O = I_1\omega_1\mathbf{i}_1 + I_2\omega_2\mathbf{i}_2 + I_3\omega_3\mathbf{i}_3 \tag{6.3}$$

(6.3) can be proved simply inserting (6.2) in the definition of $\mathbf{K}_O$

6.4
$$\mathbf{K}_O = \sum_{i=1}^{n} OP_i \wedge m_i \mathbf{v}_{P_i} = \sum_{i=1}^{n} m_i OP_i \wedge (\omega \wedge OP_i) \tag{6.4}$$

then, by the well known relation $a \wedge (b \wedge c) = (a \cdot c)b - (a \cdot b)c$ we obtain

6.5
$$\mathbf{K}_O = \sum_{i=1}^{n} m_i[(OP_i)^2\omega - (OP_i \cdot \omega)OP_i] \tag{6.5}$$

and writing (6.5) in components and exploiting (6.1) we obtain (6.3). The quantities I_1, I_2, I_3 are called *principal inertia moments* of the rigid body.
(ii) Calling T the kinetic energy, G thecenter of mass and $m = \sum_i m_i$ the total mass of the rigid body it is

6.6
$$T = \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2) + \frac{1}{2}mv_O^2 + mv_O \cdot (\omega \wedge OG) \tag{6.6}$$

(6.6) can be proved again by (6.2)

6.7
$$\begin{aligned} T &= \frac{1}{2}\sum_{i=1}^{n} m_i(\mathbf{v}_O + \omega \wedge OP_i) \cdot (\mathbf{v}_O + \omega \wedge OP_i) = \\ &= \frac{1}{2}m\mathbf{v}_O^2 + m\mathbf{v}_O \cdot (\omega \wedge OG) + \frac{1}{2}\sum_{i=1}^{n} m_i(\omega \cdot (OP_i \wedge (\omega \wedge OP_i))) \end{aligned} \tag{6.7}$$

where in the last term we used $a \cdot (b \wedge c) = c \cdot (a \wedge b)$. Such term can be therefore rewritten as

6.8
$$\frac{1}{2}\sum_{i=1}^{n} m_i(\omega \cdot (OP_i \wedge (\omega \wedge OP_i))) = \frac{1}{2}\omega \cdot \mathbf{K}_O \equiv T_O \tag{6.8}$$

and (6.6) follows from (6.3).
Remark: (6.6) shows that if $O = G$ the kinetic energy can be decomposed in the kinetic energy of G plus the kinetic energy of the motion around G (*Koenig's theorem*).

We want to define suitable coordinates in order to describe the position of $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ with respect to $(O; \bar{\mathbf{i}}, \bar{\mathbf{j}}, \bar{\mathbf{k}})$. A natural choice (*Euler's angles*) is represented in Fig. 6.1:

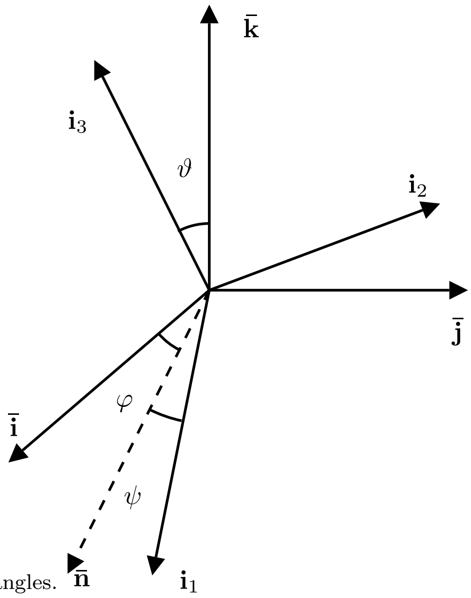

Fig. 6.1: The Euler angles.

Let $\bar{\mathbf{n}}$ be the unit vector in the direction of the intersection between the planes $\mathbf{i}_1, \mathbf{i}_2$ and $\bar{\mathbf{i}}, \bar{\mathbf{j}}$. The Euler's angles are the angle $\bar{\varphi}$ between $\bar{\mathbf{i}}$ and $\bar{\mathbf{n}}$, the angle $\bar{\psi}$ between $\mathbf{i}_1$ and $\bar{\mathbf{n}}$ and the angle $\bar{\vartheta}$ between $\bar{\mathbf{k}}$ and $\mathbf{i}_3$. The angular velocity ω can be written in terms of Euler's angles $(\bar{\vartheta}, \bar{\varphi}, \bar{\psi})$ by decomposing it in the directions $\bar{\mathbf{k}}, \mathbf{i}_3, \bar{\mathbf{n}}$:

6.9
$$\omega = \dot{\bar{\vartheta}}\bar{\mathbf{n}} + \dot{\bar{\varphi}}\bar{\mathbf{k}} + \dot{\bar{\psi}}\mathbf{i}_3 \tag{6.9}$$

The vectors $\bar{\mathbf{k}}, \mathbf{i}_3, \bar{\mathbf{n}}$ can be expressed in the frame $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ in the following way

6.10
$$\mathbf{i}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \bar{\mathbf{n}} = \begin{pmatrix} \cos\bar{\psi} \\ -\sin\bar{\psi} \\ 0 \end{pmatrix} \qquad \bar{\mathbf{k}} = \begin{pmatrix} \sin\bar{\vartheta}\sin\bar{\psi} \\ \sin\bar{\vartheta}\cos\bar{\psi} \\ \cos\bar{\vartheta} \end{pmatrix} \tag{6.10}$$

and therefore we can obtain the following expression for ω in terms of $(\bar{\vartheta}, \bar{\varphi}, \bar{\psi})$

6.11
$$\omega = \begin{pmatrix} \sin\bar{\vartheta}\sin\bar{\psi}\,\dot{\bar{\varphi}} + \cos\bar{\psi}\,\dot{\bar{\vartheta}} \\ \sin\bar{\vartheta}\cos\bar{\psi}\,\dot{\bar{\varphi}} + \sin\bar{\psi}\,\dot{\bar{\vartheta}} \\ \dot{\bar{\psi}} + \cos\bar{\vartheta}\,\dot{\bar{\varphi}} \end{pmatrix} \tag{6.11}$$

From (6.8), we obtain the expression of kinetic energy T_O in terms of Euler's angles

6.12
$$T_O = \frac{1}{2}[I_1(\sin\bar{\vartheta}\sin\bar{\psi}\,\dot{\bar{\varphi}}+\cos\bar{\psi}\,\dot{\bar{\vartheta}})2+I_2(\sin\bar{\vartheta}\cos\bar{\psi}\,\dot{\bar{\varphi}}+\sin\bar{\psi}\,\dot{\bar{\vartheta}})^2+I_3(\,\dot{\bar{\psi}}+\cos\bar{\vartheta}\,\dot{\bar{\varphi}})^2] \tag{6.12}$$

By using $(\bar{\vartheta}, \bar{\varphi}, \bar{\psi})$ as Lagrangian coordinates we obtain for the conjugate variables (1.4), by (6.8) and (6.9)

6.13
$$p_{\bar{\vartheta}} = \mathbf{K}_O \cdot \bar{\mathbf{n}} \qquad p_{\bar{\varphi}} = \mathbf{K}_O \cdot \bar{\mathbf{k}} \qquad p_{\bar{\psi}} = \mathbf{K}_O \cdot \mathbf{i}_3 \tag{6.13}$$

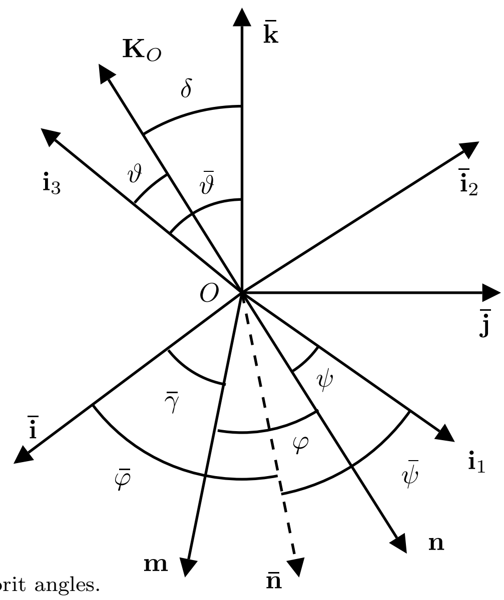

Fig. 6.2: The Deprit angles.

