dal libro di ANDREA MILANI e ZORAN KNEŽEVIĆ
« Dynamics of an asteroid » (in preparazione)

Chapter 2

PERTURBATION THEORY

2.1 Hamilton equations and canonical transformations

Given the Hamilton function $H = H(\mathbf{p}, \mathbf{q})$, of momenta $\mathbf{p}$ and coordinates $\mathbf{q}$, with $\mathbf{p}, \mathbf{q} \in \mathbb{R}^S$, of class $C^k, k \geq 2$, the **Hamilton equations** are

$$\dot{\mathbf{p}} = -\frac{\partial H(\mathbf{p}, \mathbf{q})}{\partial \mathbf{q}} \quad , \quad \dot{\mathbf{q}} = \frac{\partial H(\mathbf{p}, \mathbf{q})}{\partial \mathbf{p}} \; . \tag{2.1}$$

For another function $G(\mathbf{p}, \mathbf{q})$ we can compute the derivative along the integral flow[1] of (2.1) by

$$\dot{G}(\mathbf{p}, \mathbf{q}) = \frac{dG(\mathbf{p}(t), \mathbf{q}(t))}{dt} = \frac{\partial G}{\partial \mathbf{p}}\dot{\mathbf{p}} + \frac{\partial G}{\partial \mathbf{q}}\dot{\mathbf{q}} = -\frac{\partial G}{\partial \mathbf{p}}\frac{\partial H}{\partial \mathbf{q}} + \frac{\partial G}{\partial \mathbf{q}}\frac{\partial H}{\partial \mathbf{p}} = \{G, H\} \tag{2.2}$$

with $\{G, H\}$ being the **Poisson bracket** of G and H. Note that $\{H, H\} = 0$, and $H(\mathbf{p}, \mathbf{q})$ by (1.27) is a first integral[2].

To simplify the presentation, we can also use a notation with a single vector $\mathbf{z} = [\mathbf{p}; \mathbf{q}] \in \mathbb{R}^{2S}$ and $H = H(\mathbf{z})$; we use the **fundamental symplectic matrix**

$$J = \begin{bmatrix} \mathbf{0} & -I \\ I & \mathbf{0} \end{bmatrix} \tag{2.3}$$

where $\mathbf{0}, I$ are $S \times S$, the zero and identity matrix respectively. Then (2.1) and (2.2) can be written as

$$\dot{\mathbf{z}} = J\,\nabla H(\mathbf{z}) \quad , \quad \dot{G} = \frac{\partial G}{\partial \mathbf{z}}\dot{\mathbf{z}} = \nabla G \cdot J\,\nabla H = \{G, H\} \tag{2.4}$$

where $\nabla H = [gradH]^T$ is the gradient of $H(\mathbf{z})$ as a column vector.

We can now discuss the main property of the Hamilton equations, namely the possibility of coordinate changes: a C^k diffeomorphism $\Psi : \mathbf{z} \mapsto \mathbf{z}'$, with $k \geq 2$, is a **canonical transformation with valence** α if for every Hamilton function $H(\mathbf{z})$ (also of class $C^k, k \geq 2$) the function K of $\mathbf{z}'$ corresponding by value to αH, that is such that

$$\alpha\, H(\mathbf{z}) = K\left(\Psi(\mathbf{z})\right) \Longleftrightarrow K(\mathbf{z}') = \alpha\, H\left(\Psi^{-1}(\mathbf{z}')\right) \; , \tag{2.5}$$

[1] The integral flow is the solution that maps the initial conditions to the state at any time t: $\Phi^t : [\mathbf{p}_0; \mathbf{q}_0] \mapsto [\mathbf{p}(t); q(t)]$

[2] We are not considering yet the possibility that H and/or G are also functions of time; see later in Chapter 5

used as Hamilton function in the $\mathbf{z}'$ space, generates equivalent Hamilton equations

$$\dot{\mathbf{z}} = J\nabla H(\mathbf{z}) \Longleftrightarrow \dot{\mathbf{z}}' = J\nabla K(\mathbf{z}') \ ,$$

i.e., the integral flows $\Phi_H^t(\mathbf{z}), \Phi_K^t(\mathbf{z}')$ solving the Hamilton equations commute by Ψ, for every $\mathbf{z}, t$:

$$\Psi \circ \Phi_H^t(\mathbf{z}) = \Phi_K^t \circ \Psi(\mathbf{z}) \ .$$

The most important characterization of the canonical transformations is the following theorem: Ψ, a diffeomorphism as above, is canonical with valence α if and only if its Jacobian matrix $D\Psi = \partial\mathbf{z}'/\partial\mathbf{z}$ is, at each point $\mathbf{z}$, **symplectic** with valence α:

$$D\Psi \, J \, (D\Psi)^T = \alpha \, J \ . \tag{2.6}$$

Proof: by the chain derivatives equation → this is $\dot{\Phi}$

$$\dot{\mathbf{z}}' = \frac{d\Psi(\mathbf{z})}{dt} = D\Phi(\mathbf{z})\,\dot{\mathbf{z}} = D\Phi(\mathbf{z})\,J\nabla H$$

and by (2.5) $\alpha \, gradH = gradK \, D\Psi$, thus → $\dot{\Phi}$

(vettore riga ↙) (vettore colonna ↗)

$$\dot{\mathbf{z}}' = \frac{1}{\alpha} \, D\Phi(\mathbf{z}) \, J \, [D\Phi(\mathbf{z})]^T \, \nabla K(\mathbf{z}')$$

dell'equazione di Hamilton per K ↑

and this equation coincides with the Hamilton equation defined by K in the $\mathbf{z}'$ space, for every H, thus for every K and every possible vector $\nabla K(\mathbf{z}')$, if and only if the matrix is the same, that is if (2.6) holds. *QED*

A canonical transformation with valence $\alpha = 1$ is called a **canonical transformation**. A matrix A satisfying (2.6) with $\alpha = 1$, that is $A\,J\,A^T = J$ is called a **symplectic matrix**; then the eigenvalues of A have a special property: if λ is an eigenvalue of A, then it is also an eigenvalue of A^T, that is there is a non-zero vector $\mathbf{v}$ such that

$$A^T\mathbf{v} = \lambda\mathbf{v} \Longrightarrow AJA^T\mathbf{v} = \lambda AJ\,\mathbf{v} \Longrightarrow J\,\mathbf{v} = \lambda\, AJ\,\mathbf{v} \Longrightarrow A\,(J\,\mathbf{v}) = (1/\lambda)\,(J\,\mathbf{v}) \ .$$

Thus if $\lambda \neq 0$ is an eigenvalue of A, also $1/\lambda$ is an eigenvalue. Note that this applies also for a complex eigenvalue λ, in which case the conjugate $\overline{\lambda}$ is an eigenvalue, as well as $1/\lambda$ and $1/\overline{\lambda}$.

Another way of describing the same result is by using the Poisson brackets: the diffeomorphism $\Psi : (\mathbf{p}, \mathbf{q}) \mapsto (\mathbf{x}, \mathbf{y})$ is canonical if and only if for each $i, j = 1, S$, when the new variables are computed as a function of the old ones:

$$\{y_i, x_j\}\,(\mathbf{p}, \mathbf{q}) = \delta_{ij} \quad , \quad \{x_i, x_j\}\,(\mathbf{p}, \mathbf{q}) = 0 \quad , \quad \{y_i, y_j\}\,(\mathbf{p}, \mathbf{q}) = 0 \ , \tag{2.7}$$

which is in fact the same condition as (2.6) tested coefficient by coefficient in the matrix $D\Psi \, JD\Psi^T$.

Useful examples

The canonical coordinate changes most used in the following are: the change of the canonical polar coordinates $(R, \phi) \mapsto (p, q)$

$$p = \sqrt{2R}\,\cos\phi \quad , \quad q = \sqrt{2R}\,\sin\phi \ ,$$

where it is easy to check that the Jacobian matrix has determinant 1, hence it is a rotation matrix, and the linear transformations applied separately to the p_i and the q_i

$$\mathbf{p} = A\,\mathbf{x} \quad , \quad \mathbf{q} = B\,\mathbf{y}$$

where A, B are real $S \times S$ matrices; the condition for the $2S \times 2S$ matrix $diag[A, B]$ to be symplectic is $B = [A^{-1}]^T$. *

The most important example of canonical transformations with valence $\alpha \neq 1$ are changes of unit and/or dimension, such as the one we have implicitly used when dividing by the masses to transform equation (1.3) into (1.4). Example:

$$\mu_i\,\dot{\mathbf{x}}_i \mapsto \dot{\mathbf{x}}_i \quad , \quad \mathbf{x}_i \mapsto \mathbf{x}_i$$

→ vedi NOTA 1 alla fine del documento

is a canonical transformation with valence $1/\mu_i$, meaning that if the linear momenta are considered per unit mass, then the energy (that is the Hamilton function) has to be considered per unit mass too.

2.2 Integral Flow and Determining Function

The other critical property of Hamilton equations is that the integral flow solving them is a canonical transformation for every time t (for which the flow is defined). Let us chose a **determining function** $\chi(\mathbf{z})$ of class $C^k, k \geq 2$. Then the integral flow $\Phi^t_\chi(\mathbf{z})$ solves the Hamilton equations of χ with initial conditions $\mathbf{z}$

$$\frac{d\Phi^s_\chi(\mathbf{z})}{ds} = J\,\nabla\chi\left(\Phi^s_\chi(\mathbf{z})\right) \quad , \quad \Phi^0_\chi(\mathbf{z}) = \mathbf{z}\;, \tag{2.8}$$

where s is the independent variable, not to be necessarily interpreted as time. The **state transition matrix**, that is the Jacobian matrix of the flow with respect to the initial conditions, is

$$A^s(\mathbf{z}) = \frac{\partial\Phi^s_\chi(\mathbf{z})}{\partial\mathbf{z}} \quad , \quad A^0(\mathbf{z}) = I$$

and by taking the partials of (2.8)

$$\frac{\partial}{\partial\mathbf{z}}\frac{d\Phi^s_\chi(\mathbf{z})}{ds} = \frac{\partial\left(J\nabla\chi(\Phi^s_\chi(\mathbf{z}))\right)}{\partial\mathbf{z}} = J\nabla\nabla\chi(\mathbf{z})\,A^s(\mathbf{z})\;,$$

where $\nabla\nabla\chi$ is the **Hessian matrix** of the second derivatives of determining function χ. The question regarding the above equation is whether the partial derivatives on the left hand side do exist and are continuous. This is indeed true because of the general property of the flow being differentiable [Hartmann, 1964]. Then the order of the two derivatives, with respect to $\mathbf{z}$ and with respect to s, can be exchanged:

$$\frac{dA^s(\mathbf{z})}{ds} = \frac{\partial}{\partial\mathbf{z}}\frac{d\Phi^s_\chi(\mathbf{z})}{ds} = J\,\nabla\nabla\chi(\mathbf{z})\,A^s(\mathbf{z})\;.$$

Thus we have the **variational equation**, a matrix differential equation with its initial conditions

$$\frac{dA^s(\mathbf{z})}{ds} = J\,\nabla\nabla\chi(\mathbf{z})\,A^s(\mathbf{z}) \quad , \quad A^0(\mathbf{z}) = I\;; \tag{2.9}$$

* Sia M una matrice a blocchi $2n \times 2n$ data da $M = \begin{pmatrix} A & C \\ D & B \end{pmatrix}$ dove A, B, C, D sono matrici $n \times n$. La condizione che M sia simplettica è equivalente alle condizioni

$CA^T = AC^T$, $BD^T = DB^T$, $AB^T - CD^T = I$.

Se $n = 1$ queste condizioni si riducono a $\det M = 1$. Una matrice 2×2 è simplettica se e solo se ha determinante unitario.

its solution could also be interpreted as the matrix providing the integral flow of the linear Hamilton equations defined by $\mathcal{H}_\chi(\mathbf{v}) = 1/2\,\mathbf{v}\cdot\nabla\nabla\chi(\mathbf{z}(s))\,\mathbf{v}$, where $\mathbf{v}(t)$ represents a vector displacement propagated from a change in the initial conditions $\mathbf{v}(0)$, when terms of the second order in $|\mathbf{v}|$ are neglected.

The main result in this context is the **theorem of Liouville**: $A^s(\mathbf{z})$ is symplectic, for every $\mathbf{z}, s$ for which it is defined.

Proof: Let $C^s = A^s\,J\,[A^s]^T$; note $C^0 = J$. Then

$$\begin{aligned}\frac{dC^s}{ds} &= \frac{dA^s}{ds}\,J\,[A^s]^T + A^s\,J\left[\frac{dA^s}{ds}\right]^T = J\nabla\nabla\chi\,A^s\,J\,[A^s]^T + A^s\,J\,[J\nabla\nabla\chi\,A^s]^T = \\ &= J\nabla\nabla\chi\,C^s + C^s\,\nabla\nabla\chi\,J^T\end{aligned}$$

because the Hessian matrix$\nabla\nabla\chi$ is symmetric; this equation is satisfied by $C^s = J$, and because the solution to the differential equation for C^s with initial conditions $C^0 = J$ is unique, then $C^s = J$ for all s for which it is defined. *QED*

Thus, given a determining function $\chi(\mathbf{z})$, its integral flow $\Phi^s_\chi : \mathbf{z} \mapsto \mathbf{z}'$ for a fixed s is a canonical transformation, with each Hamilton function $H(\mathbf{z})$ transformed into $H'(\mathbf{z}')$ corresponding by value

$$H'(\mathbf{z}') = H\left(\Phi^{-s}_\chi(\mathbf{z}')\right)\ , \tag{2.10}$$

because $\left(\Phi^s_\chi\right)^{-1} = \Phi^{-s}_\chi$ by the semi-group property (cite??). Moreover, $\Phi^{-s}_\chi = \Phi^s_{-\chi}$, because a change in sign of the Hamiltonian implies a change in sign in all the Hamilton equations.

2.3 Lie series to compute the flow

The question is how to actually compute the integral flow of Hamilton equations. The simplest recipe is to use just the Taylor series, e.g., up to the second order in the independent variable s

$$\mathbf{z}' = \Phi^s_\chi(\mathbf{z}) = \mathbf{z} + s\,\frac{d}{ds}\Phi^s_\chi(\mathbf{z})\bigg|_{s=0} + \frac{1}{2}s^2\,\frac{d^2}{ds^2}\Phi^s_\chi(\mathbf{z})\bigg|_{s=0} + \mathcal{O}(s^3)\ .$$

For this, we can take advantage of (2.2) to compute the derivatives:

$$\frac{d}{ds}\Phi^s_\chi(\mathbf{z})\bigg|_{s=0} = \{\mathbf{z},\chi(\mathbf{z})\} \quad , \quad \frac{d^2}{ds^2}\Phi^s_\chi(\mathbf{z})\bigg|_{s=0} = \{\{\mathbf{z},\chi(\mathbf{z})\},\chi(\mathbf{z})\}\ ,$$

where we have used a "vectorial" notation $\{\mathbf{z},\chi\}$ to indicate the ordinary Poisson bracket applied to each coordinate of the vector $\mathbf{z}$. In this way the Taylor series becomes a **Lie series** for the canonical transformation Φ^s_χ

$$\mathbf{z}' = \mathbf{z} + s\,\{\mathbf{z},\chi\} + \frac{1}{2}s^2\,\{\{\mathbf{z},\chi\},\chi\} + \mathcal{O}(s^3)\ , \tag{2.11}$$

containing only Poisson brackets. The **transform** of a function $H(\mathbf{z})$ into the one corresponding by value[3] is obtained as the Lie series for $H'(\mathbf{z}') = H\left(\Phi^{-s}_\chi(\mathbf{z}')\right)$

$$H'(\mathbf{z}') = H(\mathbf{z}') - s\,\{H,\chi\}(\mathbf{z}') + \frac{1}{2}s^2\,\{\{H,\chi\},\chi\}(\mathbf{z}') + \mathcal{O}(s^3)\ , \tag{2.12}$$

[3]This means that the values of the function are measures of a quantity which is invariant by the transformation of coordinates, such as the energy.

$$H'(z') = H\left(\Phi^{-s}_\chi(z')\right) = H(z') - s\,\frac{d}{ds}H\left(\Phi^{-s}_\chi(z')\right)\bigg|_{s=0} + \frac{s^2}{2}\,\frac{d^2}{ds^2}H\left(\Phi^{-s}_\chi(z')\right) + O(s^3)$$

Infatti la derivata di Lie di H associata al campo vettoriale di χ è

$$\dot{H}(p,q) = \mathcal{L}_\chi H(p,q) = \{H,\chi\}(p,q) = \frac{d}{dt}H\left(\Phi^t_\chi(p,q)\right)\bigg|_{t=0}$$

vedi NOTA 2 alla fine del documento

where the minus sign in the terms of odd order with respect to s is due to the composition with the inverse canonical transformation. Note that the symbols H, χ are to be interpreted as the specific analytic expressions of the functions, not their values; however, χ is an integral of the flow defined by itself because $\{\chi, \chi\} = 0$, thus χ has the same analytical expression and the same value $\chi(\mathbf{z}') = \chi(\mathbf{z})$.

$$\chi\left(\Phi_\chi^t(\mathbf{z})\right) = \chi(\mathbf{z}) = \overline{\chi}$$

2.3.1 Small parameter and homological equations

For a series representation to be meaningful, it is essential that the terms of higher order are small. For this a necessary (not sufficient) condition is that either s is small (as shown by the remainder being of higher order in s) or χ is small (the $\mathcal{O}(s^3)$ is also $\mathcal{O}(\chi^3)$). These two conditions are the same, indeed

$$\Phi_\chi^\epsilon = \Phi_{\epsilon\chi}^1 \ , \quad \rightarrow \text{si vede dalla 2.11}$$

because rescaling the Hamiltonian by a factor ϵ slows down the solutions with respect to the independent variable s, maintaining the same trajectories. Thus we can chose either a fixed value of s, e.g., $s = 1$, and χ including a small parameter ϵ, or a small value $s = \epsilon$ for the independent variable and a χ which is not small, and the result is the same: following [Hori, 1966], in this book we select $s = 1$, which disappears from the equations, and leave ϵ as a small parameter:

$$\chi(\mathbf{z}) = \epsilon\,\chi_1(\mathbf{z}) + \epsilon^2\,\chi_2(\mathbf{z}) + \mathcal{O}(\epsilon^3) \ , \tag{2.13}$$

giving for the transform of the function $H(\mathbf{z})$, which is the function $H'(\mathbf{z}')$ corresponding by value to $H(\mathbf{z})$, a power series in ϵ

$$H'(\mathbf{z}') = H(\mathbf{z}') - \epsilon\,\{H, \chi_1\}(\mathbf{z}') + \frac{1}{2}\epsilon^2\,\{\{H, \chi_1\}, \chi_1\}(\mathbf{z}') - \epsilon^2\,\{H, \chi_2\}(\mathbf{z}') + \mathcal{O}(\epsilon^3) \ . \tag{2.14}$$

dalla 2.12

If in turn H can be expanded in powers of the same small parameter ϵ:

$$H(\mathbf{z}) = H_0(\mathbf{z}) + \epsilon\,H_1(\mathbf{z}) + \epsilon^2\,H_2(\mathbf{z}) + \mathcal{O}(\epsilon^3) \ , \tag{2.15}$$

and this expansion is substituted into (2.14), then the transformed functions are fully expanded in powers of ϵ: for the Hamiltonian

$$\begin{aligned} H'(\mathbf{z}') &= H_0(\mathbf{z}') + \epsilon\,[H_1(\mathbf{z}') - \{H_0(\mathbf{z}'), \chi_1(\mathbf{z}')\}] + \\ &+ \epsilon^2\left[H_2(\mathbf{z}') - \{H_0(\mathbf{z}'), \chi_2(\mathbf{z}')\} - \{H_1(\mathbf{z}'), \chi_1(\mathbf{z}')\} + \frac{1}{2}\{\{H_0(\mathbf{z}'), \chi_1(\mathbf{z}')\}, \chi_1(\mathbf{z}')\}\right] + \ldots \\ &= H_0'(\mathbf{z}') + \epsilon\,H_1'(\mathbf{z}') + \epsilon^2\,H_2'(\mathbf{z}') + \mathcal{O}(\epsilon^3) \ . \end{aligned}$$

Since the equality should be identically satisfied for all small ϵ, we can write a separate equation for each order:

$$H_0'(\mathbf{z}') = H_0(\mathbf{z}') \tag{2.16}$$

$$H_1'(\mathbf{z}') = H_1(\mathbf{z}') - \{H_0, \chi_1\}(\mathbf{z}') \tag{2.17}$$

$$H_2'(\mathbf{z}') = H_2(\mathbf{z}') - \{H_0, \chi_2\}(\mathbf{z}') - \{H_1, \chi_1\}(\mathbf{z}') + \frac{1}{2}\{\{H_0, \chi_1\}, \chi_1\}(\mathbf{z}') \ , \tag{2.18}$$

↳ Hori (1966) « Theory of General Perturbation with Unspecified Canonical Variable »

and so on. These equations are called **homological equations**[4], e.g., eq. (2.17) is the homological equation of order 1. This recursive procedure, in which the equation of order k, which can be solved to get χ_k, uses the functions χ_j for $j = 1, \ldots, k-1$, solutions of the previous ones, could be extended to an arbitrary order: we limit ourselves to $\mathcal{O}(\epsilon^2)$ because this is what is used in practice in a perturbation theory for the asteroids [5].

The key point is that this applies to every possible determining function χ, provided it is smooth. Thus we can suitably chose χ, i.e., use the recurrent equations above as conditions to be satisfied by χ. The first order χ_1 can be chosen in such a way to remove as much as possible of the perturbation H_1, χ_2 can be chosen to remove at least the largest terms of the order 2 perturbations, as given by eq. (2.18).

To understand how the terms of each order can be decomposed into a "removable" portion and a "non-removable" one, we should use a more specific model of the functions such as H and χ, based on an appropriate choice of coordinates.

2.4 Action-angle variables and Lie series

vedi NOTA 3 alla fine del documento sui teoremi di Liouville, di Arnol'd-Kol e sulle variabili azione - angolo

Let us assume that a coordinate change exists such that $\mathbf{z} = (\mathbf{p}, \mathbf{q})$ and the unperturbed Hamiltonian $H_0(\mathbf{z})$, thus the Hamilton equations, are reduced to the very simple form:

$$H_0 = H_0(\mathbf{p}) \quad , \quad \dot{\mathbf{p}} = -\frac{\partial H_0}{\partial \mathbf{q}} = \mathbf{0} \quad , \quad \dot{\mathbf{q}} = \frac{\partial H_0}{\partial \mathbf{p}} = \mathbf{n}(\mathbf{p}) \, ,$$

with $\mathbf{p}$ a vector of constants, that is integrals, for the dynamics defined by H_0, and $\mathbf{q}(t) = \mathbf{n}(\mathbf{p})(t - t_0) + \mathbf{q}(t_0)$ for solution. Thus the variables q_i are each a liner function of time; if they are real variables, they can for suitable times take arbitrarily large values.

If we have enough topological information on the problem defined by H_0 to be able to assume that the equations $\mathbf{p} = \mathbf{p}(t_0)$ define a set which is limited, thus compact, then it is possible to show (see Section 6.1) that such a set must be a torus of the same dimension as the vector $\mathbf{q}$, that is, each of the variables q_i is an angle variable, taking values in the quotient $\mathbb{R}/2\pi\mathbb{Z}$ that is in a circle. Then each element n_i of the vector $\mathbf{n}$ is a frequency, such that the angle q_i makes a revolution in the time $2\pi/n_i$. If this is the case, the variables $(\mathbf{p}, \mathbf{q})$ are called **action-angle variables** for the Hamiltonian H_0.

Then the equations (2.17) and (2.18) are reduced to a much simpler form, containing expressions such as

$$\frac{\partial H_0(\mathbf{p}')}{\partial \mathbf{p}'} = m(\mathbf{p}') \; ; \quad \{H_0(\mathbf{z}'), \chi_1(\mathbf{z}')\} = -\mathbf{n}(\mathbf{p}') \cdot \frac{\partial \chi_1(\mathbf{z}')}{\partial \mathbf{q}'} \, .$$

dipende solo da $\mathbf{p}'$

These expressions are even more constrained taking into account that the $\mathbf{q}$ are angle variable, thus all the functions $H_k(\mathbf{p}, \mathbf{q})$ and $\chi(\mathbf{p}, \mathbf{q})$ need to be periodic of period 2π in each of the

[4] This name comes from the analogy with sequences of equations used in algebraic topology and other branches of geometry.

[5] Even χ_2 either is not used or is not used in full in practical computations, due to a large number of terms to be handled. Such is the situation with the computation of asteroid mean elements which can be limited to χ_1, see Section 2.9

variables q_i. This implies that each of them can be developed in a Fourier series, e.g.:

$$H_1 = \sum_{\mathbf{k}} H_{1\mathbf{k}}(\mathbf{p}) \cos(\mathbf{k}\cdot\mathbf{q}) \;\; ; \;\; \chi_1 = \sum_{\mathbf{k}} \chi_{1\mathbf{k}}(\mathbf{p}) \sin(\mathbf{k}\cdot\mathbf{q}) \;\; ; \;\; H'_1 = \sum_{\mathbf{k}} H'_{1\mathbf{k}}(\mathbf{p}') \cos(\mathbf{k}\cdot\mathbf{q}') \;, \quad (2.19)$$

where $\mathbf{k}$ is a multi-index, that is each coefficient is an integer (with sign, that is in $\mathbb{Z}$). The use of cosine terms only for H_1, H'_1 shall be justified in Section 2.8, and the use of sine only for χ_1 is explained below. Note that $H(\mathbf{z})$ is expanded in the original variables $\mathbf{z}$, then $H(\mathbf{z}')$ has the same expansion, but not the same value. To the contrary $H'(\mathbf{z}')$ is expanded in the new variables $\mathbf{z}'$, and $H'(\mathbf{z}') = H(\mathbf{z})$. As for χ, it has the property that $\chi(\mathbf{z}) = \chi(\mathbf{z}')$, see eq. (2.10), thus it can be expanded in both ways, with no difference.

Then the computation of the Poisson brackets contained in the Lie series expansion can be performed term by term in the Fourier series: equation (2.17) becomes (*) vedi sotto

$$H'_{1\mathbf{k}}(\mathbf{p}') = H_{1\mathbf{k}}(\mathbf{p}') + (\mathbf{k}\cdot\mathbf{n})\, \chi_{1\mathbf{k}}(\mathbf{p}', \mathbf{q}') \;,$$

where we have used the presence of sine functions in χ to guarantee cosine functions in $\{H_0, \chi_1\}$. If H_1 is given as the first order perturbation function to be added to H_0, we can choose χ_1 to simplify the term of the same order H'_1: we select

$$\chi_{1\mathbf{k}}(\mathbf{p}') = -\frac{H_{1\mathbf{k}}(\mathbf{p}')}{\mathbf{k}\cdot\mathbf{n}} \quad (2.20)$$

for each multi-index $\mathbf{k}$ for which this division is possible. If we can assume that the **divisor** $\mathbf{k}\cdot\mathbf{n}$ is never zero apart from the case of $\mathbf{k} = \mathbf{0}$, then χ_1 can be selected in such a way that each $H'_{1\mathbf{k}} = 0$, apart from $H'_{1\mathbf{0}}$, and the new Hamiltonian $H'_1 = H'_{1\mathbf{0}}(\mathbf{p}')$ is a function of the actions only, thus it is also trivially integrable. However, the hypothesis $\mathbf{k}\cdot\mathbf{n} \neq 0$ for all multi-indexes $\mathbf{k} \neq \mathbf{0}$ is far from trivial, and, even if it applies, the convergence of the Fourier series for H_1 and χ_1 does not guarantee the convergence of the Fourier series for χ_1 as computed from (2.20).

Anyway, if $\mathbf{k}\cdot\mathbf{n} \neq 0$ is assumed , it is possible to apply recursively the procedure as an operation on formal series, and push the simplification of H' to higher order; in this task the following formalism is useful. Let

$$\Psi \qquad \Psi : \chi \mapsto \{H_0, \chi\}$$

be an operator between suitable spaces[6]. The **kernel** of Ψ is a subspace $Ker\ g$ containing the Fourier series g such that $\Psi(g) = 0$, that is, the subset of prime integrals of the dynamics defined by H_0. Let us assume we can decompose any Fourier series g, into terms belonging to the kernel of Ψ and terms belonging to the image $Im\ g$ of Ψ:

$$g = \overline{g} + \tilde{g} \quad , \quad \overline{g} \in Ker\ \Psi \quad , \quad \tilde{g} \in Im\ \Psi$$

If we use, as assumed in this section, an $H_0 = H_0(\mathbf{p})$ independent from the angles $\mathbf{q}$, such decomposition is possible by elementary computations, e.g., if $\mathbf{k}\cdot\mathbf{n} \neq 0$ (for $\mathbf{k} \neq \mathbf{0}$) can be assumed

$$\overline{H_1} = H_{1\mathbf{0}}(\mathbf{p}) \quad , \quad \tilde{H}_1 = H_1 - H_{1\mathbf{0}}(\mathbf{p}) \;, \qquad \left(\{H_0, H_1\} = -\frac{\partial H_0}{\partial p} \cdot \frac{\partial H_1}{\partial q} \right)$$

[6]These spaces could be functional spaces of smooth functions, but also spaces containing formal series, without any requirement of convergence, thus not necessarily corresponding to functions.