Another set of natural coordinates (*Deprit's angles*) is obtained as follows, see Fig. 6.2. Consider a frame $(O;\mathbf{i},\mathbf{j},\mathbf{k})$ such that $\mathbf{k}$ is parallel to $\mathbf{K}_O$ and $\mathbf{i}$ is on the intersection between the plane orthogonal to $\mathbf{k}$ and the plane $\bar{\mathbf{i}},\bar{\mathbf{j}}$; $\mathbf{j}$ is then uniquely determined. Let (ϑ,φ,ψ) be the Euler's angles of $(O;\mathbf{i}_1,\mathbf{i}_2,\mathbf{i}_3)$ with respect to $(O;\mathbf{i},\mathbf{j},\mathbf{k})$, and $(\delta,\gamma,0)$ the Euler's angles of $(O;\mathbf{i},\mathbf{j},\mathbf{k})$ with respect to $(O;\bar{\mathbf{i}},\bar{\mathbf{j}},\bar{\mathbf{k}})$

Note that the angles $(\bar\vartheta,\bar\varphi,\bar\psi)$, (ϑ,φ,ψ) and $(\delta,\gamma,0)$ are related by a spherical triangle (*Deprit's spherical triangle*). Denoting as $(A,\alpha),(B,\beta),(C,\gamma)$ the pairs of opposite angles, the Deprit's triangle is $(\varphi,\pi-\bar\vartheta),(\bar\psi-\psi,\delta),(\bar\varphi-\gamma,\vartheta)$ (see Fig. 6.2).

Moreover let us define the quantities $A=|\mathbf{K}_O|$, $K=\mathbf{K}_O\cdot\bar{\mathbf{k}}=p_{\bar\varphi}$, $L=\mathbf{K}_O\cdot\mathbf{i}_3=p_{\bar\psi}$. Consider now the transformation

6.14
$$(p_{\bar\vartheta},p_{\bar\varphi},p_{\bar\psi},\bar\vartheta,\bar\varphi,\bar\psi)\to(A,K,L,\varphi,\gamma,\psi) \tag{6.14}$$

Such transformation is well defined: the knowledge of $(p_{\bar\vartheta},p_{\bar\varphi},p_{\bar\psi},\bar\vartheta,\bar\varphi,\bar\psi)$ gives immediately, by (6.13), the components of $\mathbf{K}_O$ and therefore the Deprit's variables; on the other side, from the knowledge of $(A,K,L,\varphi,\gamma,\psi)$ it is immediate to find ϑ, given by $\cos\vartheta=\frac{L}{A}$, and δ, given by $\cos\delta=\frac{K}{A}$; by definition

6.15
$$p_{\bar\psi}=L \qquad p_{\bar\varphi}=K \tag{6.15}$$

and it is easy to check that

6.16
$$p_{\bar\vartheta}=A\sin\vartheta\sin(\psi-\bar\psi) \tag{6.16}$$

finally the Euler's angles $(\bar\vartheta,\bar\varphi,\bar\psi)$, and therefore $p_{\bar\vartheta}$, are obtained by solving the Deprit's spherical triangle by means of (5.1)-(5.5).

A remarkable result is the fact that the transformation (6.14) is canonical. This can be proved by applying (5.6) to the Deprit's triangle, obtaining

6.17
$$d\varphi=\cos\delta d(\bar\varphi-\gamma)+\cos\vartheta d(\bar\psi-\psi)-sin\vartheta\sin(\bar\psi-\psi)d\bar\vartheta \tag{6.17}$$

which can be studied by considering the quantity $Kd\gamma+Ad\varphi+Ld\psi$. By (6.17) we have

6.18
$$Kd\gamma+Ad\varphi+Ld\psi=p_{\bar\varphi}d\gamma+p_{\bar\psi}d\psi+A\cos\delta d(\bar\varphi-\gamma)+A\cos\vartheta d(\bar\psi-\psi)-Asin\vartheta\sin(\bar\psi-\psi)d\bar\vartheta \tag{6.18}$$

Substituting (6.15) and (6.16) in (6.18) we obtain

6.19
$$Kd\gamma+Ad\varphi+Ld\psi=p_{\bar\varphi}d\bar\varphi+p_{\bar\psi}d\bar\psi+p_{\bar\vartheta}d\bar\vartheta \tag{6.19}$$

which shows the canonicity of (6.14) by (1.6) with $F=.p_{\bar\varphi}\bar\varphi+p_{\bar\psi}\bar\psi+p_{\bar\vartheta}\bar\vartheta$.

It is easy to write now the expression of the kinetic energy T_O in terms of Deprit's variables: by (6.3) and (6.8) we have

6.20
$$T_O=\frac{1}{2}\mathbf{K}_O\cdot I^{-1}\mathbf{K}_O \tag{6.20}$$

By definition of Deprit's variables the components of $\mathbf{K}_O$ in the frame $(O;\mathbf{i}_1,\mathbf{i}_2,\mathbf{i}_3)$ are

6.21
$$\mathbf{K}_O=(\sqrt{A^2-L^2}\sin\psi,\sqrt{A^2-L^2}\cos\psi,L) \tag{6.21}$$

and this gives

6.22
$$T_O=\frac{1}{2}\Big[\frac{L^2}{I_3}+(A^2-L^2)\Big(\frac{\sin^2\psi}{I_1}+\frac{\cos^2\psi}{I_2}\Big)\Big] \tag{6.22}$$

7. First order perturbation theory. The averaging method.

In general solving dynamical problems is very difficult and one has to develop approximation methods. The simplest method is the *averaging method*: our purpose here is precisely to show how it can be used in order to obtain a first description of the main phenomena of the Solar system that cannot be reduced to the Kepler's laws.

It is convenient to study general Hamiltonians of the form

7.1 $$\mathcal{H}(\mathbf{A},\alpha) = H_0(\mathbf{A}) + \eta H_I(\mathbf{A},\alpha) \tag{7.1}$$

where $\mathbf{A} \in V \subset \mathbf{R}^l$, $\alpha \in \mathbf{T}^l$, are canonically conjugated action ($\mathbf{A} = (A_1,\ldots,A_\ell)$) and angle $((\alpha_1,\ldots,\alpha_n\ell))$ coordinates, H_0 is analytic in V and H_I is analytic in $V \times T^l$ and such that for some $N > 0$

7.2 $$H_I(\mathbf{A},\alpha) = \sum_{\substack{\nu\in\mathbf{Z}^l\\ |\nu|\le N}} H_I(\mathbf{A})_\nu e^{i\nu\cdot\alpha} \tag{7.2}$$

Our main goal is to determine a canonical transformation $(\mathbf{A},\alpha) \longleftrightarrow (\mathbf{A}',\alpha')$ in a suitable neighborhood of the initial condition $\mathbf{A}_0$ such that the Hamiltonian (7.1) takes the form

7.3 $$\mathcal{H}^{(1)}(\mathbf{A}',\alpha') = H^{(1)}(\mathbf{A}') + \eta^2 H_I^{(1)}(\mathbf{A}',\alpha') \tag{7.3}$$

Such a canonical transformation can be constructed if suitable conditions are met: a typical condition on $\mathbf{A}_0$ is

7.4 $$(\omega(\mathbf{A}_0)\cdot\nu) \neq 0 \qquad \forall \nu \in \mathbf{Z}^l, \quad 0 < |\nu| \le N \tag{7.4}$$

where

7.5 $$\omega(\mathbf{A}) \equiv \frac{\partial H_0}{\partial \mathbf{A}}(\mathbf{A}) \tag{7.5}$$

For a generic choice of $\mathbf{A}_0$ the condition (7.4) is verified. Since $\eta \ll 1$, we look for a generating function "near the identity", i.e. of the form

7.6 $$F(\mathbf{A}',\alpha) = \mathbf{A}'\cdot\alpha + \eta\Phi(\mathbf{A}',\alpha) \tag{7.6}$$

and we want to impose that the Hamiltonian in the new variables defined by the relations

7.7 $$\mathbf{A} = \mathbf{A}' + \eta\frac{\partial\Phi}{\partial\alpha}(\mathbf{A}',\alpha) \qquad \alpha' = \alpha + \eta\frac{\partial\Phi}{\partial\mathbf{A}'}(\mathbf{A}',\alpha) \tag{7.7}$$

is α–independent up to first order in η. Hence we have to impose that

7.8 $$H_0(\mathbf{A}' + \eta\frac{\partial\Phi}{\partial\alpha}(\mathbf{A}',\alpha)) + \eta H_I(\mathbf{A}' + \eta\frac{\partial\Phi}{\partial\alpha}(\mathbf{A}',\alpha),\alpha) \tag{7.8}$$

is α–independent up to first order in η. By expanding in powers of η and by collecting the terms up to first order we find the following equation for Φ, G

7.9 $$\frac{\partial H_0}{\partial\mathbf{A}'}(\mathbf{A}')\cdot\frac{\partial\Phi}{\partial\alpha}(\mathbf{A}',\alpha) + H_I(\mathbf{A}',\alpha) = G(\mathbf{A}') \tag{7.9}$$

This can be written in terms of Fourier components of Φ as

7.10 $$i(\omega(\mathbf{A}')\cdot\nu)\Phi(\mathbf{A}')_\nu = H_I(\mathbf{A}')_\nu \tag{7.10}$$

This is a first order algebraic equation which can be solved if $\mathbf{A}'$ is chosen in a sufficiently small neighborhood of $\mathbf{A}_0$ in such a way that

7.11
$$(\omega(\mathbf{A}')\cdot\nu)\neq 0 \qquad \forall\nu\in\mathbf{Z}^l,\quad 0<|\nu|\le N \tag{7.11}$$