(*) $H'_1(z') = H_1(z') - \{H_0, \chi_1\}(z')$ $\qquad = m(p')$ $\qquad = \chi_{1,k}(p')\, k \cos(k\cdot q')$

$\longrightarrow$ $H'_{1,k}(p')\cos(k\cdot p') = H_{1,k}(p')\cos(k\cdot q') + \left(\frac{\partial H_{0,k}}{\partial p'} \cdot \frac{\partial \chi'_{1,k}}{\partial q'} \right)$

con $H_{0,k}(p') = m(p')\cdot p'$ e $\tilde{\chi}_{1,k}(p', q') = \chi_{1,k}(p')\,\mathrm{sen}\,(k\cdot q')$

that is, the kernel is just the term of the series without the angles, the image contains all the other terms. For χ_1, since it is chosen only to satisfy eq. (2.20), we can also assume that

$$\overline{\chi_1}(\mathbf{p}') = 0 \quad , \quad \tilde{\chi_1}(\mathbf{p}', \mathbf{q}') = \chi_1 \ .$$

With this notation, the homological equation of order 1 can be written as

$$H_1'(\mathbf{p}', \mathbf{q}') = \overline{H_1}(\mathbf{p}') + \tilde{H_1}(\mathbf{p}', \mathbf{q}') - \{H_0(\mathbf{p}'), \chi_1(\mathbf{p}', \mathbf{q}')\} = \overline{H_1}(\mathbf{p}') \ .$$

Now let us apply the formalism to the second order homological equation (2.18): the right hand side contains four terms, but by substituting $\{H_0, \chi_1\} = \tilde{H_1}$

$$\begin{aligned} H_2'(\mathbf{p}') &= \overline{H_2} + \tilde{H_2} - \{H_0, \chi_2\} - \{H_1, \chi_1\} + \frac{1}{2}\left\{\tilde{H_1}, \chi_1\right\} = \\ &= \overline{H_2} - \frac{1}{2}\left\{\tilde{H_1}, \chi_1\right\} + \left[\tilde{H_2} - \{\overline{H_1}, \chi_1\} - \{H_0, \chi_2\}\right] \ . \end{aligned} \tag{2.21}$$

By introducing the function $K_2 = \left\{\tilde{H_1}, \chi_1\right\}$ we get

$$\begin{aligned} \overline{H_2'} &= \overline{H_2} - \frac{1}{2}\overline{K_2} \\ \tilde{H_2'} &= \tilde{H_2} - \{\overline{H_1}, \chi_1\} - \frac{1}{2}\tilde{K_2} - \{H_0, \chi_2\} \end{aligned} \tag{2.22}$$

where the decomposition $H_2' = \overline{H_2'} + \tilde{H_2'}$ is much more complex than what happens in the homological equation of order 1, but was somewhat simplified because the portions inside the square bracket in (2.21) all belong to the image of Ψ, that is all terms contain the angles $\mathbf{q}'$.

Now we can write the equation defining the desired χ_2 by using only $\tilde{H_2'}$ as portion to be removed:

$$G_2 = \tilde{H_2} - \{\overline{H_1}, \chi_1\} - \frac{1}{2}\tilde{K_2} = \tilde{G_2} \tag{2.23}$$

$$\tilde{H_2'} = \tilde{G_2} - \{H_0, \chi_2\} = 0 \tag{2.24}$$

where the first equation assumes again $\mathbf{k} \cdot \mathbf{n} \neq 0$ for $\mathbf{k} \neq \mathbf{0}$. Because $G_2 = \tilde{G_2}$ is a series without terms independent from the angles, the coefficients $\chi_{2\mathbf{k}}$ of the Fourier series of χ_2 can be computed from the coefficients $G_{2\mathbf{k}}$ of the Fourier series of G_2, with the same divisors appearing in eq. (2.20):

$$\chi_{2\mathbf{k}}(\mathbf{p}') = -\frac{G_{2\mathbf{k}}}{\mathbf{k} \cdot \mathbf{n}} \ . \tag{2.25}$$

Assuming again the same $\mathbf{k} \cdot \mathbf{n} \neq 0$ condition for $\mathbf{k} \neq \mathbf{0}$ we can solve for χ_2 in such a way that it satisfies eq. (2.24). Thus the new Hamiltonian H' can be, at the level of formal series, normalized (up to order 2 in ϵ) by reducing to an integrable Hamiltonian

$$H'(\mathbf{p}', \mathbf{q}') = H_0(\mathbf{p}') + \epsilon\, H_1'(\mathbf{p}') + \epsilon^2\, H_2'(\mathbf{p}') + \mathcal{O}(\epsilon^3) \ .$$

We can easily figure out that the computations would be formally the same, only more complicated, to find a Fourier series for χ_3 in such a way that also H_3' is integrable, and so on, up to any desirable finite order in ϵ. Of course, the computations can in practice become very complex and the convergence issue become worse and worse, as the order in ϵ increases.

ho seguito un'altra presentazione (vedi file separato)

2.5 Action-angle variables for the 2-body problem

To be able to apply the Lie series method, as outlined in the previous two sections, to the gravitational multi-body problem we need to show that suitable action-angle variables can be found, thus integrating the unperturbed Hamiltonian.

Let $\mathbf{r} = (x, y, z)$ be Cartesian coordinates of a point mass moving under the attraction of a (gravitationally active) mass μ fixed at the center in $\mathbf{0}$; of course this dynamical problem can be obtained by reducing a 2-body problem as in Section 1.4. The Lagrange function (for unit point mass) is

$$L_c(x, y, z, \dot{x}, \dot{y}, \dot{z}) = \frac{1}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) + \frac{\mu}{r}\,, \tag{2.26}$$

with r the distance from the center, and μ can be either the mass of a fixed central body or the sum of the two masses in a 2-body problem. In spherical polar coordinates

$$\begin{cases} x &= r\cos\theta\cos\phi &, \quad -\pi/2 < \theta < \pi/2 \\ y &= r\cos\theta\sin\phi &, \quad \phi \in S^1 \\ z &= r\sin\theta &, \quad r > 0 \end{cases}$$

the Lagrange function in polar coordinates corresponding by value to (2.26) is

$$L_p(r, \theta, \phi, \dot{r}, \dot{\theta}, \dot{\phi}) = \frac{1}{2}\left[\dot{r}^2 + r^2\left(\dot{\theta}^2 + \dot{\phi}^2\cos^2\theta\right)\right] + \frac{\mu}{r}\,. \tag{2.27}$$

The Legendre transformation is

$$p_r = \frac{\partial L_p}{\partial \dot{r}} = \dot{r} \quad , \quad p_\theta = \frac{\partial L_p}{\partial \dot{\theta}} = r^2\dot{\theta} \quad , \quad p_\phi = \frac{\partial L_p}{\partial \dot{\phi}} = r^2\cos^2\theta\,\dot{\phi}\,, \tag{2.28}$$

and the corresponding Hamilton function:

$$H_p(p_r, p_\theta, p_\phi, r, \theta, \phi) = \frac{1}{2}\left[{p_r}^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2\cos^2\theta}\right] - \frac{\mu}{r}\,. \tag{2.29}$$

Commuting integrals

Snce H_p does not contain the coordinate ϕ, then $\dot{p}_\phi = \{p_\phi, H_p\} = -\partial H_p/\partial\phi = 0$ and p_ϕ is an integral, which can be interpreted as the component of the angular momentum (of the relative motion, per unit mass) along the z axis of the current reference system; this we shall denote D. On the other hand the length of the angular momentum vector $C = |\mathbf{c}_1|$ is the distance r from the center times the component of the velocity ~~tangential~~ transversal to the orbit; by using (2.28)

$$C = r\sqrt{\left(r\dot{\theta}\right)^2 + \left(r\dot{\phi}\cos\theta\right)^2} = r^2\sqrt{\dot{\theta}^2 + \dot{\phi}^2\cos^2\theta} = \sqrt{p_\theta^2 + \frac{p_\phi^2}{\cos^2\theta}}\,.$$

componente trasversale della velocità (vedi 2.27)

We know C is an integral, as it is clear also because of the separation of variables:

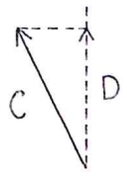
$$\{C, r\} = 0 = \{C, p_r\}$$

and the Hamilton function

$$H_p = \frac{1}{2}\left[{p_r}^2 + \frac{C}{r^2}\right] - \frac{\mu}{r}$$

contains only the variables r, p_r with no action on C, and C itself (which is a function of θ, p_θ, p_ϕ only), thus the functions H_p and C have no variable in common, implying $\{C, H_p\} = 0$. Moreover, $\{C, D\} = 0$ because $D = p_\phi$ and C does not contain ϕ. Thus (H_p, C, D) are 3 **commuting integrals**, that is with Poisson bracket zero for each couple of them.

The fact that a Hamiltonian with 3 degrees of freedom has 3 commuting integrals has a lot of consequences, but for this we need also to check that the integrals are **functionally independent**, that is there is no identically satisfied functional relationship between them. To check functional independence we need to compute the gradients of all the integrals and check if they are linearly independent, at each point in phase space.



$$\begin{aligned}
\nabla H_p &= \left(p_r, \frac{p_\theta}{r^2}, \frac{p_\phi}{r^2\cos^2\theta}, \frac{\partial H_p}{\partial r}, \frac{p_\phi^2 \sin\theta}{r^2\cos^3\theta}, 0\right) \\
\nabla C &= \left(0, \frac{p_\theta}{C}, \frac{p_\phi}{C\cos^2\theta}, 0, \frac{p_\phi^2\sin\theta}{C\cos^3\theta}, 0\right) \\
\nabla D &= (0,0,1,0,0,0) \ .
\end{aligned}$$

The conditions for parallelism are:

$$\begin{aligned}
\nabla H_p \ &|| \ \nabla C \Longleftrightarrow 0 = p_r \ , \ 0 = \frac{\partial H_p}{\partial r} = -\dot{p}_r = -\ddot{r} \Longleftrightarrow r = const \Longleftrightarrow e = 0 \\
\nabla C \ &|| \ \nabla D \Longleftrightarrow \theta = 0, \pi \ , \ \underbrace{p_\theta}_{= r^2\dot{\theta}} = 0 \Longleftrightarrow z \equiv 0 \Longleftrightarrow \sin I = 0 \ .
\end{aligned}$$

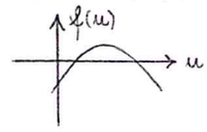
implying that the inclination I, defined by $\cos I = D/C$, is either 0 or π. $\nabla D || \nabla H_p$ cannot occur unless all three are parallel, that is $e = 0 = \sin I$. Moreover, no linear combination of ∇C and ∇D could be a multiple of ∇H_p unless $e = 0$: this implies linear independence.

Let us define a manifold $V = V(E, \overline{C}, \overline{D})$ by assigning a fixed value to each of the 3 integrals:

$$H_p = E \quad , \quad \sqrt{p_\theta^2 + \frac{p_\phi}{\cos^2\theta}} = \overline{C} \quad , \quad p_\phi = \overline{D} \ ;$$

because of the independence of the gradient vectors V is a smooth manifold at each of its points, provided $e > 0$ and $\sin I > 0$. If the energy integral is negative: $E < 0$, then we can solve for p_r from the equation $H_p = E$

$$p_r = \pm\sqrt{2E + \frac{2\mu}{r} - \frac{C^2}{r^2}} \ ,$$



and find that under the square root the polynomial in $u = 1/r$ has 2 positive roots and a negative degree 2 term, thus real values of p_r are possible only in the interval between the roots, which are the distance q at pericenter and Q at apocenter: the values of both r and p_r are limited, on V, to closed intervals. $E < 0$ implies that the eccentricity $e < 1$ (see Section 1.4). We do not need to set limits to the angle variables because they belong to a circle; on V, $p_\phi = D$ is fixed and $|p_\theta| < C$. Thus $V(E, C, D)$ for $E < 0$ is a closed, limited set.

$$2E + 2\mu u - C^2u^2 = 0 \quad \rightarrow \quad C^2u^2 - 2\mu u - 2E = 0$$

$$u_{1,2} = \frac{\mu \pm \sqrt{\mu^2 + 2EC^2}}{C^2} = \frac{\mu}{C^2}(1 \pm e) = \begin{cases} 1/Q \\ 1/q \end{cases}$$

In conclusion, for $E < 0$, $0 < e < 1$, $\sin I > 0$ the set $V(E, C, D)$ is a compact, smooth 3-manifold, and the following inequalities apply:

$$0 < e^2 = 1 + \frac{2EC^2}{\mu^2} < 1 \quad , \quad C > |D| \, .$$

The entire domain in the phase space of positions and velocity corresponding to the conditions $0 < e < 1, \sin I > 0$ is the **Delaunay domain**, foliated in smooth 3-manifolds $V(E, C, D)$.

Angle variables

si riferisce al problema di Keplero

To extend the map $(r, \theta, \phi, p_r, p_\theta, p_\phi) \mapsto (E, C, D)$ to a six dimensional canonical coordinate change we can exploit the well known fact that the solutions of the fixed center problem can be represented by **Keplerian elements** $(a, e, I, \ell, \omega, \Omega)$, where the angle Ω is the longitude of the ascending node and ℓ is the mean anomaly, an angle which is 0 at pericenter and π at apocenter and moves with a constant rate, the **mean motion** n, such that Kepler's third law applies:

$$\dot{\ell} = \{\ell, H_p\} = n \quad , \quad n^2 a^3 = \mu \, . \tag{2.30}$$

The angle variables (ℓ, ω, Ω) are suitable to complete the coordinate system together with the action variables (E, C, D). To show this we can use the notion of conjugated variables, in particular the coordinate conjugated with C is ω, and the coordinate conjugated with D is Ω (apart from constants, which we are allowed to set to 0 in both cases). Indeed D is the Hamiltonian of a rotation around the z axis: the Hamilton equations with D as Hamiltonian are just $d\phi/ds = \partial D/\partial p_\phi = 1$, all the other derivatives being 0.

* vedi NOTA 4 alla fine del documento

As for the dynamics defined by C, we can consider it in a cartesian reference system in which the z axis is the axis of the angular momentum (as in Section 1.4). In this reference system, $C = D$ and its dynamics is again a rotation around the z axis, which is the axis of the angular momentum. Thus the Poisson brackets

$$\{\omega, C\} = 1 \quad , \quad \{\Omega, D\} = 1 \qquad \{q_i, p_j\} = \delta_{ij}$$

but also the ones indicating that the action of C, D change nothing else but the conjugated variables

$$\{\omega, D\} = 0 \quad , \quad \{\Omega, C\} = 0 \quad , \quad \{\ell, C\} = 0 \quad , \quad \{\ell, D\} = 0 \, .$$

Action variables

Now we need to find a conjugated variable to ℓ, which we are going to call L. After the coordinate change

$$\Psi : (a, e, I, \ell, \omega, \Omega) \mapsto (L, C, D, \ell, \omega, \Omega)$$

$K(L) = H_p$ is the function corresponding by value, such that there are constant frequencies (on a given $V(E, C, D)$) for each angle variable

$$\dot{\ell} = \frac{\partial K(L)}{\partial L} = n(L) \quad , \quad \dot{\omega} = \frac{\partial K(L)}{\partial C} = 0 \quad , \quad \dot{\Omega} = \frac{\partial K(L)}{\partial D} = 0 \, .$$

→ The Hamilton function defines the Hamilton equations, which are the equations of motion as a function of time (taken as independent variable); this can be described by the statement that (H, t) are conjugated variables. Similarly, other integrals can be taken as Hamilton functions, and provide with the corresponding Hamilton equations the motion under the action of 1-parameter symmetry groups; e.g. D is the Hamiltonian associated to the rotation around the z axis of the angle Ω, with Ω (in radiants) as independent variable.

To find the analytic expression of the single variable function $K(L)$ we use the third Kepler law (2.30) and the relationship between a and the energy E of the 2-body problem (1.37)

$$n = \sqrt{\frac{\mu}{a^3}} = \frac{(-2\,E\,)^{3/2}}{\mu} \,. \qquad E = -\frac{\mu}{2a} \,, \quad a = -\frac{\mu}{2E}$$

By using the correspondence by value $K(L) = E$

$$\frac{dK(L)}{dL} = \frac{(-2\,K(L))^{3/2}}{\mu} \iff dL = \frac{\mu\, dK(L)}{(-2\,K(L))^{3/2}}$$

with solution (unique apart from a constant we set to 0)

$$L = \frac{\mu}{\sqrt{-2\,K(L)}} = \sqrt{\mu\, a} \quad , \quad K(L) = -\frac{\mu^2}{2\,L^2} \,.$$

The following Poisson brackets are consequences of $\{\ell, K\} = n(L)$

$$\{\ell, L\} = 1 \quad , \quad \{\omega, L\} = 0 \quad , \quad \{\Omega, L\} = 0$$

We already know the integrals are commuting, and L is functionally dependent on H_p, thus the action variables (L, C, D), with $C = L\sqrt{1-e^2}$ from eq.(1.38), then $D = L\sqrt{1-e^2}\,\cos I$, are commuting

$$\{L, C\} = 0 \quad , \quad \{L, D\} = 0 \quad , \quad \{C, D\} = 0 \,.$$

In conclusion, the **Delaunay variables** $(L, C, D, \ell, \omega, \Omega)$ are action-angle variables; to complete the proof that they are canonical variables, that is the check of the symplectic condition, eq. (2.7), we need only to compute the few missing Poisson brackets.

vedi NOTA 3 alla fine del documento

$$\{\ell, \omega\} = 0 \quad , \quad \{\ell, \Omega\} = 0 \quad , \quad \{\omega, \Omega\} = 0 \,.$$

[If we assume the Arnold-Jost theorem is known, then the action-angle canonical variables exist and it is possible to adapt the standard proof, based on the generating function method, in such a way that they coincide with the Delaunay variables. This section was meant as a separate proof of this theorem for the special case of the Kepler problem. NO]

[The three zero Poisson brackets above are zero if the choice of the origins for the angle variables, which is of course arbitrary and is not dictated in an unique way by the Arnold-Jost theorem, are such that they do not depend at all from the action variables. See the proof of the canonical nature of the Delaunay variables in (Moser and Zehnder). ~~Find citation~~]

↳ «Notes on Dynamical Systems»

2.5.1 Delaunay variables with mass

The Delaunay action variables, as defined in this Section, are per unit mass of each planet. To handle a set of simultaneous 2-body problems, which is used as unperturbed Hamiltonian H_0 for a multi-body problem, we need to define variables such that the action variables have the dimension of an action, containing also the mass of the body; in this way it is possible to sum the action variables to obtain the integrals, such as the angular momentum from (1.52).

Thus we define **Delaunay variables with mass** for each planet $i = 1, N$ and asteroid $i = N+1$

$$L_i = \mu_i \sqrt{(\mu_0 + \mu_i)\, a_i} \quad , \quad C_i = L_i \sqrt{1 - e_i^2} \quad , \quad D_i = C_i \cos(I_i) \tag{2.31}$$

where a_i, e_i, I_i are the relative 2-body elements for each planet and $\mu_0 + \mu_i$ is the active central mass for the corresponding fixed center problem[7]. The angles $\lambda_i, \omega_i, \Omega_i$ are just the same defined for each 2-body subset.

The transformation for each 2-body sub-problem with index i from the Delaunay variables per unit mass to the variables with mass $(L_i, C_i, D_i, \ell_i, \omega_i, \Omega_i)$ is canonical with valence μ_i, thus for each 2-body problem with index i the corresponding unperturbed Hamiltonian, with value multiplied by μ_i, and the mean motion are

$$H_{0i} = -\frac{(\mu_0 + \mu_i)^2\, \mu_i^3}{2\, L_i^2} \tag{2.32}$$

$$n_i = \frac{\partial H_{0i}}{\partial L_i} = \frac{(\mu_0 + \mu_i)^2\, \mu_i^3}{L_i^3} = \frac{(\mu_0 + \mu_i)^2 \mu_i^3}{\mu_i^3\, (\mu_0 + \mu_i)^{3/2}\, a_i^{3/2}} = \frac{(\mu_0 + \mu_i)^{1/2}}{a_i^{3/2}} \tag{2.33}$$

as in the Third Kepler law. If we use a suitable coordinate system for the handling of all the planets, and for the removal of the center of mass, such as the Heliocentric Canonical of Section 1.6, then it is possible to express the total angular momentum (scalar and z component) as sum of the contributions from each 2-body Sun-planet subsystem[8]. The simplest consequence of this approach is that the z component of the total angular momentum is $c_z = \sum_{i=1}^{N+1} D_i$, where the $i = N+1$ term gives a negligible contribution but has to be kept to generate the appropriate symmetry involving also the asteroid; the same applies to the scalar total angular momentum $c = \sum_{i=1}^{N+1} C_i$, see Section 1.6. Note that this scaling of the action variables does not change the fact that (C_i, ω_i) and (D_i, Ω_i) are couples of conjugated variables for all $i = 1, N+1$, because the scaling applies equally to the momenta and to the 2-body Hamiltonian in each group of 6 variables referring to one planet/asteroid.

2.6 Lie series in Delaunay variables

Let us use action-angle Delaunay variables for all bodies, N planets and 1 asteroid:

$$\mathbf{L} = (L_1, \ldots, L_N, L_{N+1}) \quad , \quad \mathbf{\Theta} = (C_1, D_1, \ldots, C_N, D_N, C_{N+1}, D_{N+1})$$

$$\boldsymbol{\ell} = (\ell_1, \ldots, \ell_N, \ell_{N+1}) \quad , \quad \boldsymbol{\theta} = (\omega_1, \Omega_1, \ldots, \omega_N, \Omega_N, \omega_{N+1}, \Omega_{N+1}) \; . \tag{2.34}$$

The Hamilton function for the complete $N+2$ body problem expressed as a function of these variables is:

$$H(\mathbf{L}, \mathbf{\Theta}, \boldsymbol{\ell}, \boldsymbol{\theta}) = H_0(\mathbf{L}) + \epsilon\, H_1(\mathbf{L}, \mathbf{\Theta}, \boldsymbol{\ell}, \boldsymbol{\theta}) \; .$$

[7] For $i = N+1$, that is the asteroid, we can neglect μ_{N+1} with respect to the much larger μ_0, thus $L_{N+1} = \mu_{N+1} \sqrt{\mu_0\, a_i}$; similar simplifications occur in the formulas for the Hamiltonians.

[8] Also the Jacobian Coordinates [Milani and Gronchi 2010][Section 4.4] have this property of simplifying the formulas for the angular momentum.

The unperturbed part, the Hamiltonian of $N+1$ 2-body problems each evolving as if there were no other perturbing bodies, is just the sum (without any coupling terms):

$$H_0(\mathbf{L}) = \sum_{i=1}^{N+1} H_{0i}(L_i) = \sum_{i=1}^{N+1} \frac{-(\mu_0+\mu_i)^2\,\mu_i^3}{2\,L_i^2} \ .$$

Let also the determining function $\chi = \epsilon\,\chi_1 + \epsilon^2\,\chi_2$ be a function of the same action-angle variables, and the canonical map defined by χ

$$\Phi_\chi^1(\mathbf{L},\Theta,\ell,\theta) = (\mathbf{L}',\Theta',\ell',\theta') \ .$$

Then the transformed Hamiltonian is, to first order, from eq. (2.16), (2.17):

$$H'(\mathbf{L}',\Theta',\ell',\theta') = H_0(\mathbf{L}') + \epsilon\,[H_1(\mathbf{L}',\Theta',\ell',\theta') - \{H_0(\mathbf{L}'),\chi_1(\mathbf{L}',\Theta',\ell',\theta')\}] + O(\epsilon^2)$$

and the possible simplification of the Hamilton function in the new coordinates is described by the homological equation of order 1 in ϵ

$$H_1 - \{H_0,\chi_1\} = \overline{H}_1 \quad , \quad \{H_0,\overline{H}_1\} = 0 \ , \tag{2.35}$$

where the null Poisson bracket with $H_0 = H_0(\mathbf{L})$ implies that $\overline{H}_1 = \overline{H}_1(\mathbf{L},\Theta,\theta)$ is independent from ℓ.

To solve explicitly the homological equation (2.35) we use a Fourier series expansion of the perturbing Hamilton function H_1, where the arguments contain all the angle variables:

$$H_1 = \sum_{\mathbf{j},\mathbf{k}} H_{1\mathbf{j},\mathbf{k}}(\mathbf{L},\Theta)\cos(\mathbf{j}\cdot\ell + \mathbf{k}\cdot\theta) \tag{2.36}$$

where $\mathbf{j}\in\mathbb{Z}^{N+1}$ and $\mathbf{k}\in\mathbb{Z}^{2N+2}$ are multi-indexes. The fact that H_1 is even in the angles, thus we can use only the cosine function, will be justified later in Section 2.8.

Since χ_1 must contain terms corresponding to the ones of H_1 to be able to solve (2.35), we also expand

$$\chi_1 = \sum_{\mathbf{j},\mathbf{k}} \chi_{1\mathbf{j},\mathbf{k}}(\mathbf{L},\Theta)\sin(\mathbf{j}\cdot\ell + \mathbf{k}\cdot\theta) \tag{2.37}$$

and compute

$$\{H_0,\chi_1\} = -\sum_{i=1}^{N+1} n_i \frac{\partial\chi_1}{\partial\ell_i}$$

with $n_i = \partial H_{0i}/\partial L_i$ the mean motion of the body $i = 1, N+1$, from eq. (2.33). By using the expansion (2.37)

$$\{H_0,\chi_1\} = -\sum_{i=1}^{N+1} n_i \sum_{\mathbf{j},\mathbf{k}} j_i\,\chi_{1\mathbf{j},\mathbf{k}}\cos(\mathbf{j}\cdot\ell + \mathbf{k}\cdot\theta)$$

where j_i is the i-th coordinate of $\mathbf{j}$. In order to reduce the complexity of the transformed Hamiltonian H', as many as possible terms in this expression need to be the same as their counterparts in the expansion of H_1 from (2.36), thus term by term:

$$H_{1\mathbf{j},\mathbf{k}} = -\mathbf{n}\cdot\mathbf{j}\ , \chi_{1\mathbf{j},\mathbf{k}} \iff \chi_{1\mathbf{j},\mathbf{k}} = -\frac{H_{1\mathbf{j},\mathbf{k}}}{\mathbf{n}\cdot\mathbf{j}} \ . \tag{2.38}$$

where $\mathbf{n} = (n_1, \ldots, n_{N+1})$ is the vector of mean motions, and each scalar product (integer combination of mean motions) is a divisor.

The conclusion from (2.38) is that the only obstruction to solving for the coefficients $\chi_{1\mathbf{j},\mathbf{k}}$ can arise from zero divisors $\mathbf{n} \cdot \mathbf{j} = 0$. These can occur in two ways. First case: $\mathbf{j} = \mathbf{0}$, that is the corresponding term in H_1 is a **secular term**, not containing $\boldsymbol{\ell}$; in this case the term containing the coefficient $H_{10,\mathbf{k}}$ must be included in $\overline{H}_1$; the term $\chi_{10,\mathbf{k}}$ can be set to 0 because it does not reduce in any way the complexity of H'.

The other case of zero divisor, namely $\mathbf{n} \cdot \mathbf{j} = 0$ with $\mathbf{j} \neq \mathbf{0}$, indicates a **mean motion resonance**. This needs to be handled in special way, by keeping the resonant terms in the reduced order 1 perturbation $\overline{H}_1$, see Chapter 8.

In this chapter we shall assume that $\mathbf{n} \cdot \mathbf{j} = 0$ does not occur for $\mathbf{j} \neq \mathbf{0}$; this is a meaningful assumption if we can assume that the Fourier expansion of H_1 has just a finite number of terms, because of some limitation such as $||\mathbf{j}|| < j_{max}$, where the norm is just $||\mathbf{j}|| = \sum_{i=1}^{N+1} |j_i|$. This "ultraviolet cutoff" (truncation for too large multiples of the mean motion frequencies) is a legitimate approximation because of the fast convergence of the Fourier expansion for the smooth function H_1. Nevertheless the absence of resonances with small coefficients is a hypothesis, to be tested case by case in each region of, e.g., the asteroid belt.

Anyway, if we can assume the **non-resonance hypothesis** $\mathbf{n} \cdot \mathbf{j} = 0 \Longrightarrow \mathbf{j} = \mathbf{0}$, then

$$H_1' = \overline{H}_1 = \overline{H}_1(\mathbf{L}', \boldsymbol{\Theta}', \boldsymbol{\theta}') = \sum_{\mathbf{k}} H_{10,\mathbf{k}}(\mathbf{L}', \boldsymbol{\Theta}') \cos(\mathbf{k} \cdot \boldsymbol{\theta}') \tag{2.39}$$

a sum of pure **secular terms**, without any mean anomaly. We use also the notation $\tilde{H}_1 = H_1 - \overline{H}_1$ for the sum of terms in the expansion of H_1 which contain the fast variables $\boldsymbol{\ell}$.

Higher order perturbation theory

Having solved in this way the first homological equation, we can tackle the second one, namely (2.18), which can be somewhat simplified if we use $\{H_0, \chi_1\} = \tilde{H}_1$ from the first, and also that $H_2 = 0$ as it is indeed the case in our problem, see Section 1.6:

$$\begin{aligned} H_2' &= -\{H_1, \chi_1\} + \frac{1}{2}\{\{H_0, \chi_1\}, \chi_1\} - \{H_0, \chi_2\} = \\ &= -\left\{\overline{H}_1 + \tilde{H}_1, \chi_1\right\} + \frac{1}{2}\left\{\tilde{H}_1, \chi_1\right\} - \{H_0, \chi_2\} \\ &= -\left\{\overline{H}_1, \chi_1\right\} - \frac{1}{2}\left\{\tilde{H}_1, \chi_1\right\} - \{H_0, \chi_2\} \ . \end{aligned}$$

Pursuing the same idea of reducing the complexity of H', from this we can write the second homological equation in the form

$$K_2 = \left\{\tilde{H}_1, \chi_1\right\} \quad , \quad G_2 = -\left\{\overline{H}_1, \chi_1\right\} - \frac{1}{2}\tilde{K}_2 = \tilde{G}_2 \quad , \quad \tilde{G}_2 - \{H_0, \chi_2\} = 0 \tag{2.40}$$

where the second order reduced form $H_2' = \overline{K_2}$, contains again only secular terms. They can only arise from beats between non-secular terms contained in $\tilde{H}_1$ and χ_1. To the contrary,

$\{\overline{H}_1, \chi_1\}$ cannot contain secular terms, because all derivatives of $\overline{H}_1$ do not contain ℓ and all derivatives of χ_1 do, thus no beat term without ℓ can occur. Thus

$$H'_2 = \overline{K_2}(\mathbf{L}', \Theta', \theta') = \overline{-\frac{1}{2}\left\{\tilde{H}_1, \chi_1\right\}} ,$$

meaning that H'_2 is just the collection of the secular terms of K_2. All of G_2, that is all non-secular terms, can be removed by composing χ_2 with sine terms of the same arguments contained in the terms of $\tilde{G}_2 = G_2$, and with the same divisors as in the solution of the first homological equation:

$$G_{2\mathbf{j},\mathbf{k}} = -\mathbf{n}\cdot\mathbf{j}\,\chi_{2\mathbf{j},\mathbf{k}} \Longleftrightarrow \chi_{2\mathbf{j},\mathbf{k}} = -\frac{G_{2\mathbf{j},\mathbf{k}}}{\mathbf{n}\cdot\mathbf{j}} . \tag{2.41}$$

It would be possible to prove by recursion that the same procedure could be repeated up to an arbitrary order in ϵ, thus pushing the transformation of H into a Hamilton function in the new variables H' to an arbitrary order:

$$H' = H_0 + \epsilon\, H'_1 + \epsilon^2\, H'_2 + \ldots + \epsilon^r\, H'_r + O(\epsilon^{r+1}) .$$

However, before attempting to push these computations to a higher order we need a deeper understanding of their significance and computational value.