In such domain the equation for $\Phi(\mathbf{A}',\alpha), G(\mathbf{A}')$ can be solved:

7.12
$$\Phi(\mathbf{A}',\alpha)=\sum_{\substack{\nu\in\mathbf{Z}^l\\ 0<|\nu|\le N}}\frac{H_I(\mathbf{A}')_\nu e^{i\nu\cdot\alpha}}{i(\omega(\mathbf{A}'),\cdot\nu)},\qquad G(\mathbf{A}')=H_I(\mathbf{A}')_{\mathbf{0}} \tag{7.12}$$

The required canonical transformation will be defined on the set of the $\mathbf{A}'$ such that the second of the (7.7) may be inverted and it yields

7.13
$$\alpha=\alpha'+\Delta(\mathbf{A}',\alpha') \tag{7.13}$$

It is worth to remark that by substituting (7.12) in(7.8) and by expanding up to first order in η, the resulting Hamiltonian has the following simple form:

7.14
$$H'(\mathbf{A}')=H_0(\mathbf{A}')+H_I(\mathbf{A}')_{\mathbf{0}} \tag{7.14}$$

This result, known also as *averaging theorem*, shows that the first order of the formal perturbation theory, for a perturbation $H_I(\mathbf{A},\alpha)$ which is a trigonometric polynomial in the angle variables, corresponds to computing the average of $H_I(\mathbf{A},\alpha)$ over the angles α and the change of coordinates generated by Φ.

Neglecting the higher order corrections allows us to find approximate solutions of the equations of motion and this is a method that will be used in the following sections.

One should however keep in mind that a rigorous justification of the averaging approximation is usually very difficult if at all possible. It is nevertheless quite easy to establish bounds on the time within which the averaging approximation can be regarded as valid up to prefixed errors: such times can be pushed, possibly after improving the approximation to higher order (second order averaging, or higher, method), to an extent of becoming of interest even for accurate astronomical predictions (from thousands to millions of years depending on the problem considered).

8. Earth precession Hamiltonian

Imagine that the Earth $\mathcal{E}$ is an ideally rigid homogeneous solid of rotation with equatorial radius R. Assume that the center T revolves on a purely Keplerian orbit $t\to\underline{r}_T(t)$ and fix the frame $(O;\bar{\mathbf{i}},\bar{\mathbf{j}},\bar{\mathbf{k}})$ to be with center $O\equiv T$ and with $\bar{\mathbf{k}}$ axis orthogonal to the plane of the Earth orbit, while the $\bar{\mathbf{i}}$ axis is at the equinox line at a prefixed time (*epoch*). The motion of the Earth is described in the coordinates $(\bar\vartheta,\bar\varphi,\bar\psi)$ by the Lagrangian:

8.1
$$\mathcal{L}=\frac12 J(\dot{\bar\varphi}\cos\bar\vartheta+\dot{\bar\psi})^2+\frac12 I(\dot{\bar\vartheta}^2+\dot{\bar\varphi}^2\sin^2\bar\vartheta)+\int_{\mathcal{E}}\frac{kM_SM_T}{|\underline{r}_T+\underline{x}|}\frac{d\underline{x}}{|\mathcal{E}|} \tag{8.1}$$

with

8.2
$$J=I_3=\int_{\mathcal{E}}(x_1^2+x_2^2)\frac{d\underline{x}}{|\mathcal{E}|},\qquad I=I_1=I_2=\int_{\mathcal{E}}(x_2^2+x_3^2)\frac{d\underline{x}}{|\mathcal{E}|} \tag{8.2}$$

being the Earth inertia moments, M_T, M_S being the masses of the Earth and of the Sun, k being the gravitational constant and $|\mathcal{E}|$ being the Earth volume: in the case of an ellipsoid with polar radius $(1-\eta)R$ it is $J=(2/5)R^2M_T$, $I=J(1-\eta+\eta^2/2)$.

The Lagrangian (8.1) is obtained by writing the kinetic energy of a rigid body given in (6.12), exploiting the fact that $I_1=I_2$, and then adding the integral representing, in the

chosen comoving frame, the potential of the gravitational attraction for an ideal solid body.
In order to obtain an explicitly computable integral we write the integrand in (8.1) in terms of Taylor expansion in the quantity $|\underline{x}|/|\underline{r}_T|$ and we obtain

8.3
$$\begin{aligned}\int_{\mathcal{E}} \frac{d\underline{x}}{|\underline{r}_T+\underline{x}|} &= \frac{1}{|\underline{r}_T|}\int_{\mathcal{E}} \frac{d\underline{x}}{\sqrt{1+2\frac{\underline{x}\cdot\underline{r}_T}{|\underline{r}_T|^2}+\frac{|x|^2}{|\underline{r}_T|^2}}} = \\ &= \frac{1}{|\underline{r}_T|}\int_{\mathcal{E}} d\underline{x}\Big(1-\frac{\underline{x}\cdot\underline{r}_T}{|\underline{r}_T|^2}-\frac{1}{2}\frac{|x|^2}{|\underline{r}_T|^2}+\frac{3}{2}\frac{(\underline{x}\cdot\underline{r}_T)^2}{|\underline{r}_T|^4}+O\Big(\frac{|\underline{x}|}{|\underline{r}_T|}\Big)^3\Big)\end{aligned} \tag{8.3}$$

By symmetry considerations it is possible to show that the error is actually of order $O\big(|\underline{x}|/|\underline{r}_T|\big)^4$.
The gravitational potential may be rewritten, in this approximation, as

8.4
$$-V=\frac{kM_SM_T}{|\underline{r}_T|}+\frac{kM_SM_T}{2|\underline{r}_T|^3}\int_{\mathcal{E}}(3(u_{\underline{r}_T}\cdot\underline{x})^2-|\underline{x}|^2)\frac{d\underline{x}}{|\mathcal{E}|} \tag{8.4}$$

where $u_{\underline{r}_T}$ is the unit vector parallel to $\underline{r}_T$. Using the symmetry between the coordinates x_1 and x_2 and denoting by α the angle between $\underline{r}_T$ and the axis $\mathbf{i}_3$ we obtain

8.5
$$\begin{aligned}-V=&\frac{kM_SM_T}{|\underline{r}_T|}+\frac{kM_SM_T}{2|\underline{r}_T|^3}\int_{\mathcal{E}}(3(x_3^2\cos^2\alpha+x_1^2(1-\cos^2\alpha))-2x_1^2-x_3^2)\frac{d\underline{x}}{|\mathcal{E}|}=\\ =&\frac{kM_SM_T}{|\underline{r}_T|}+\frac{kM_S}{2|\underline{r}_T|^3}(I-J)(3\cos^2\alpha-1)=\\ =&\frac{kM_SM_T}{|\underline{r}_T|}-\frac{kM_S}{2a^3}(\frac{a}{|\underline{r}_T|})^3\eta_1 J(3\cos^2\alpha-1)\end{aligned} \tag{8.5}$$

where a is the major semiaxis of the Earth orbit, and the *mechanical ellipticity* η_1 is defined by

8.6
$$\eta_1=(J-I)/J \tag{8.6}$$

Supposing the Earth be an ellipsoid with polar radius $(1-\eta)R$ it is $\eta_1=\eta-\eta^2/2$.
It is now easy to see that $\cos\alpha$ can be written in terms of Euler's angles as

8.7
$$\cos\alpha=\sin\bar\vartheta\sin(\bar\varphi-\lambda_T) \tag{8.7}$$

where λ_T is the angle between $\underline{r}_T$ and the axis $\bar{\mathbf{i}}$. This allows us to write the Lagrangian of our problem as

8.8
$$\begin{aligned}\mathcal{L}=&\frac{1}{2}J(\dot{\bar\varphi}\cos\bar\vartheta+\dot{\bar\psi})^2+\frac{1}{2}I(\dot{\bar\vartheta}^2+\dot{\bar\varphi}^2\sin^2\bar\vartheta)+\frac{kM_SM_T}{|\underline{r}_T|}-\\ &-\frac{kM_S}{2a^3}(\frac{a}{|\underline{r}_T|})^3\eta_1 J(3(\sin\bar\vartheta\sin(\bar\varphi-\lambda_T))^2-1)\end{aligned} \tag{8.8}$$

The corrresponding motions are difficult to study, even approximately, because the Lagrangian does not pertain to a one-dimensional case.
The problem is better studied by using Deprit's variables. Due to the canonicity of the transformation (6.14) the Hamiltonian of the system becomes

8.9
$$H=\frac{1}{2}\Big[\frac{L^2}{J}+\frac{(A^2-L^2)}{I}\Big]-\frac{kM_SM_T}{|\underline{r}_T|}+\frac{kM_S}{2a^3}(\frac{a}{|\underline{r}_T|})^3\eta_1 J(3\cos^2\alpha-1) \tag{8.9}$$

In (8.9) the quantity $\cos^2\alpha$ has to be written in terms of Deprit's variables. This can be done as follows

8.10
$$\cos\alpha=\sin\bar\vartheta\sin(\bar\varphi-\lambda_T)=\sin\bar\vartheta\sin(\bar\varphi-\gamma)\cos(\lambda_T-\gamma)+\sin\bar\vartheta\cos(\bar\varphi-\gamma)\sin(\lambda_T-\gamma) \tag{8.10}$$

by (5.3) and (5.4) applied to the Deprit's spherical triangle we obtain

$$\begin{aligned}\cos\alpha &= \sin(\lambda_T-\gamma)(\cos\varphi\cos\delta\sin\vartheta+\sin\delta\cos\vartheta)-\cos(\lambda_T-\gamma)\sin\vartheta\sin\varphi = \\ &= \sin(\lambda_T-\gamma)\Big(\cos\varphi\frac{K}{A}\frac{\sqrt{A^2-L^2}}{A}+\frac{L}{A}\frac{\sqrt{A^2-K^2}}{A}\Big)-\cos(\lambda_T-\gamma)\sin\varphi\frac{\sqrt{A^2-L^2}}{A}\end{aligned} \tag{8.11}$$

8.11 The *simple solar precession* Hamiltonian is now given by (8.9) posing $|\underline{r}_T| = a$ and $\lambda_T = \lambda_0 + \omega_T t$, $\omega_T^2 = \frac{kM_S}{2a^3}\mu$, $\mu = \frac{M_S+M_T}{M_S}$, *i.e.* assuming a circular Keplerian orbit for the Earth, and setting $\mu = 1$, obtaining

8.12

$$\mathcal{H} = \frac{A^2}{2J} + \eta_2\frac{A^2-L^2}{2J} + \frac{3}{2}\omega_T^2 J\eta_1\cos^2\alpha \tag{8.12}$$

with $\eta_2 = (J-I)/I \approx \eta(1+\eta/2)$ (η_1 is defined in (8.6)) and $\cos\alpha$ written in terms of (8.11). Although the Hamiltonian (8.12) is not integrable it is suitable to be easily studied ny the averaging method. In fact $\eta \ll 1$, and therefore (8.12), up to the first order in η, has the form

8.13

$$\mathcal{H} = H_0(\mathbf{A}) + \eta H_I(\mathbf{A},\alpha) \tag{8.13}$$

where $(\mathbf{A},\alpha)$ is a set of action-angle variables, $H_0(\mathbf{A})$ is an integrable Hamiltonian, giving rise trivially to quasi-periodic motions, and η is a small constant. In our case

8.14

$$H_0(\mathbf{A}) = \frac{A^2}{2J} \qquad H_I(\mathbf{A},\alpha) = \frac{\eta_1}{\eta}\frac{3}{2}\omega_T^2 J\cos^2\alpha + \frac{\eta_2}{\eta}\frac{A^2-L^2}{2J} \tag{8.14}$$

and in section 9 the above Hamiltonian systems will be studied in a first approximation (via the averaging method of section 7).

9. Equinox precession.

To apply the averaging method described in section 7 we have to compute the average value $\langle\cos^2\alpha\rangle$. From (8.11), exploiting the fact that for a generic angle variable β conjugated to an action variable we have $\langle\cos^2\beta\rangle = \langle\sin^2\beta\rangle = 1/2$, we obtain

9.1

$$\langle\cos^2\alpha\rangle = \frac{1}{4}\frac{K^2}{A^2}\frac{A^2-L^2}{A^2} + \frac{1}{2}\frac{L^2}{A^2}\frac{A^2-K^2}{A^2} + \frac{1}{4}\frac{A^2-L^2}{A^2} \tag{9.1}$$

and therefore

$$\mathcal{H}'(A,K,L) = \frac{A^2}{2J} + \eta_2\frac{A^2-L^2}{2J} + \frac{3}{8}\omega_T^2 J\eta_1\Big(1+\frac{K^2}{A^2}\Big)\frac{A^2-L^2}{A^2} + \frac{3}{4}\omega_T^2 J\eta_1\frac{L^2}{A^2}\frac{A^2-K^2}{A^2} \tag{9.2}$$

9.2 The system described by the Hamiltonian (9.2) is now integrable, and the equation of motions are

9.3

$$\begin{aligned} &A = A_0 \quad K = K_0 \quad L = L_0 \\ &\dot\varphi = \frac{\partial\mathcal{H}'}{\partial A} \quad \dot\psi = \frac{\partial\mathcal{H}'}{\partial L} \quad \dot\gamma = \frac{\partial\mathcal{H}'}{\partial K}\end{aligned} \tag{9.3}$$

The explicit expression of the frequencies $\dot\varphi, \dot\psi, \dot\gamma$, however, are still quite involved. In the application of (9.2) to the motion of the Earth axis one can simplify them by using the fact that the angle ϑ is extremely small, namely $\vartheta \ll \eta$. It is therefore easy to see that by neglecting in (9.2) the term $\frac{3}{8}\omega_T^2 J\eta_1\Big(1+\frac{K^2}{A^2}\Big)\frac{A^2-L^2}{A^2}$, and by setting $\frac{L^2}{A^2} = 1$ one introduces in the equations of motion an error negligible up to the first order in η.
With this assumption the Hamiltonian becomes

9.4
$$\mathcal{H}'(A,K,L) = \frac{A^2}{2J} + \eta_2 \frac{A^2 - L^2}{2J} + \frac{3}{4}\omega_T^2 J \eta_1 \frac{A^2 - K^2}{A^2} \tag{9.4}$$

and one obtains simply

9.5
$$\dot\varphi = \frac{A}{J}(1+\eta_2) \qquad \dot\psi = -\frac{L}{J}\eta_2 = -\frac{A\cos\vartheta}{J}\eta_2 \qquad \dot\gamma = -\frac{3}{2}\omega_T^2 \eta_1 \frac{JK}{A^2} \tag{9.5}$$

The motion of the Earth axis may be described, in this approximation, as follows. Defining as ω_D the angular velocity along the $\mathbf{i}_3$ axis, we have $\omega_D = \dot\varphi\cos\vartheta + \dot\psi$. Neglecting the terms in $O(\vartheta^2)$ we have $\omega_D = A/J$. Within the same approximation

9.6
$$\dot\varphi = (1+\eta_2)\omega_D \qquad \dot\psi = -\eta_2\omega_D \qquad \dot\gamma = -\frac{3}{2}\frac{\omega_T^2}{\omega_D}\eta_1 \cos\delta \stackrel{def}{=} -\omega_p^S \tag{9.6}$$

This is the contribution due to the Sun to the motion of Earth axis. The Moon gives a contribution which can be computed along the same lines: supposing the Moon, as done previously with the Sun, on a circular Keplerian orbit, the potential due to the Moon can be written as

9.7
$$\frac{kM_L}{2a^3}\eta_1 J(3\cos^2\alpha_L - 1) \tag{9.7}$$

where M_L is the Moon mass, a_L is the radius of its orbit and α_L is the angle between the axis $\mathbf{i}_3$ and the vector connecting the centers of the Earth and the Moon. Assuming the Moon on the ecliptic and writing $\cos^2\alpha_L$ in terms of de Prit's variables and averaging the Moon potential on the fast angles as before one obtains

9.8
$$\dot\gamma = -\omega_p^L = -\eta_1 \frac{3kM_L}{2a_L^3}\frac{JK}{A^2} + O(\eta_1 e^2) = -\omega_p^S (\frac{a}{a_L})^3 \frac{M_L}{M_S} \tag{9.8}$$

so that the total luni-solar (simple) precession is:

9.9
$$\omega_p = \omega_p^S + \omega_p^L = \lambda_p^S\big(1 + (\frac{a}{a_L})^3\frac{M_L}{M_S}\big) \tag{9.9}$$

Since $\mathbf{i}_3$ and $\mathbf{k}$ are almost coinciding, we can study the motion of the Earth axis with respect to the fixed frame up to the first order in ϑ. In this approximation we have

9.10
$$\bar\vartheta = \delta + \sin\vartheta\cos\varphi \tag{9.10}$$

9.11
$$\bar\varphi - \gamma = \sin\vartheta \frac{\sin\varphi}{\sin\delta} \tag{9.11}$$

9.12
$$\bar\psi - \psi = \varphi + \sin\vartheta\sin\varphi\cot\delta \tag{9.12}$$

Therefore neglecting the terms in ϑ in (9.11) we have $\gamma \approx \bar\varphi$, and ω_p represents $\dot{\bar\varphi}$ and the total (average) precession of the equinox.

The total rate of lunisolar precession in the above approximation gives $T_p \sim 2.51\,10^4$ years, or a yearly precession of the equinoxes of $\sim 51.6''$ per year. Note that only 1/3 of the lunisolar precession is due to the Sun.

The observed value of the lunisolar precession is $50.38''$ per year. It is easy to show that the discrepancy does not come from other contributions: *e.g.* considering the attraction of planet Jupiter, which can be estimated simply by assuming that it gravitates around the Earth on a circular orbit, one sees that its contribution to the precession would give a much smaller to the precession: with obvious notations it would be a fraction of the order of $(a/a_J)^3 M_J/M_S$, *i.e.* $O(10^{-5})$ of the solar precession.