2.7 Non-singular variables, Poincaré domain

The solution of the homological equations such as (2.17) and (2.18) can be computed term by term in the Fourier series. However, if we wish to understand which, among the terms of the same order in ϵ, are the most important, then we need to introduce some small parameters, such as powers of e, I. To do this we need to be able to expand in the neighborhood of circular orbits $e = 0$ and planar direct orbits $I = 0°$, that is to consider a **Poincaré domain** which adds $e = 0$ and $I = 0°$ orbits to the Delaunay domain without giving rise to singularities in the equations of motion. Thus we need to define coordinates for the Poincaré domain which are smooth on both $e = 0$ and $I = 0°$. Note that planar retrograde orbits $I = 180°$ are not included: the definition of the Poincaré domain is $e < 1, I < 180°$.

To obtain such coordinates we proceed in two steps. First, starting from Delaunay variables $(L, C, D, \ell, \omega, \Omega)$ we perform a global canonical transformation such that the angle variables go into angle variables (that is, the 3-tori $V(E, C, D)$ are mapped into themselves by a smooth diffeomorphism), and such that the small quantities $L - C$ and $C - D$ are among the new action variables. Note that $L - C = L\left[1 - \sqrt{1 - e^2}\right] = L\,\mathcal{O}(e^2)$ and $C - D = C\,[1 - \cos I] = C\,\mathcal{O}(sin^2 I)$.

As shown in the examples of Section 2.1, if A is a 3×3 matrix then the following is locally a canonical transformation

$$\begin{bmatrix} \Lambda \\ V \\ Z \end{bmatrix} = A \begin{bmatrix} L \\ C \\ D \end{bmatrix} \quad , \quad \begin{bmatrix} \lambda \\ v \\ z \end{bmatrix} = \left[A^{-1}\right]^T \begin{bmatrix} \ell \\ \omega \\ \Omega \end{bmatrix} ;$$

for it to be a global canonical transformation both A and A^{-1} need to be matrices with integer coefficients. Thus A must be a matrix with integer coefficients and[9] $det\, A = 1$. We select

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \quad , \quad \left[A^{-1}\right]^T = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

that is we define new action and angle variables

$$\begin{bmatrix} \Lambda \\ V \\ Z \end{bmatrix} = \begin{bmatrix} L \\ C-L \\ D-C \end{bmatrix} \quad , \quad \begin{bmatrix} \lambda \\ v \\ z \end{bmatrix} = \begin{bmatrix} \ell+\omega+\Omega \\ \omega+\Omega \\ \Omega \end{bmatrix} . \tag{2.42}$$

The variables: λ, the **mean longitude** and $v = \varpi$ **longitude of pericenter** are "broken-legged", that is sum of angles which are in different planes, but as we will see they can be continuously extended to $e = 0, I = 0°$. Of course $z = \Omega$ is the longitude of ascending node.

In the second step we define "Cartesian like" canonical variables

$$\begin{aligned} \xi &= \sqrt{-2V}\sin v = \sqrt{2L}\sqrt{1-\sqrt{1-e^2}}\,\sin\varpi = \sqrt{2L}\,\sin\varpi\,\mathcal{O}(e) \\ \eta &= \sqrt{-2V}\cos v = \sqrt{2L}\sqrt{1-\sqrt{1-e^2}}\,\cos\varpi = \sqrt{2L}\,\cos\varpi\,\mathcal{O}(e) \\ \alpha &= \sqrt{-2Z}\sin z = \sqrt{2L}\sqrt{\sqrt{1-e^2}\,(1-\cos I)}\,\sin\Omega = \sqrt{2L}\,\sin\Omega\,\mathcal{O}(\sin I) \\ \beta &= \sqrt{-2Z}\cos z = \sqrt{2L}\sqrt{\sqrt{1-e^2}\,(1-\cos I)}\,\cos\Omega = \sqrt{2L}\,\cos\Omega\,\mathcal{O}(\sin I)\ , \end{aligned} \tag{2.43}$$

in this way we have obtained variables which, unlike Delaunay one, can be used even when $\xi = \eta = 0$, that is $e = 0$ and/or $\alpha = \beta = 0$, that is $I = 0°$. We need to prove that the coordinates $(\Lambda, \xi, \alpha, \lambda, \eta, \beta)$ are actually smooth functions of the Cartesian coordinates of the position and velocity over the entire Poincaré domain. The simplest way is to use as intermediaries the **equinoctial elements** [Broucke and Cefola, 1972]. → «On the equinoctial orbital elements»

$$(a, \lambda, h, k, p, q) \quad with \quad h = e\sin\varpi\ ,\ k = e\cos\varpi\ ,\ p = \tan(I/2)\sin\Omega\ ,\ q = \tan(I/2)\cos\Omega$$

where a, λ are semimajor axis and mean longitude as previously defined. Of course $\Lambda = L = \sqrt{\mu\, a}$ is a smooth function of a and the variable λ is the same in both coordinate systems. Since

$$e^2 = h^2 + k^2 \quad , \quad \tan^2(I/2) = p^2 + q^2 \quad , \quad \cos I = \frac{1-p^2-q^2}{1+p^2+q^2}$$

the elements $(\xi, \eta, \alpha, \beta)$ can be computed as a smooth function of (h, k, p, q). By expanding in $e = h = k = 0$

$$\xi = [\mu\, a]^{1/4}\ h + \mathcal{O}(h^2+k^2) \quad , \quad \eta = [\mu\, a]^{1/4}\ k + \mathcal{O}(h^2+k^2) \quad *$$

and by expanding in $I = p = q = 0$

$$\alpha = \sqrt{2}\,[\mu\, a]^{1/4}\,\left[1-h^2-k^2\right]^{1/4}\sqrt{\frac{2(p^2+q^2)}{1+p^2+q^2}}\,\sin\Omega = 2\,[\mu\, a]^{1/4}\,\left[1-h^2-k^2\right]^{1/4}\,p + \mathcal{O}(p^2+q^2)$$

[9] Actually, $det\, A = -1$ would still define a canonical transformation, but we prefer to use a matrix not changing the orientation.

$$\frac{1}{\sqrt{1+x}} = 1 + O(x)$$

* basta notare che $\sqrt{1-\sqrt{1-e^2}} = \frac{e}{\sqrt{1+\sqrt{1-e^2}}}$ e che $\frac{1}{\sqrt{1+\sqrt{1-x}}} = \frac{1}{\sqrt{2}} + O(x)$

and similarly

$$\beta = 2\,[\mu\,a]^{1/4}\,\left[1-h^2-k^2\right]^{1/4}\,q + \mathcal{O}(p^2+q^2)$$

we find that the map is invertible even where $e = 0$ and/or $I = 0$.

Thus it is enough to prove that the equinoctial elements are obtained from the Cartesian position and velocity with a diffeomorphism smooth on the Poincaré domain. This is a classical result from [Broucke and Cefola, 1972]; indeed the only singularities occur for $e = 1$ and/or $I = 180°$. On the other hand the map from Cartesian position and velocity to Delaunay variables is canonical on the Delaunay domain, and the same applies to the map from Delaunay to Poincaré elements. Note that the extension of both maps to the Poincaré domain, once proven to exist and be smooth, is also canonical because the Poisson brackets are continuous functions and the equations proving the map is canonical can be continued to $e = 0$ and/or $I = 0$.

2.8 D'Alembert rules

Now we can exploit both arguments, the existence of commuting integrals (from Section 2.5) and the expansions in Taylor series around $e = 0$ and $I = 0$ (from Section 2.7) to constrain the expansion of the perturbing Hamiltonian H_1 in a Fourier series: these constraints are the **D'Alembert rules**.

First rule

The first D'Alembert rule arises from the fact that the perturbing Hamiltonian must be invariant with respect to the reference system change, for example, if we turn the x, y axes by the same angle s around the z axis. In other words, it arises from the 1-parameter group of symmetries generated by the z-component of the angular momentum, that is the Delaunay variable D. Of course, D and Ω are conjugated variables, that is the action of D (as determining function) on a 2-body problem expressed in Delaunay variables is just a translation, with frequency 1, of the variable Ω with all other variables being fixed:

$$\Phi_D^s(L, C, D, \ell, \omega, \Omega) = (L, C, D, \ell, \omega, \Omega + s)\ .$$

When the same action of D is applied to the same 2-body problem, but as expressed in the variables $(\Lambda, V, Z, \lambda, \varpi, \Omega)$, then from the definition of these variables (2.42) the one parameter group changes all the angles in the same way:

$$\Phi_D^s(\Lambda, V, Z, \lambda, \varpi, \Omega) = (\Lambda, V, Z, \lambda + s, \varpi + s, \Omega + s)\ .$$

On the other hand, $\{H, c_z\} = \{H_1, c_z\} = 0$, where c_z is the z component of the total angular momentum of the system, thus the expansion of H_1 in a Fourier series is constrained. Let us suppose the expansion (2.36) has been converted to the variables $(\Lambda_i, V_i, Z_i, \lambda_i, \varpi_i, \Omega_i)$, one for each planet and for the asteroid, $i = 1, N+1$;

$$H_1 = \sum_{\mathbf{j},\mathbf{f},\mathbf{g}} H_{1\mathbf{j},\mathbf{f},\mathbf{g}}(\mathbf{\Lambda}, \mathbf{V}, \mathbf{Z})\,\cos\left(\mathbf{j}\cdot\boldsymbol{\lambda} + \mathbf{f}\cdot\boldsymbol{\varpi} + \mathbf{g}\cdot\boldsymbol{\Omega}\right)\ .$$

where the vectors $\mathbf{\Lambda}, \mathbf{V}, \mathbf{Z}, \boldsymbol{\lambda}, \boldsymbol{\omega}, \boldsymbol{\Omega}$ contain the corresponding variable for all $N+1$ planets+ asteroid. Then the action of the determining function $c_z = \sum_{i=1}^{N+1} D_i$ on this expansion is

$$H_1 \circ \Phi_{c_z}^s = \sum_{\mathbf{j},\mathbf{f},\mathbf{g}} H_{1\mathbf{j},\mathbf{f},\mathbf{g}}(\mathbf{\Lambda}, \mathbf{V}, \mathbf{Z}) \cos\left(\mathbf{j}\cdot\boldsymbol{\lambda} + \mathbf{f}\cdot\boldsymbol{\varpi} + \mathbf{g}\cdot\boldsymbol{\Omega} + s\sum_{i=1}^{N+1}(j_i + f_i + g_i)\right)$$

because all angles for all planets (and for the asteroid) are changed by the same amount s. Since the derivative of the cosine term with respect to its argument is generally not zero, the only possibility for preserving the value of H_1 is that the additional part of the argument is zero, for all the non-zero Fourier terms:

$$H_{1\mathbf{j},\mathbf{f},\mathbf{g}} \neq 0 \Longleftrightarrow \sum_{i=1}^{N+1}(j_i + f_i + g_i) = 0 \ ; \tag{2.44}$$

this is the **first D'Alembert rule**, stating that the integer coefficients appearing in the angular arguments of the Fourier series expansion have a zero sum.

Second rule

The second D'Alembert rule arises from the symmetry of the Hamilton function with respect to mirror transformations: they are also isometries. Let us use the mirror symmetry with respect to the (x, y) plane, that is $\Sigma : (x, y, z) \mapsto (x, y, -z)$. Such a transformation implies that the orbital nodes exchange their places - the ascending one becomes descending, and conversely; reckoned from a fixed direction, the longitude of ascending node, Ω, changes by π. The argument of pericenter, ω, reckoned from the new ascending node, changes by the same amount. When this transformation is written in Keplerian elements, it takes the form

$$\Sigma : (a, e, I, \ell, \omega, \Omega) \mapsto (a, e, I, \ell, \omega + \pi, \Omega + \pi)$$

To prove the statement above, we can write the Cartesian coordinates of a Keplerian orbit as a function of the elements by means of the rotation matrix with Euler angles (Ω, I, ω)

$$\begin{cases} x &= r\left[\cos(v+\omega)\cos\Omega - \sin(v+\omega)\sin\Omega\cos I\right] \\ y &= r\left[\cos(v+\omega)\sin\Omega + \sin(v+\omega)\cos\Omega\cos I\right] \\ z &= r\sin(v+\omega)\sin I \end{cases} , \tag{2.45}$$

where the true anomaly is a function of the mean anomaly $v = v(\ell)$, both invariant, since reckoned from the pericenter direction. By adding π to both ω and Ω there are two changes in sign in x, y, which remain unchanged, and one in z, changing it to $-z$. Of course the same transformation to angles applies in Delaunay variables

$$\Sigma : (L, C, D, \ell, \omega, \Omega) \mapsto (L, C, D, \ell, \omega + \pi, \Omega + \pi)$$

and when converted to the non-singular angles

$$\Sigma : (\Lambda, V, Z, \lambda, \varpi, \Omega) \mapsto (\Lambda, V, Z, \lambda + 2\pi, \varpi + 2\pi, \Omega + \pi) \ ,$$

thus modulo 2π all the angles are conserved, but $\Omega \mapsto \Omega + \pi$. Then a term in H_1 to be invariant must be such that

$$H_1 \circ \Sigma = \sum_{\mathbf{j},\mathbf{f},\mathbf{g}} H_{1\mathbf{j},\mathbf{f},\mathbf{g}}(\mathbf{\Lambda}, \mathbf{V}, \mathbf{Z}) \cos\left(\mathbf{j}\cdot\boldsymbol{\lambda} + \mathbf{f}\cdot\boldsymbol{\varpi} + \mathbf{g}\cdot\boldsymbol{\Omega} + \pi\sum_{i=1}^{N+1} g_i\right) = H_1 \ .$$

This implies that the additional term in the argument must contain an even number of π, that is

$$H_{1\mathbf{j},\mathbf{f},\mathbf{g}} \neq 0 \Longleftrightarrow \sum_{i=1}^{N+1} g_i \equiv 0 (mod\ 2) \ , \tag{2.46}$$

and this also implies that the sum of integer coefficients standing with Ω_i in the arguments of the non-zero perturbing terms must be even. This is the **second D'Alembert rule.**

The Hamiltonian is even in the angles

The next rule[10] is implied by the possibility of a symmetry with respect to a line, namely $\Sigma_x : (x, y, z) \mapsto (x, -y, -z)$. In Delaunay variables

$$\Sigma_x : (L, C, D, \ell, \omega, \Omega) \mapsto (L, C, D, -\ell, -\omega, -\Omega)$$

all angles change in sign, and in the non-singular angles the same:

$$\Sigma : (\Lambda, V, Z, \lambda, \varpi, \Omega) \mapsto (\Lambda, V, Z, -\lambda, -\varpi, -\Omega) \ .$$

The proof is as above, from the Cartesian position as a function of the elements, eq. (2.45): the change of sign in all angles implies the change of sign in y, z, containing one sine (with angle variables as arguments)[11] in all terms, but not in x, containing only cosines and even number of sines in all terms. In order for the Hamilton function to be symmetric with respect to a line, each term in the Fourier series for H_1 must be even in the argument (i.e. $f(x) = f(-x)$), and the series can thus contain only cosines, not sines[12].