Therefore it less crude approximations are needed to obtain a better agreement with the observed data. We can consider the corrections due to the fact that the orbits of the

Earth around the Sun and of the Moon around the Earth are not circular, and to the fact that the orbit plane of the Moon does not coincide with the ecliptic.
The resulting corrections, however, are still too small, and the discrepancy between the observed and the computed precession remains of the order of some percent.
Note that by (9.10)-(9.11) one expects that $\bar{\varphi}$ and $\bar{\vartheta}$ oscillate around their average value with an amplitude of the order of ϑ and a frequency of the order of ω_D and ω_T. The results obtained by (9.10)-(9.11), however, are not quantitatively reliable, because we obtained them by averaging on angles with period of the same magnitude of the effects we want to study. Therefore in next section we will study the Hamiltonian (8.9) averaging it only on the faster angle φ, and taking in account the terms proportional to ϑ. This is the so-called *nutation theory*.

10. Nutation.

The Hamiltonian of the classical theory of nutation averages the Hamiltonian (8.12) over the fast angles φ, but *neither* over the relatively slower angles λ *nor* over the very slow γ. The Hamiltonian thus obtained should reliably describe motions over a time scale $\gg 2\pi/\omega_D = 1$ day and it is:

10.1
$$H = \frac{A^2}{2J} + \eta_2 \frac{A^2 - L^2}{2J} + \frac{3}{2}\eta_1 \omega_T^2 J\Big(1 - \frac{K^2}{A^2}\Big)\sin^2(\lambda_T - \gamma) \tag{10.1}$$

Note that, since the quantities A and L are constant, we can simply study the one dimensional Hamiltonian

10.2
$$H_D = \frac{3}{2}\eta_1 \omega_T^2 J\Big(1 - \frac{K^2}{A^2}\Big)\sin^2(\lambda_T - \gamma) \tag{10.2}$$

in which the quantity A plays the role of a constant parameter.
The new Hamiltonian H_D depends on time through the angle $\lambda_T = \lambda_0 + \omega_T t$. Nevertheless H_D is integrable by quadratures, because there is a conserved quantity: since

10.3
$$\dot{K} = -\frac{\partial H_D}{\partial \gamma} = -\frac{3}{2}\eta_1 \omega_T^2 J\Big(1 - \frac{K^2}{A^2}\Big) 2\sin(\lambda_T - \gamma)\cos(\lambda_T - \gamma) \tag{10.3}$$

and

10.4
$$\frac{dH_D}{dt} = \frac{\partial H_D}{\partial t} = -\frac{3}{2}\eta_1 \omega_T^2 J\Big(1 - \frac{K^2}{A^2}\Big) 2\sin(\lambda_T - \gamma)\cos(\lambda_T - \gamma)\omega_T \tag{10.4}$$

one has

10.5
$$\frac{d}{dt}(H_D - \omega_T K) = 0 \tag{10.5}$$

Setting $\gamma - \omega_T t = \tilde{\gamma}$ and calling $K_0, \tilde{\gamma}_0, t_0$ the values of $K, \tilde{\gamma}, t$ when $\gamma - \omega_T t = 0$ one has, from the conservation of $H_D - \omega_T K$, that

10.6
$$K - K_0 = \frac{H_D}{\omega_T} \tag{10.6}$$

Moreover one obtains from the Hamilton equation for γ

10.7
$$\dot{\gamma} = \frac{\partial H_D}{\partial K} = -\frac{3}{2} J \eta_1 \omega_T^2 (2K/A^2)\sin^2(\lambda_T - \gamma) \tag{10.7}$$

and therefore one can write the following quadrature:

10.8
$$\begin{gathered} \int_{\tilde{\gamma}_0}^{\tilde{\gamma}} \frac{d\gamma'}{-\omega_T - \frac{3}{2} J \eta_1 \omega_T^2 (2K/A^2)\sin^2\tilde{\gamma}'} = t - t_0 \\ \frac{3}{2}\eta_2 J \omega_T^2 (1 - \frac{K^2}{A^2})\sin^2\tilde{\gamma} - \omega_T K = -\omega_T K_0 \end{gathered} \tag{10.8}$$

To solve (10.3) and (10.7) to first order in η one remarks that from (10.6) the fluctuation of K around its mean value are of order η, and from the first of (10.8) it is clear that $\gamma = \omega_T t + o(\varepsilon)$. Then, neglecting the variations of K and γ of higher order in η, one finds that the motion is:

10.9
$$\begin{aligned}\dot\gamma =& -\omega_p t - \frac{3}{2}\omega_T^2 J\eta_1 \frac{2K_0}{A^2}(\sin^2(\omega_T t + \lambda_0 - \gamma_0) - \frac{1}{2})\\ \dot K =& -\frac{3}{2}\omega_T^2 J\eta_1 (1 - \frac{K_0^2}{A^2})(\sin 2(\omega_T t + \lambda_0 - \gamma_0)\end{aligned} \tag{10.9}$$

Recalling that $\cos\delta = K/A$, $A/J = \omega_D$ and setting $\beta_0 = 2(\lambda_0 - \gamma_0)$, $\delta = \delta_0 + \delta'$ and $\gamma + \omega_p t = \gamma'$ we rewrite (10.9) as

10.10
$$\begin{aligned}\delta' =& \frac{3}{4}\eta\Big(\frac{\omega_T}{\omega_D}\Big)\sin\delta_0 \sin(2\omega_T t + \beta_0)\\ \gamma' =& \frac{3}{4}\eta\Big(\frac{\omega_T}{\omega_D}\Big)\cos\delta_0 \cos(2\omega_T t + \beta_0)\end{aligned} \tag{10.10}$$

where to obtain the first line of (10.10) we used the fact that, up to the first order in $\delta - \delta_0$,

$$\dot K = \frac{d}{dt}(K - K_0) = A\frac{d}{dt}(\cos\delta - \cos\delta_0) = -A\sin\delta_0 \frac{d}{dt}(\delta - \delta_0)$$

The motion in (10.10) express the deviations from the *mean* precession motion and it shows that the Earth axis moves on a small ellipse with a period equal, in this approximation, to $2\pi/\omega_T$. This is the solar *nutation motion*.
To compute the nutation motion due to the Moon it is easy to see that the main contribution does not come from the revolution of the Moon around the Earth, but to the precession motion of the plane of the revolution of the Moon with respect to the ecliptic on a cone of angle equal to the Moon inclination $i_L \sim 5^o$ and with period $T_{pL} = 2\pi/\omega_{pL} \sim 19$ years. The nutation motion due to the Moon turns out to describe an ellipse about 10 times larger than the one found above for the Sun contribution and with a period of the order of $2\pi/2\omega_{pL}$.

One can check that the precession of the Moon plane is (mainly) due to the gravitational force of the Sun, as the following argument shows. One can imagine, for the purpose of studying phenomena that take place over a time scale large with respect to the Moon period of revolution ($T_L =\sim 27$ days) that the Moon is uniformly spread on its orbit on an ring of radius a_L whose plane is inclined by i_L over the ecliptic and which is rotating around its center T at velocity ω_L equal to the mean angular velocity of the Moon $\omega_L = 2\pi/T_L$ with $T_L \sim 27$ days. The ring is at a distance a from the Sun and gravitates around it with angular velocity ω_T, (neglecting the eccentricities of Earth and Moon), hence it has a precession that can be calculated from that of the Earth simply by using (see (8.6)) the value η_1 appropriate for an ring, *i.e.* 1/2 because the inertia moments of a ring are $J = M_L a_L^2$ and $I = J/2$. Hence the precession velocity, from the last of (9.6), is $\omega_{pL} = -(3/4)\omega_T^2\omega_L^{-1}\cos i_L$, which gives the approximate period $T_{pL} = 2\pi/\omega_{pL} \sim 19$ years mentioned above.
This computation is remarkable because it simply uses the geometry of the ring, *without any free parameter* η: hence it can be regarded as a test of the gravitation law.

11. Planar restricted three body problem

The motion of the planets around the Sun is approximatively described by Kepler's laws. If one wants to study the orbits of the planets with more accuracy, one has to take into account the influence of the other planets on the motion, and the fact that, most of

all for the inner planets, the potential has not exactly the form (3.1) due to *relativistic corrections.*
We split an example problem in two parts: we first shall study the effect of Jupiter on the orbit of an inner planet, say of Mercury and later we take into account the further corrections due to general relativity effects.
We shall adopt the following simplifying assumptions

(i) The Sun is at rest in the center of the frame of reference.
(ii) Mercury does not influence the motion of Jupiter.
(iii) The system is supposed to be on a fixed plane.
(iv) Jupiter is supposed to be on a circular Keplerian orbit.

Then the Hamiltonian of the system will be

11.1
$$H_\varepsilon(L,G,T_G,\lambda,\gamma,\lambda_G) = -\frac{g^2}{2L^2} + \omega_G T_G - \frac{g\varepsilon}{|\vec\rho_G - \vec\rho|} \tag{11.1}$$

where the term $\omega_G T_G$ gives the equation of motion $\dot\lambda_G = \omega_G$ for the Jupiter anomaly, according to assumption *(iv)*, T_G represents the excess of energy due to assumption *(ii)*, and $\vec\rho_G$ and $\vec\rho$ represent respectively the (vectorial) position of Jupiter and Mercury respectively; $|\vec\rho_G| = \rho_G$ is supposed to be constant. The small parameter ε is given by $\varepsilon = \frac{M_G}{M_S}$.
Denoting by ϑ the angle between $\vec\rho$ and the major semiaxis of the orbit of Mercury (ϑ is also called *true anomaly*), the angle between $\vec\rho_G$ and $\vec\rho$ is given by $\vartheta + \gamma - \lambda_G$. Hence (11.1) can be rewritten as

11.2
$$\begin{aligned} H_\varepsilon(L,G,T_G,\lambda,\gamma,\lambda_G) &= -\frac{g^2}{2L^2} + \omega_G T_G - \frac{g\varepsilon}{\Big(\rho_G^2 + \rho^2 - 2\rho\rho_G\cos(\vartheta+\gamma-\lambda_G)\Big)^{1/2}} = \\ &= -\frac{g^2}{2L^2} + \omega_G T_G - \frac{g\varepsilon}{\rho_G\Big(1 + \Big(\frac{\rho}{\rho_G}\Big)^2 - 2\Big(\frac{\rho}{\rho_G}\Big)\cos(\vartheta+\gamma-\lambda_G)\Big)^{1/2}} \end{aligned} \tag{11.2}$$

We expand (11.2) in powers of ρ/ρ_G: neglecting the constant terms and the terms in $\big(\rho/\rho_G\big)^3$ we obtain

11.3
$$\begin{aligned} H_\varepsilon(L,G,T_G,\lambda,\gamma,\lambda_G) = &-\frac{g^2}{2L^2} + \omega_G T_G - \frac{g\varepsilon}{\rho_G}\Big(\frac{\rho}{\rho_G}\Big)^2 \frac{3\cos^2(\vartheta+\gamma-\lambda_G) - 1}{2} - \\ &- \frac{g\varepsilon}{\rho_G}\frac{\rho}{\rho_G}\cos^2(\vartheta+\gamma-\lambda_G) \end{aligned} \tag{11.3}$$

where H_ε is implicitly function of λ through ρ and ϑ.
The above analysis allows us to compute the precession of the perihelion of Mercury via the results of section 7.

12. The precession of the perihelion of Mercury due to Jupiter

The Hamiltonian (11.3), following the results of section 7, has to be averaged on λ, γ and λ_g. However the dependence on λ is quite implicit. To average (11.3) it is convenient to introduce an auxiliary coordinate, the so called *eccentric anomaly.* The construction of eccentric anomaly is presented in Fig. 12.1.
The following relations hold

12.1
$$\rho = a(1 - e\cos\xi), \qquad \lambda = \xi - e\sin\xi \tag{12.1}$$

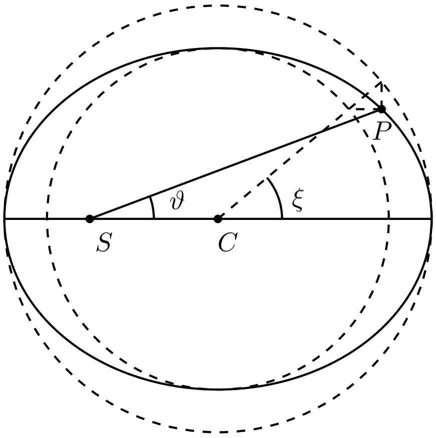

Fig. 12.1: The eccentric anomaly ξ and the true anomaly ϑ of P.

The first has a simple geometrical proof: from Fig. 12.1 if we call, respectively, x and y the components of ρ along the major semiaxis a and the minor semiaxis b of the orbit, it is $x^2 = (a\cos\xi - ae)^2$ and $y^2 = b^2\sin^2\xi = a^2(1-e^2)\sin^2\xi$; so that $\rho^2 = x^2 + y^2 = a^2(1-e\cos\xi)^2$. The second follows from (2.14), which in our case takes the form

12.2
$$t = \int \frac{d\rho'}{A\sqrt{\left(\frac{1}{\rho_-} - \frac{1}{\rho'}\right)\left(\frac{1}{\rho'} - \frac{1}{\rho_+}\right)}} \tag{12.2}$$

From $a = (\rho_+ + \rho_-)/2$ and $b^2 = \rho_+\rho_- = a^2(1-e^2)$ one obtains

12.3
$$t = \frac{T}{2\pi}\int d\rho' \frac{\rho'}{a}\frac{1}{\sqrt{a^2e^2 - (\rho'-a)^2}} \tag{12.3}$$

which implies

12.4
$$\lambda = \int d\rho' \frac{\rho'}{a}\frac{1}{\sqrt{a^2e^2 - (\rho'-a)^2}} = \int (1-e\cos\xi)d\xi = \xi - e\sin\xi \tag{12.4}$$

where in the second equality we have used (12.1)
In order to compute the average of (11.3) with respect to λ, γ and λ_G we write

12.5
$$\begin{aligned} \cos(\vartheta+\gamma-\lambda_G) &= \cos(\vartheta+\gamma)\cos(\lambda_G) + \sin(\vartheta+\gamma)\sin(\lambda_G) \\ (\cos(\vartheta+\gamma-\lambda_G))^2 &= \cos^2(\vartheta+\gamma)\cos^2(\lambda_G) + \sin^2(\vartheta+\gamma)\sin^2(\lambda_G) + \\ &\quad + 2\cos(\vartheta+\gamma)\cos(\lambda_G)\sin(\vartheta+\gamma)\sin(\lambda_G) \end{aligned} \tag{12.5}$$

and we compute the averages

12.6
$$\frac{1}{(2\pi)^3}\int_0^{2\pi} d\lambda \int_0^{2\pi} d\lambda_G \int_0^{2\pi} d\gamma \rho(\lambda)(\cos(\vartheta+\gamma)\cos(\lambda_G) + \sin(\vartheta+\gamma)\sin(\lambda_G)) = 0 \tag{12.6}$$

and

12.7
$$\begin{aligned} &\frac{1}{(2\pi)^3}\int_0^{2\pi} d\lambda \int_0^{2\pi} d\lambda_G \int_0^{2\pi} d\gamma \rho^2(\lambda) \cdot \\ &\qquad \cdot \Big(\cos^2(\vartheta+\gamma)\cos^2(\lambda_G) + \sin^2(\vartheta+\gamma)\sin^2(\lambda_G) + \\ &\qquad + 2\cos(\vartheta+\gamma)\cos(\lambda_G)\sin(\vartheta+\gamma)\sin(\lambda_G)\Big) = \\ &= \frac{1}{2\pi}\int_0^{2\pi} d\lambda \rho^2(\lambda)(\cos^2(\vartheta+\gamma)\frac{1}{2} + \sin^2(\vartheta+\gamma)\frac{1}{2}) = \frac{1}{2\pi}\,\frac{1}{2}\int_0^{2\pi} d\lambda\rho^2(\lambda) \end{aligned} \tag{12.7}$$

Since the average of $\frac{1}{2}(3\cos^2(\vartheta+\gamma-\lambda_G)-1)$ over λ_G is $\frac{1}{4}$ we obtain, if $\langle H_\varepsilon\rangle$ is the average of H_e,

12.8
$$\langle H_\varepsilon\rangle = -\frac{g^2}{2L^2}+\omega_G T_G - \frac{g\varepsilon}{\rho_G}\frac{1}{2\pi}\,\frac{1}{4}\int_0^{2\pi} d\lambda\Big(\frac{a}{\rho_G}\Big)^2(1-e\cos\xi)^2 \tag{12.8}$$

Hence we have to compute the integral

12.9
$$\frac{1}{2\pi}\int_0^{2\pi} d\lambda(1-e\cos\xi)^2 = \frac{1}{2\pi}\int_0^{2\pi} d\xi(1-e\cos\xi)^3 = 1-\frac{3}{2}e^2 \tag{12.9}$$

Note that, by (4.12), one has that $1-\frac{3}{2}e^2=\frac{3}{2}\Big(\frac{G}{L}\Big)^2-\frac{1}{2}$. Neglecting constant terms it is

12.10
$$\langle H_\varepsilon\rangle = -\frac{g^2}{2L^2}+\omega_G T_G - \frac{g\varepsilon}{\rho_G}\frac{1}{4}\Big(\frac{a}{\rho_G}\Big)^2\frac{3}{2}\Big(\frac{G}{L}\Big)^2 \tag{12.10}$$

And the precession of the Mercury's orbit is

12.11
$$\dot\gamma = \frac{\partial\langle H_\varepsilon\rangle}{\partial G} = \frac{g\varepsilon}{L\rho_G}\frac{3}{4}\Big(\frac{a}{\rho_G}\Big)^2\sqrt{1-e^2} \tag{12.11}$$

The relations $L=\sqrt{ga}$ (see (4.10)) and (3.10) imply, finally,

12.12
$$\dot\gamma = \frac{3\varepsilon}{4}\frac{T_M}{T_G^2}2\pi\sqrt{1-e^2} \tag{12.12}$$

The latter gives the perihelion precession of Mercury due to Jupiter. Similar formulae can be obtained for the contributions to the precession dues to the other planets. The result does not match the observed data by an amount of $43''$ per century. The latter can be explained by general relativity as discussed in sections 13,14.