As for the determining function $\chi = \epsilon\chi_1 + \epsilon^2\chi_2 + \ldots$, when used to solve the homological equation, the first order part χ_1 is obtained from a derivative, with respect to some angle variable, of H_1, thus it must contain only sines in place of cosines of the arguments, and it must be odd in the angles; the same for χ_2 and so on.

Third rule

The third D'Alembert rule arises from the regularity of the Hamilton function, as expressed in the Poincaré variables, on the points corresponding to $e = 0, I = 0$ (for each Λ, λ): then it must be possible to expand it in a Taylor series either in the variables ξ, η, α, β, centered at $\xi = \eta = \alpha = \beta = 0$, or equivalently, in the variables h, k, p, q at $h = k = p = q = 0$. However, not all monomials in these variables are allowed, because of (2.44), (2.46), and the cosine rule.

[10] Some authors, e.g., [Morbidelli, 2002], consider this also as a D'Alembert rule, thus he lists four of them.

[11] Note that I is not an angle variable.

[12] This is not considered a D'Alembert rule in the literature, but it is analogous to the second one.

As the simplest case, let us consider secular terms like those in (2.39): for these, from the first and second D'Alembert rules:

$$\mathbf{j} = \mathbf{0} \quad , \quad \sum_{i=1}^{N+1} k_{2i} \equiv 0 (mod\, 2) \quad , \quad \sum_{i=1}^{N+1} k_{2i-1} = 0 (mod\, 2) \quad , \quad \sum_{i=1}^{2N+2} k_i = 0$$

because, according to (2.34), the even secular angles are the longitudes of node Ω_i and the odd ones are the arguments of perihelion ω_i. As an example, the arguments which can be expressed as combinations of basic arguments $\varpi_i - \varpi_l$ and $\Omega_i - \Omega_l$ are allowed, and the corresponding basic expressions, appearing in the Hamiltonian, in non-singular elements are

$$e_i\, e_l\, \cos(\varpi_i - \varpi_l) = h_i h_l + k_i k_l \quad , \quad \tan(I_i/2)\, \tan(I_l/2)\, \cos(\Omega_i - \Omega_l) = p_i p_l + q_i q_l \ ,$$

and similar expressions in Poincaré elements. In the portion of the secular Hamiltonian $\overline{H}_1$ of homogeneous degree 2 in h_i, k_i, p_i, q_i (also in the canonical $\xi_i, \eta_i, \alpha_i, \beta_i$) there are no other possible terms, because only two angles are allowed and the combinations like $\varpi_i - \Omega_l$ are forbidden by the second D'Alembert rule.

To extend this analysis to an arbitrary degree homogeneous polynomial in the non-singular variables we need to use the representation of

$$\cos(\mathbf{f} \cdot \boldsymbol{\varpi} + \mathbf{g} \cdot \boldsymbol{\Omega}) = P(\boldsymbol{\varpi}, \boldsymbol{\Omega})$$

as a trigonometric polynomial containing sines and cosines of all the angles. It is easy to prove by recursion that P is a homogeneous trigonometric polynomial of degree

$$s = \sum_{i=1}^{2N+2} (|k_i|) \ ,$$

with each monomial containing $|k_{2i-1}|$ sines/cosines of ϖ_i and $|k_{2i}|$ sines/cosines of Ω_i, for each i. From this we can deduce that the term containing $P(\boldsymbol{\varpi}, \boldsymbol{\Omega})$, when expressed in non-singular variables, must contain a homogeneous polynomial in h_i, k_i, p_i, q_i of at least the same degree s. There are further constraints on the powers of h_i, k_i and p_i, q_i contained in each monomial: e.g., the powers of either h_i of k_i must add to at least $|k_{2i-1}|$, the the powers of either p_i of q_i must add to at least $|k_{2i}|$. This is the secular version of the **third D'Alembert rule**.

[Can we give a set of exhaustive rules? that is, each terms consistent with the rules actually exists?]

As an example, a prominent term of degree 4 contains

$$e_i^2\, \tan^2(I_i/2)\, \cos(2\omega_i) = e_i^2\, \tan^2(I_i/2)\, \cos(2\varpi_i - 2\Omega_i) = (h_i^2 - k_i^2)\, (p_i^2 - q_i^2) \ ,$$

which is not a product of the expressions appearing in the degree 2 portion: the argument is a combination of $\varpi_i + \varpi_i$ and $\Omega_i + \Omega_i$, with the opposite sign.

In particular the third D'Alembert rule implies that there are no terms in $\overline{H}_1$ of degree odd in the non-singular variables, because they would correspond to trigonometric terms with odd s, which do not satisfy the first D'Alembert rule.

[How to define exactly the functions generating the algebraic functions of h_i, k_i, p_i, q_i allowed by all the D'Alembert rules?]

The third D'Alembert rule for the terms containing the mean longitudes is based on the definition of **order of resonance**, which is just $|\mathbf{j}| = \sum_i j_i$ (note that it can be assumed to be positive, given that the argument appears in a cosine term). Then $|\mathbf{f}| = \sum_{i=1}^{N+1} k_{2i}$ must be even by (2.46) and $|\mathbf{g}| = \sum_{i=1}^{N+1} k_{2i-1}$ is such that $|\mathbf{j}|+|\mathbf{f}| = -|\mathbf{g}|$. From this we can deduce that the monomials in h_i, k_i, p_i, q_i have at least degree $|\mathbf{j}|$ when appearing as factors of $\cos(\mathbf{j}\cdot\boldsymbol{\lambda})$.

There are even more strict rules, that is the argument $k_i\,\varpi_i$ cannot appear without $e_i^{k_i}$ in front of the cosine, the same with $h_i\Omega_i$ and $\tan^{h_i}(I_i/2)$.

Thus, as the simplest example, the terms corresponding to a 2/1 resonance between the asteroid and planet i of the lowest degree contain a factor either $e\,\cos(\lambda-2\lambda_i+\varpi)$ or $e_i\cos(\lambda-2\lambda_i+\varpi_i)$; the nodes Ω, Ω_i are not allowed to appear alone, thus there is no term containing the first powers of I, I_i. (Here we are using a notation in which the index $N+1$ is dropped, and the elements without indexes are those of the asteroid).

Decomposition of the perturbation in 2-body portions

One additional rule on the terms occurring in the expansion of H_1 depends upon the superposition principle for the gravitational forces, which implies the gravitational potential is just a sum of 2-body terms. Given the particular form assumed by the indirect perturbation function in the case of Heliocentric canonical coordinates, see (1.51), the same applies to the indirect portion of H_1. Then the terms contained in H_1 are all of the form

$$H_{1j_ij_rf_if_rg_ig_r}(\Lambda_i,\Lambda_r,V_i,V_r,Z_i,Z_r)\,\cos\left(j_i\lambda_i+j_r\lambda_r+f_i\varpi_i+f_r\varpi_r+g_i\Omega_i+g_r\Omega_r\right)$$

where $1\le i<r\le N+1$ are the two bodies whose interaction is considered (note that this rule applies not just to U_1 but also to T_1). It follows that there cannot be terms corresponding to 3-body resonances, with $j_i\lambda_i+j_r\lambda_r+j_v\lambda_v$ in the argument, in the first order Hamiltonian H_1, and because of the homological equation such terms cannot arise in the first order perturbations. But of course at the second order in ϵ such terms with three mean longitudes can appear, as a result of Poisson brackets between terms with two longitudes.

[Why so many secular terms of degree 4 with frequencies from 2 planets occur in the secular terms of our own theory? There are even terms with fundamental frequencies from 3 planets.]

2.9 Mean elements and mean semimajor axis

The **mean elements** could be defined by a canonical transformation; e.g., by using the non-singular action-angle variables, a canonical map

$$\Psi : (\boldsymbol{\Lambda},\mathbf{V},\mathbf{Z},\boldsymbol{\lambda},\mathbf{v},\mathbf{z}) \mapsto (\boldsymbol{\Lambda}',\mathbf{V}',\mathbf{Z}',\boldsymbol{\lambda}',\mathbf{v}',\mathbf{z}')$$

such that

$$H\circ\Psi^{-1}(\boldsymbol{\Lambda}',\mathbf{V}',\mathbf{Z}',\boldsymbol{\lambda}',\mathbf{v}',\mathbf{z}') = H'(\boldsymbol{\Lambda}',\mathbf{V}',\mathbf{Z}',\mathbf{v}',\mathbf{z}') + \mathcal{O}(\epsilon^2)$$

would define mean elements to order 1. Because the new Hamilton function H' does not contain the fast variables $\boldsymbol{\lambda}'$, the evolution in time of the mean elements would not contain, at order

1 in ϵ, any short periodic perturbation resulting from terms with mean longitudes (that is, containing mean anomalies).

The question is whether this could be defined by a Lie series. This question has two different answers depending upon the context.

If the idea was to achieve a total removal of the short periodic terms, containing combination of mean longitudes $\mathbf{j} \cdot \boldsymbol{\lambda}'$ for all values of $|\mathbf{j}|$, then the corresponding Lie series needs to have an infinite number of terms for each order and the problem would be whether the series converge, both as Fourier series for each order and as a power series in ϵ. As we will discuss later, this would not be the case unless the original problem is integrable, that is admitting a set of canonical action-angle variables.

If the idea was to remove all the terms whose size, as estimated from D'Alembert third rule*, is below an assigned order of magnitude, then a series expansion for χ containing a finite number of terms would be needed. A classical way to do this is to assume that there is an integer r such that $\mathcal{O}(e_i^r) = \mathcal{O}(\sin^r I_i) = \mathcal{O}(\epsilon)$; e.g., for low e, I we can assume $r = 2$, for example

$$\epsilon\, e^2 \cos(\lambda - 3\lambda_i + 2\varpi) = \mathcal{O}(\epsilon^2) \quad , \quad \epsilon\, e^2 \sin^2 I \cos(2\varpi - 2\Omega) = \mathcal{O}(\epsilon^3) \ .$$

This assumption is often used by the classical authors in panetary dynamics. However, $\epsilon \simeq \mu_5/\mu_0 \simeq 10^{-3}$, thus $e < \sqrt{\epsilon} \simeq 0.03$ is appropriate only for a small portion of the asteroid belt.

Thus [Milani and Knežević, 1990] use implicitly $r = 4$ and select all the terms in the perturbing Hamiltonian H_1 with order of resonance ≤ 4 for the direct portion U_{DIR} and ≤ 3 for the indirect portion U_{IND}. Thus for each couple of mutually perturbing bodies they compute 189 direct and 62 indirect terms[13]. They perform these computations in Delaunay variables, and limited to order 1 in ϵ, that is they use only a determining function $\epsilon\,\chi_1$ expanded as in (2.37).

With this limited accuracy approach (see Section 9.1), and with the methods of expansion discussed in Chapter 3, we can explicitly compute a finite number of coefficients $\chi_{1\mathbf{j},\mathbf{k}}(\mathbf{L}, \boldsymbol{\Theta})$ for values of the action variables corresponding to the ordinary instantaneous[14] orbital elements. Then the transformed mean elements are computed as in (2.11), note the + sign in the first order term [15]:

$$\begin{aligned} \mathbf{L}' &= \mathbf{L} + \epsilon\, \{\mathbf{L}, \chi_1\} = \mathbf{L} - \epsilon \frac{\partial \chi_1}{\partial \boldsymbol{\ell}} \\ \boldsymbol{\Theta}' &= \boldsymbol{\Theta} + \epsilon\, \{\boldsymbol{\Theta}, \chi_1\} = \boldsymbol{\Theta} - \epsilon \frac{\partial \chi_1}{\partial \boldsymbol{\theta}} \\ \boldsymbol{\theta}' &= \boldsymbol{\theta} + \epsilon\, \{\boldsymbol{\theta}, \chi_1\} = \boldsymbol{\theta} + \epsilon \frac{\partial \chi_1}{\partial \boldsymbol{\Theta}} \ , \end{aligned}$$

where by χ_1 we indicate just the sum of the finite number of terms which are available from the theory. Then the new Hamilton function in the space of the mean elements is $H' = H_0 + \epsilon\, [H_1 - \{H_o, \chi_1\}]$ from which the "large" short periodic terms, containing the fast variables $\boldsymbol{\ell}'$, have been removed. Then it is possible to approximate the dynamics of the mean elements

[13] In Knežević (1993) the indirect part has been expanded to $r = 4$, to include 192 indirect terms. Some direct terms were also added to end up with 197 of these in the final version of the theory.

[14] Usually called osculating, but this expression does not actually apply to heliocentric canonical elements.

[15] The + sign is due to the use of the direct map to compute mean elements from instantaneous ones.

guardare la Bibliografia in fondo al file (anche per gli altri riferimenti)

$z' = z + s\{z, \chi\} + \frac{1}{2} s^2 \{\{z, \chi\}, \chi\} + O(s^3)$

$\chi_1 = \sum_{j,k} \chi_{1jk}(L, \Theta) \operatorname{sen}(j \cdot \ell + k \cdot \theta)$

* il termine di grado minimo nello sviluppo in serie di Taylor del coefficiente dell'argomento $j_1\lambda' + j_2\lambda + j_3\varpi' + j_4\varpi + j_5\Omega' + j_6\Omega$ è uguale al prodotto $e'^{|j_3|}\, e^{|j_4|}\, s'^{|j_5|}\, s^{|j_6|}$

Inoltre $|j_3| + |j_4| + |j_5| + |j_6| \geq |j_1 + j_2|$

by using (2.39) and the truncated Hamiltonian H' as a function of the mean elements

$$H' = H'_0(\mathbf{L}') + \epsilon\, H'_1(\mathbf{L}', \mathbf{\Theta}', \boldsymbol{\theta}') = H_0(\mathbf{L}') + \epsilon\, \overline{H}_1(\mathbf{L}', \mathbf{\Theta}', \boldsymbol{\theta}') = \tag{2.47}$$

$$= H_0(\mathbf{L}') + \epsilon \sum_{\mathbf{k}} H_{10,\mathbf{k}}(\mathbf{L}', \mathbf{\Theta}')\, \cos(\mathbf{k}\cdot\boldsymbol{\theta}')\ , \tag{2.48}$$

where the summation is limited again to a finite number of terms explicitly computed. This method is conceptually simple (although the computation of the coefficients is far from trivial, see Chapter 3), and provides mean elements which, at least in the simplified dynamics defined by the Hamiltonian (2.48), have $\mathbf{L}'$ as integrals. Thus the corresponding semimajor axes a', a'_i are **mean semimajor axes**, integrals of an approximation of the $N+2$ body problem, and it is reasonable to assume that these **quasi-integrals** shall change slowly with time. In this approach, it is not even necessary to compute the "mean mean anomalies" $\boldsymbol{\ell}'$ because they are not used in the following computations based upon mean elements.