13. The relativistic effects on the Sun attraction.

While the forces due to the other planets make the motion of any of the planets no longer integrable the relativistic corrections change only the form of the central potential.
More precisely, the solution of the Einstein's equation with central symmetry (Schwartzchild, 1916) implies, for a planet moving around the Sun, an equation of motion equivalent to a central motion with a classical potential, depending on initial condition through the quantity A, given by

13.1
$$V_R(\rho) = -\frac{kmM}{\rho}-\frac{kmMA^2}{c^2\rho^3} \tag{13.1}$$

with $m=$ planet's mass, $M=$ Sun's mass, $c=$ speed of light. See appendix 1 for a sketchy derivation of (13.1) from the Schwartzchild solution of Einstein equation.
The quantity $kM/c^2=r_S$ has the dimension of a length, and it is usually called *gravitational radius* of the Sun. The value of such constant is $r_S\approx 1480m$ Due to the small perturbations in (13.1) the trajectories of the planet are no longer closed, and after each revolution the perihelion moves slightly. This movement, called *relativistic precession of the perihelion*, can be described in terms of the small angle $\delta\vartheta$ spanned by the perihelion in a revolution. To compute it we can use the Lagrangian formalism, recalling that the potential is still central and therefore the system is integrable, or we can use the averaging method of section 7. To show the simplicity of the Hamiltonian formalism we will present both approaches, starting from the averaging method. The Lagrangian computations for the same relativistic precession are presented in the appendix A1 to section 14.

The basic axiom of general relativity is that motion is simply a constant speed flow on the geodesics of a pseudo-euclidean metric of a four-dimensional space. This means that the trajectory in general relativity is determined by the variational principle

13.2 $$\delta \int ds = 0 \tag{13.2}$$

where the line element ds is defined by the space metric $g^{\mu\nu}$ as

13.3 $$ds = \sqrt{g^{\mu\nu} \tfrac{dx_\mu}{d\lambda} \tfrac{dx_\nu}{d\lambda}}\; d\lambda \tag{13.3}$$

and λ is a generic parametrization of the trajectory.
We are concerned with the particular case of a point mass, moving with a speed smal light, and therefore moving along geodesics with positive length element. Hence the quantity $\sqrt{g^{\mu\nu} \frac{dx_\mu}{d\lambda} \frac{dx_\nu}{d\lambda}}$ is strictly positive, for any regular parametrization of the trajectory.
Note that, if we choose as parametrization of the motion exactly the arc length of the searched trajectory (geodesic), we have from (13.3), $ds = \sqrt{g^{\mu\nu} \frac{dx_\mu}{ds} \frac{dx_\nu}{ds}}\; ds$ and therefore on the geodesics

13.4 $$\sqrt{g^{\mu\nu} \tfrac{dx_\mu}{ds} \tfrac{dx_\nu}{ds}} = 1 \tag{13.4}$$

It is also useful to define the *proper time* τ simply writing $ds = cd\tau$. The name comes from the fact that τ is the time measured by an observer posed in the reference frame of the moving object. Denoting by u_μ the quadrivelocity $u_\mu = \frac{dx_\mu}{d\tau}$, the variational principle (13.2) can be rewritten as

13.5 $$\delta \int L \; d\tau = 0 \tag{13.5}$$

where

13.6 $$L = \sqrt{g^{\mu\nu} \tfrac{dx_\mu}{d\tau} \tfrac{dx_\nu}{d\tau}} = \sqrt{g^{\mu\nu} u_\mu u_\nu} \tag{13.6}$$

and, on the geodesics,

13.7 $$g^{\mu\nu} u_\mu u_\nu = c^2 \tag{13.7}$$

The gravitational field of the Sun corresponds to the Schwartzchild metric which, in spherical space coordinates, is

13.8 $$(ds)^2 = \Big(1 - \frac{2r_s}{r}\Big)(dt)^2 - \frac{(dr)^2}{1 - \frac{2r_s}{r}} - r^2\Big((d\vartheta)^2 + \sin^2\vartheta (d\phi)^2\Big) \tag{13.8}$$

where $r_s = kM_s/c^2$ is the *gravitational radius* of the Sun, and has the value $r_s \approx 1480m$.
The metric (13.8) has the following features:

(a) $g^{\mu\nu}$ is stationary, *i.e.* it is time independent.
(b) $g^{\mu\nu}$ is rotation invariant.
(c) $g^{\mu\nu}$ is invariant under the symmetry $\vartheta \to \pi - \vartheta$. This implies that a point initially in the plane $\vartheta = \pi/2$ remains indefinitely on such plane. From b), we can always choose, without loss of generality, the initial value $\vartheta = \pi/2$.
(d) $g^{\mu\nu}$ is independent on ϕ.

Properties (a) and (d) imply that t and ϑ are cyclic coordinates for the Lagrangian (13.6). The corresponding conserved quantities are

13.9 $$\frac{\partial L}{\partial \frac{dt}{d\tau}} = \frac{\Big(1 - \frac{2r_s}{r}\Big)\frac{dt}{d\tau}}{\sqrt{g^{\mu\nu} \frac{dx_\mu}{d\tau} \frac{dx_\nu}{d\tau}}}, \qquad \frac{\partial L}{\partial \frac{d\vartheta}{d\tau}} = \frac{r^2 \frac{d\vartheta}{d\tau}}{\sqrt{g^{\mu\nu} \frac{dx_\mu}{d\tau} \frac{dx_\nu}{d\tau}}} \tag{13.9}$$

From (13.7) we deduce

13.10 $$\Big(1-\frac{2r_s}{r}\Big)\frac{dt}{d\tau}=\text{const},\qquad r^2\frac{d\vartheta}{d\tau}=\text{const}\equiv A \tag{13.10}$$

On the other hand (13.7) can be explicitly written as

13.11 $$g^{\mu\nu}u_\mu u_\nu=c^2=\Big(1-\frac{2r_s}{r}\Big)\Big(\frac{dt}{d\tau}\Big)^2-\frac{\Big(\frac{dr}{d\tau}\Big)^2}{1-\frac{2r_s}{r}}-r^2\Big(\Big(\frac{d\vartheta}{d\tau}\Big)^2+\sin^2\vartheta\Big(\frac{d\phi}{d\tau}\Big)^2\Big) \tag{13.11}$$

From (c) we deduce

13.12 $$c^2=\Big(1-\frac{2r_s}{r}\Big)\Big(\frac{dt}{d\tau}\Big)^2-\frac{\Big(\frac{dr}{d\tau}\Big)^2}{1-\frac{2r_s}{r}}-r^2\Big(\frac{d\phi}{d\tau}\Big)^2 \tag{13.12}$$

Substituting (13.10) in (13.12) we finally obtain

13.13 $$\text{const}=2E=\Big(\frac{dr}{d\tau}\Big)^2+\Big(1-\frac{2r_s}{r}\Big)r^2\Big(\frac{d\phi}{d\tau}\Big)^2-\frac{2r_s}{r}c^2 \tag{13.13}$$

and therefore

13.14 $$E=\frac{1}{2}\Big(\dot r^2+r^2\dot\phi^2\Big(1-\frac{2r_s}{r}\Big)\Big)-\frac{g}{r} \tag{13.14}$$

The relativistic one-dimensional potential (13.1) is obtained by substituting (13.10) in (13.14). We also remark that defining the Hamiltonian

13.15 $$H=\frac{1}{2}\Big(p_r^2+\frac{p_\phi^2}{r^2}\Big)-\frac{g}{r}-A^2\frac{r_s}{r^3} \tag{13.15}$$

parameterized by the constant A, we have that for $A=p_\vartheta$ (*but only for this value of* A) the solutions of the equations of motion have (13.10) and (13.14) as conserved quantities.

14. Relativistic precession of the perihelion of Mercury. Averaging method.

The precession of the perihelion of Mercury due to the relativistic corrections can be found by applying the general method of section 7 to the Hamiltonian

14.1 $$H(p_\rho,p_\vartheta,\rho,\vartheta)=\frac{p_\rho^2}{2}+\frac{p_\vartheta^2}{2\rho^2}-\frac{g}{\rho}-\frac{r_sA^2}{\rho^3}\qquad g=kM_s,\ r_s=\frac{kM_s}{c^2} \tag{14.1}$$

where r_s is the gravitational radius of the sum ($rs\approx 1480m$) and A is a fixed parameter. As discussed in section 13, the Hamiltonian system in (14.1) has two conserved quantities (*i.e.* H and p_ϑ), and if the parameter A is fixed to be numerically equal to p_ϑ,the Hamiltonian flow generated by (14.1) on initial data with $p_\vartheta=A$ coincides with the geodesic flow under the Schwartzchild metric and the same initial data.