2.10 Secular perturbations in semimajor axis

The simple, first order and truncated, method to compute mean elements of the previous section is anyway a step in the right direction, and indeed for a large portion of the asteroid main belt it leads to a useful approximation. However, it is clear that it contains many simplifications and therefore can lead to poor approximations in some cases.

The main cause of these unsatisfactory results are mean motion resonances. Let us take as an example the 2/1 mean motion resonance with Jupiter, with main H_1 term

$$H_{2/1} = \frac{\mu_5}{\mu_0}\, g(a, a_5)\, e\, \cos(\lambda - 2\lambda_5 + \varpi)\ ,$$

where $\mu_5/\mu_0 \simeq 10^{-3}$ is the mass of Jupiter in solar masses, and $g(a, a_5)$ is a function of the semimajor axes. By assuming $e > e_5$ this term is larger than the similar one containing ϖ_5. In Delaunay-like variables the argument of the trigonometric term is $\ell - 2\ell_5 + 2(\omega - \omega_5) + 2(\Omega - \Omega_5)$. The corresponding term in $\epsilon\,\chi_1$

$$\chi_{2/1} = \frac{\mu_5\, g(a, a_5)\, e}{n - 2n_5} \sin(\lambda - 2\lambda_5 + \varpi)\ . \tag{2.49}$$

Then the mean element Λ can be computed with the addition of a single selected second order term by using a χ_1 term computed as in eq. (2.38):

$$\Lambda' = \Lambda - \epsilon \frac{\partial \chi_1}{\partial \lambda} - \frac{1}{2}\left\{\frac{\partial \chi_{2/1}}{\partial \lambda}, \chi_{2/1}\right\} + \dots$$

$$= \Lambda - \epsilon \frac{\partial \chi_1}{\partial \lambda} - \frac{1}{2}\,\mu_5^2\, e^2 \left\{\frac{g(a, a_5)}{n - 2n_5}\, \cos(\lambda - 2\lambda_5 + \varpi), \frac{g(a, a_5)}{n - 2n_5}\, \sin(\lambda - 2\lambda_5 + \varpi)\right\} + \dots\ .$$

To compute the main portion of the Poisson bracket in the formula above we can neglect the part containing derivatives with respect to e, ϖ because they contain the denominator $(n - 2n_5)^2$, while the portion containing derivatives with respect to Λ (that is a) has a much smaller denominator $(n - 2n_5)^3$.

We need to take into account that the important contribution to the order of magnitude of the second order correction arises from the divisor $n - 2n_5$: indeed, in the neighborhood of the 2/1 resonance it can be $\mathcal{O}(\sqrt{\epsilon})$. In computing the Poisson bracket of the formula above, the terms with derivatives with respect to Λ and Λ contain the inverse cube of the divisor and either $\sin^2(\lambda - 2\lambda_5 + \varpi)$ or $\cos^2(\lambda - 2\lambda_5 + \varpi)$. When averaging over the argumentwe we get secular terms with mean 1/2, with amplitude enhanced by the inverse cube of a small quantity.

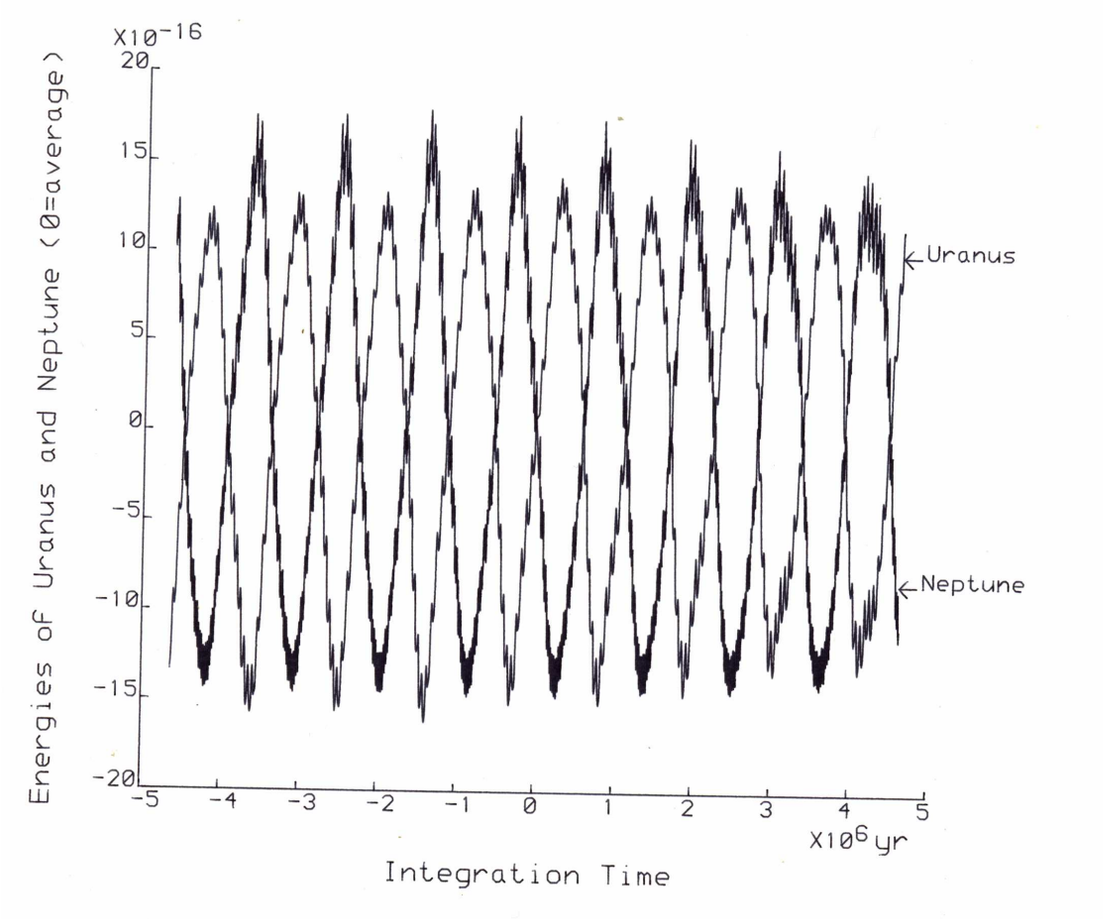


Figure 2.1: Energy of Uranus and Neptune over the 10 million years of the LONGSTOP 1A numerical integration (performed in 1984).

We would like to discuss two examples of secular perturbations on the semimajor axes, which remain in the mean elements after removing only the first order short periodic terms. The first one is amplified by the near-resonance 2/1 of the mean motions of Uranus and Neptune. The small divisor $n_7 - 2\,n_8$ is about 51 times smaller than n_7, while $\mu_8/\mu_0 = 4.4 \times 10^{-5}$, thus the small parameter

$$\left[\frac{n_7}{n_7 - 2\,n_8}\right]^3 \left[\frac{\mu_8}{\mu_0}\right]^2 \simeq 2.5 \times 10^{-4}$$

has to be multiplied by e_7^2 with a mean value of 2.3×10^{-3} and, by using a full computation of the quantities appearing in the second order effect due to the main term, [Milani *et al.* 1987b] have succeeded in predicting, although only approximately because of a simplified computation with very few terms, the values of the corresponding spectral line with frequency $g_7 - g_8$ in the semimajor axis of Uranus, which is $3.7 \times 10^{-6}\,a_7$ (the other coefficients result in an increase by

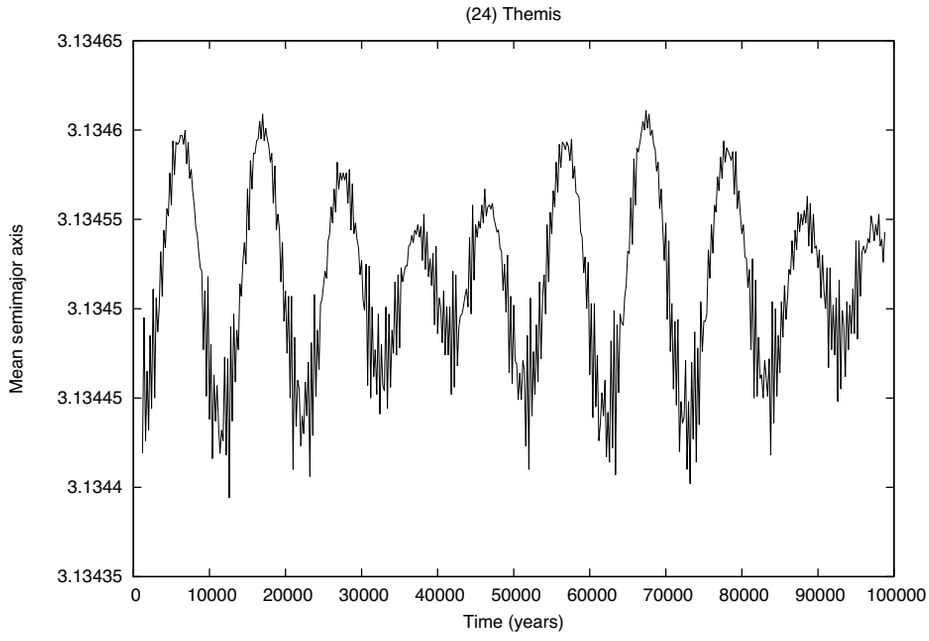


Figure 2.2: Mean semimajor axis, computed by digital filtering, for the asteroid (24) Themis, as a function of time over 100 000 years.

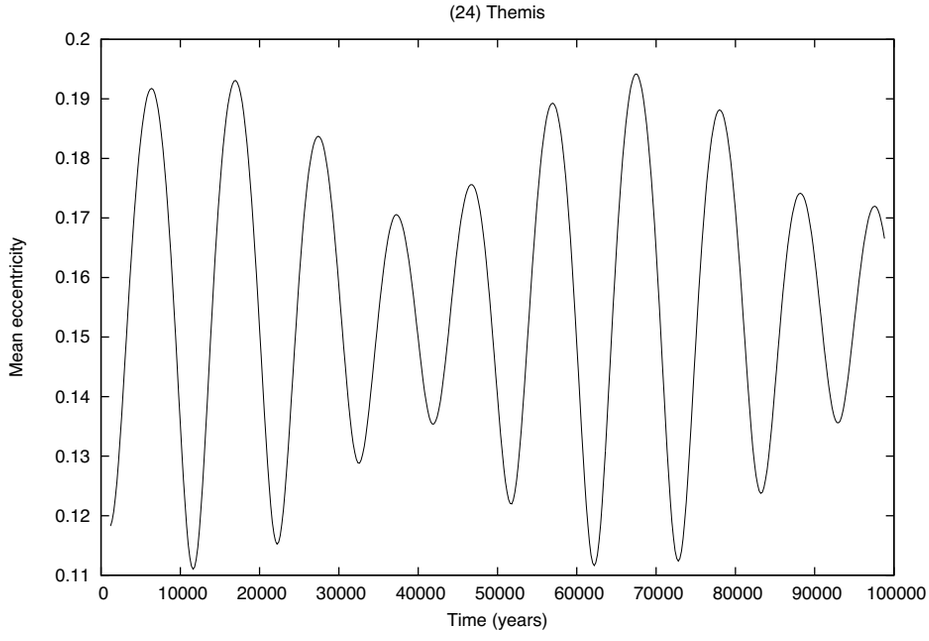


Figure 2.3: Mean eccentricity, computed by digital filtering, for the asteroid (24) Themis, as a function of time over 100 000 years.

about an order of magnitude). Figure 2.1 shows the exchange in 2-body energy between M_7E_7 and M_8E_8 as obtained by the 10 Million years integration LONGSTOP 1A [Milani *et al.* 1987b].

The second example is the secular change of the mean semimajor axis of the asteroid (24) Themis, affected by the $n - 2\,n_5$ divisor, which is smaller than n_5 by a factor $\simeq 7.4$; the main second order term with the inverse cube of the divisor contains e^2. [Milani and Knežević, 1990] show that it is possible to approximately predict the amplitude of the secular change in a. These effects are anyway at the level of few parts in 10^{-4} au, thus their removal is not essential for the study of asteroid families.

In both examples, however, to measure these small secular perturbations on the semimajor axis we have been forced to use the digital filtering methods from Chapter 4, thus obtaining for (24) Themis the Figure 2.2. The accuracy of the analytic removal is not really enough to allow detection of these secular second order effects. The strict correlation between the secular perturbations in a and in e is apparent by comparing with Figure 2.3.

2.11 Iterative mean elements

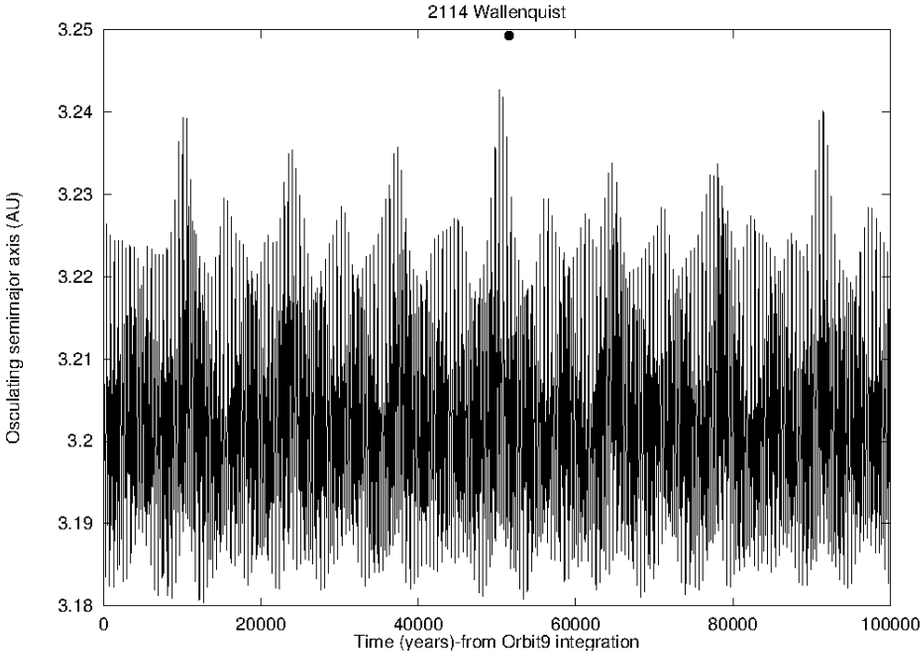


Figure 2.4: Semimajor axis of the asteroid (2114) Wallenquist, as a function of time over 100 000 years.

To understand the performance in the computation of mean elements, e.g., the mean semimajor axis a', let us chose an asteroid very close to the 2/1 mean motion resonance with Jupiter, namely (2114) Wallenquist. This asteroid has values of a lower then the one corresponding to the resonance, thus $n - 2\,n_5 > 0$, but this divisor is very small.