Note that A has to be considered as an independent parameter: *i.e.* the solutions of the Hamilton equations for (14.1) are interesting form our purposes only if the initial data are such that $p_\vartheta=A$. Replacing p_ϑ with A in (14.1) would be *wrong*.

The Hamiltonian (14.1) is actually integrable, due to the presence of two constants of motion. However, since the quantity $r_s/\rho\approx 10^{-7}$ is extremely small, it can be considered a small perturbation of the Keplerian Hamiltonian (4.1) and therefore it willbe conveniently written in terms of Delaunay variables. This approach has the advantage, with respect to the direct computation presented in section 4, to give a precession computable in terms of the average method.

Following the previous analysis, we write

14.2
$$H(L,G,\lambda,\gamma) = -\frac{g^2}{2L^2} - \frac{r_s A^2}{\rho^3} \tag{14.2}$$

where ρ has to be expressed in terms of the canonical variables. To do this we use the relations (12.1), *i.e.*

$$\rho = a(1 - e\cos\xi) \qquad \lambda = \xi - e\sin\xi$$

We have

14.3
$$H(L,G,\lambda,\gamma) = -\frac{g^2}{2L^2} - \frac{r_s A^2}{a^3(1-e\cos\xi)^3} \tag{14.3}$$

and averaging on λ we have

14.4
$$\langle H\rangle = -\frac{g^2}{2L^2} - \frac{r_s A^2}{a^3}\frac{1}{2\pi}\int_0^{2\pi}\frac{d\lambda}{(1-e\cos\xi)^3} = -\frac{g^2}{2L^2} - \frac{r_s A^2}{a^3}\frac{1}{2\pi}\int_0^{2\pi}\frac{d\xi}{(1-e\cos\xi)^2} \tag{14.4}$$

The integral can be explicitly computed

14.5
$$\langle H\rangle = -\frac{g^2}{2L^2} - \frac{r_s A^2}{a^3(1-e^2)^{3/2}} \tag{14.5}$$

and, using $1 - e^2 = G^2/L^2$, we obtain

14.6
$$\langle H\rangle = -\frac{g^2}{2L^2} - \frac{r_s A^2}{a^3}\frac{L^3}{G^3} \tag{14.6}$$

Hence the Hamilton's equation for γ is

14.7
$$\dot\gamma = 3\frac{r_s A^2}{a^3}\frac{L^3}{G^4} \tag{14.7}$$

Now we can use the fact that numerically the constant A is fixed in such a way that $A = G$, and again $1 - e^2 = G^2/L^2$, obtaining

14.8
$$\dot\gamma = 3\frac{r_s}{a(1-e^2)}\frac{L}{a^2} \tag{14.8}$$

The $\delta\vartheta$, which is the movement of the perihelion in a single revolution, is given by

14.9
$$\delta\vartheta = \dot\gamma T = \frac{6\pi r_S}{a(1-e^2)} \tag{14.9}$$

In appendix the same expression is found in (A1.7).
The result gives a Mercury perihelion precession of $43''$ per century, which is in excellent agreement with observations. This has been one of the most striking confirmations of the validity of the general relativity theory.

Appendix A1: Lagrangian calculation of the precession of Mercury.

Here we come back to (2.15), and we write the angle $\Delta\vartheta$ done in a revolution as

A1.1
$$\Delta\vartheta = 2\int_{\rho_-}^{\rho_+}\frac{\frac{A}{\rho^2}d\rho}{\sqrt{\frac{2}{m}(E - V_R(\rho)) - \frac{A^2}{\rho^2}}} = 2\int_{\rho_-}^{\rho_+}\frac{\frac{A}{\rho^2}d\rho}{\sqrt{2(\frac{E}{m} + \frac{kM}{\rho}) - \frac{A^2}{\rho^2}(1 - \frac{2r_S}{\rho})}} \tag{A1.1}$$

In order to avoid fictitious divergencies it is convenient to rewrite (A1.1) as

$$\Delta\vartheta = 2\int_{\rho_-}^{\rho_+} \frac{\frac{A}{\rho^2}(1-\frac{2r_S}{\rho})d\rho}{\sqrt{2(\frac{E}{m}+\frac{kM}{\rho})-\frac{A^2}{\rho^2}(1-\frac{2r_S}{\rho})}} + 2\int_{\rho_-}^{\rho_+} \frac{\frac{A}{\rho^2}\frac{2r_S}{\rho}d\rho}{\sqrt{2(\frac{E}{m}+\frac{kM}{\rho})-\frac{A^2}{\rho^2}(1-\frac{2r_S}{\rho})}} =$$
$$= -2\frac{\partial}{\partial A}\int_{\rho_-}^{\rho_+} d\rho\sqrt{2(\frac{E}{m}+\frac{kM}{\rho})-\frac{A^2}{\rho^2}(1-\frac{2r_S}{\rho})} + 2\int_{\rho_-}^{\rho_+} \frac{\frac{A}{\rho^2}\frac{2r_S}{\rho}d\rho}{\sqrt{2(\frac{E}{m}+\frac{kM}{\rho})-\frac{A^2}{\rho^2}(1-\frac{2r_S}{\rho})}} \qquad (A1.2)$$

A1.2 Denote now with ρ_-^0, ρ_+^0 the solutions of the unperturbed equation

A1.3
$$2\Big(\frac{E}{m}+\frac{kM}{\rho}\Big) - \frac{A^2}{\rho^2} = 0 \qquad (A1.3)$$

It is easy to see that $|\rho_- - \rho_-^0| = o(\frac{r_S}{\rho})$ and $|\rho_+ - \rho_+^0| = o(\frac{r_S}{\rho})$. Therefore up to the first order in $\frac{r_S}{\rho}$ we can write

$$-2\frac{\partial}{\partial A}\int_{\rho_-}^{\rho_+} d\rho\sqrt{2(\frac{E}{m}+\frac{kM}{\rho})-\frac{A^2}{\rho^2}(1-\frac{2r_S}{\rho})} + 2\int_{\rho_-}^{\rho_+} \frac{\frac{A}{\rho^2}\frac{2r_S}{\rho}d\rho}{\sqrt{2(\frac{E}{m}+\frac{kM}{\rho})-\frac{A^2}{\rho^2}(1-\frac{2r_S}{\rho})}} =$$

A1.4
$$= -2\frac{\partial}{\partial A}\int_{\rho_-^0}^{\rho_+^0} d\rho\sqrt{2(\frac{E}{m}+\frac{kM}{\rho})-\frac{A^2}{\rho^2}(1-\frac{2r_S}{\rho})} + 2\int_{\rho_-^0}^{\rho_+^0} \frac{\frac{A}{\rho^2}\frac{2r_S}{\rho}d\rho}{\sqrt{2(\frac{E}{m}+\frac{kM}{\rho})-\frac{A^2}{\rho^2}(1-\frac{2r_S}{\rho})}} \qquad (A1.4)$$

Now we expand both integrals in (A1.4) up to first order in $\frac{r_S}{\rho}$ obtaining

A1.5
$$\begin{aligned}\Delta\vartheta =& 2\int_{\rho_-^0}^{\rho_+^0} \frac{\frac{A}{\rho^2}d\rho}{\sqrt{2(\frac{E}{m}+\frac{kM}{\rho})-\frac{A^2}{\rho^2}}} - 2\frac{\partial}{\partial A}\int_{\rho_-^0}^{\rho_+^0} \frac{\frac{A^2}{\rho^2}\frac{r_S}{\rho}d\rho}{\sqrt{2(\frac{E}{m}+\frac{kM}{\rho})-\frac{A^2}{\rho^2}}} + \\ &+ 4\int_{\rho_-^0}^{\rho_+^0} \frac{\frac{A}{\rho^2}\frac{r_S}{\rho}d\rho}{\sqrt{2(\frac{E}{m}+\frac{kM}{\rho})-\frac{A^2}{\rho^2}}}\end{aligned} \qquad (A1.5)$$

Since we are interested in the result of the integrals in (A1.4) up to the first order in $\frac{r_S}{\rho}$ we compute them on the Keplerian trajectory. We obtain from zero-th order the unperturbed result 2π, while for the first order, exploiting (2.15), we have

A1.6
$$\delta\vartheta = -2\frac{\partial}{\partial A}\int_0^\pi \frac{Ar_S}{\rho}d\vartheta + 4\int_0^\pi \frac{r_S}{\rho}d\vartheta = -2\frac{\partial}{\partial A}\Big(\frac{\pi kMr_S}{A}\Big) + \frac{4\pi kMr_S}{A^2} \qquad (A1.6)$$

where in last equality we used (3.7) and (3.9) and the fact that the term proportional to $\cos\vartheta$ vanishes. We obtain finally, exploiting again (3.7)

A1.7
$$\delta\vartheta = 6\pi r_S\frac{1}{2}\Big(\frac{1}{\rho_+}+\frac{1}{\rho_-}\Big) = \frac{6\pi r_S}{a(1-e^2)} \qquad (A1.7)$$

where a and e are the major semiaxis and the eccentricity of the orbit.