The short periodic oscillations in the elements, especially in the osculating semimajor axis a, are very large, see Figure 2.4, and would create difficulty in the family classification, actually the highest values would place (2114) right in the Kirkwood gap, while the lowest values would place this asteroid right in the middle of the family of (24) Themis.

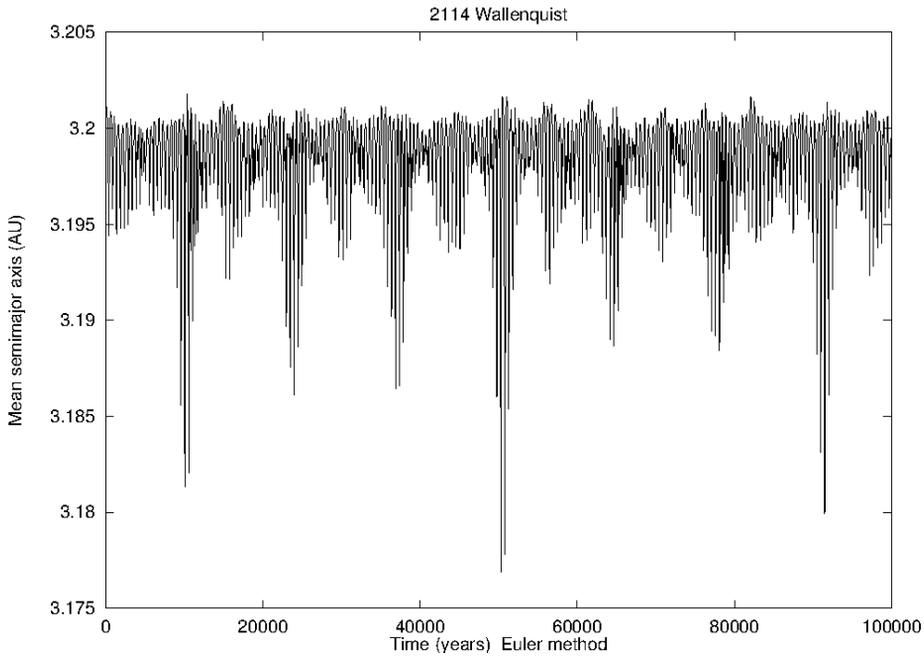


Figure 2.5: Mean semimajor axis, computed by an explicit first order theory, for the asteroid (2114) Wallenquist as a function of time over 100 000 years.

The mean elements computed with the first order theory [Milani and Knežević, 1990] exhibit a lower excursion, in particular avoiding the values too close to the resonance, see Figure 2.5. Still there are negative peaks with values > 0.02 au smaller than the average.

To understand this behavior, the best way is to look at the Figure 2.6 which allows to appreciate the way the correction, which should approximate $a' - a$, behaves as a changes. Indeed, if osculating a has a large value, very close to the resonance, such as 3.24 au, then the divisor $n - 2\,n_5$ is still positive but small, and from the $\chi_{2/1}$ term a correction containing $-\partial\chi_{2/1}/\partial\lambda$ is introduced, containing the inverse of the divisor, thus quite large. But then this results in overshooting, that is the computed value is $a' \simeq 3.18$ au, forming one of the negative peaks.

To the contrary, if the osculating value is $a \simeq 3.18$ au, then the divisor is much larger and the correction results in an undershooting, that is it is not enough to bring the computed a' to the intuitively right value, which should be just above 3.20 au.

The above empirical observation of the behavior of the computation of a' by a first order theory suggests a method which could give better results [Milani and Knežević, 1999]. Let us suppose that we knew the values of the mean semimajor axes, both for the planet and for the asteroid, namely a'_5 and a', thus we can compute the values of the **mean mean motion** for both, n'_5 and n'; then the divisor would have a value $n' - 2\,n'_5$ which would change little with time. Then we could use a value of $\chi_{2/1}$ containing the divisor as $n' - 2\,n'_5$, and the computed correction could be subtracted from a' to give a with better accuracy. (We assume here we know a'_5, although to compute it a similar argument should be applied, taking into account the perturbations from the other planets.)

This is not an algorithm for computing a', since we use its value to start with. However, if we assume we know the map

$$\Lambda = \Lambda' + \frac{\partial\chi_{2/1}(\Lambda', \lambda')}{\partial\lambda} \ ,$$

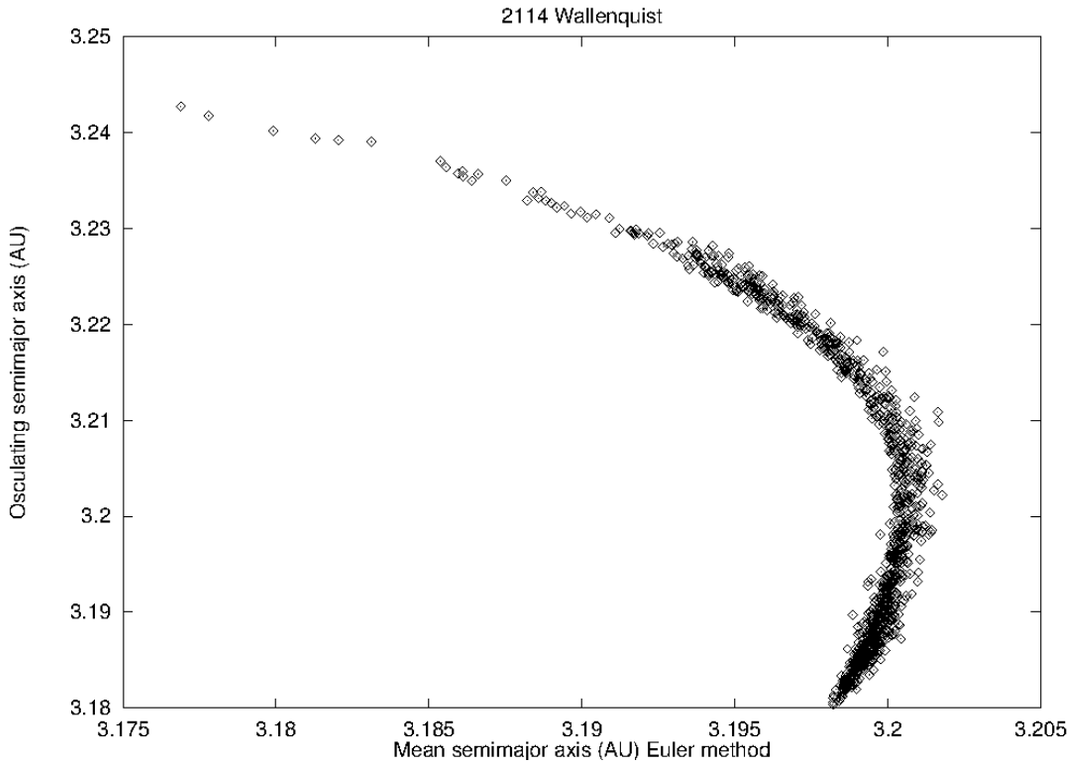


Figure 2.6: Mean semimajor axis, computed by an explicit first order theory, and osculating semimajor axis for the asteroid (2114) Wallenquist, over a time span of 100 000 years.

and that it is close to the identity, its inverse can be computed by using the fixed point method[16], that is an iterative procedure starting from $\Lambda^{(0)} = \Lambda$ and $\lambda^{(0)} = \lambda$ with recursive equation

$$\Lambda'^{(k+1)} = \Lambda - \frac{\partial \chi_{2/1}(\Lambda'^{(k)}, \lambda'^{(k)})}{\partial \lambda} \quad , \quad \lambda'^{(k+1)} = \lambda + \frac{\partial \chi_{2/1}(\Lambda'^{(k)}, \lambda'^{(k)})}{\partial \Lambda} \; .$$

For ϵ small enough the iterations may converge, then the limit for $k \to +\infty$ would give (Λ', λ').

The formulas above are somewhat simplified, because in a real computation we cannot consider just the resonant term, but also all the others, including the one with argument $\lambda - 2\,\lambda_5 + \varpi_5$. The changes in the elements e, I between osculating and mean are relevant too, thus the iteration should be done on all the coordinates, including λ, while in the first order theory λ' does not need to be computed. Moreover, the convergence of the iterations cannot always be guaranteed, e.g., if the divisor changes sign then divergent iterations can be expected; indeed, there are resonant orbits, for which the divisor is permanently very close to zero.

The results of the iterative method are shown in Figure 2.7, which indicate there is still an effect not completely removed by the iterations. However, the values of a' at convergence are closely clustered, and Figure 2.8 shows that, although not all the short periodic perturbations have been removed, still the maximum excursion is now $\simeq 0.005$ au, more than an order of magnitude smaller than the excursion of the osculating a. For comparison, the mean semimajor axis computed by digital filtering, plotted in Figure 4.3, clearly shows the higher order long period changes, which are of the same nature as those for (24) Themis but almost an order of magnitude stronger, due to the asteroid being closer to the 2/1 resonance.

[16][Milani and Knežević, 1999] have actually tested several different methods, finding that the iterative fixed point method best suits this specific application.

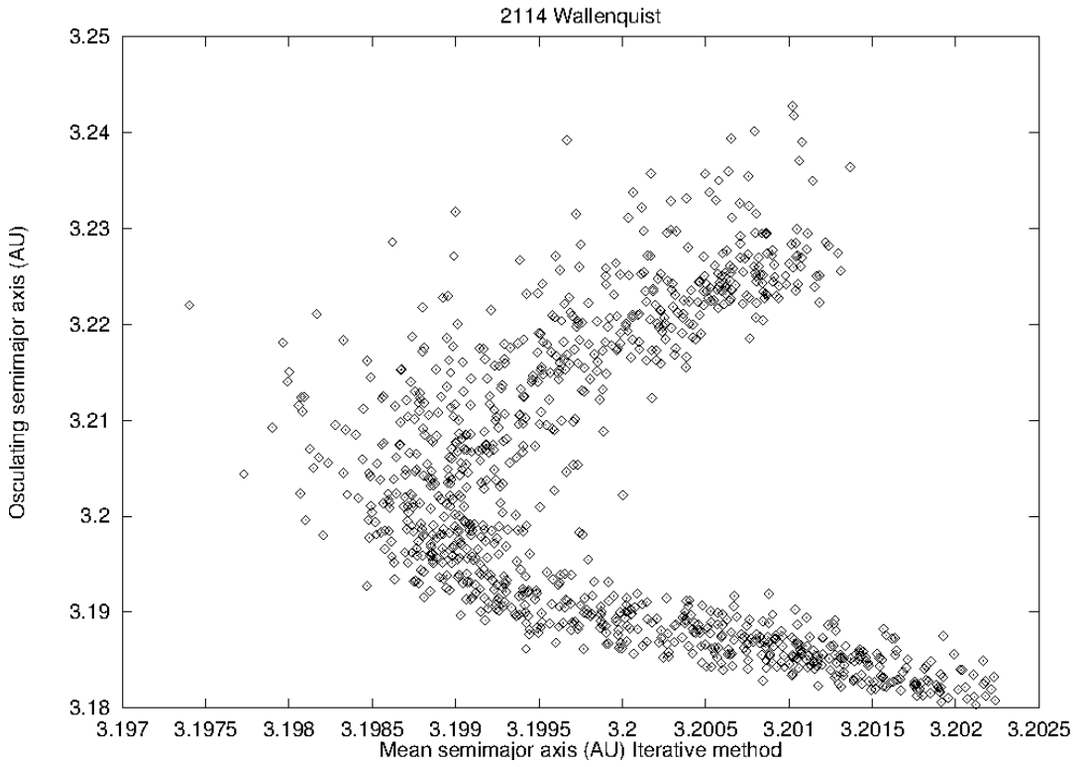


Figure 2.7: Mean semimajor axis, computed by an iterative theory, and osculating semimajor axis for the asteroid (2114) Wallenquist, over a time span of 100 000 years.

A more global view of the situation with the computation of mean elements can be appreciated from Figure 2.9, where crosses mark values of mean (a', e') for which the computation with the first order theory and with the iterative theory results in a difference $> 10^{-3}$ au in a'. The location of the crosses near the main Kirkwood gaps excavated by the strongest resonances indicates that the simple model, based on just the largest resonant term, used above is a correct qualitative explanation.

All the above is a purely empirical argument, and indeed it leaves many open problems: first and most important, why should a computation of the inverse map $a' \mapsto a$ work better than the computation of $a \mapsto a'$, the one we are actually looking for? Does this contradict the statement we have made previously that the map defined by the determining function $-\chi(\mathbf{z}')$ is the inverse of the one defined by $\chi(\mathbf{z})$?

In fact we have shown that the main term $\chi_{2/1}$ has a very different values when computed in the mean elements space. How can a smooth function χ have zero derivative with respect to the one-parameter group of transformations determined by itself, that is $\{\chi, \chi\} = 0$, and then not be invariant? The answer, of course, is that $\chi(\mathbf{z})$ does not exist at all as a smooth function, but only as a formal series. In fact, $\chi(\mathbf{z}')$ might be in some sense better defined. To understand all these apparent paradoxes, please wait until Chapter 6.

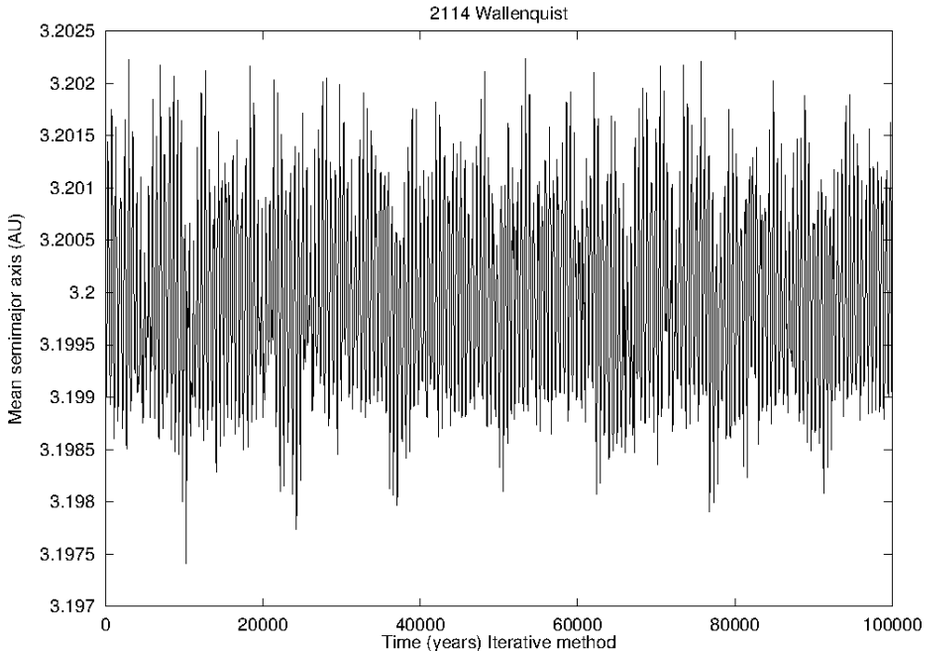


Figure 2.8: Mean semimajor axis, computed by an iterative theory, for the asteroid (2114) Wallenquist, as a function of time over 100 000 years.

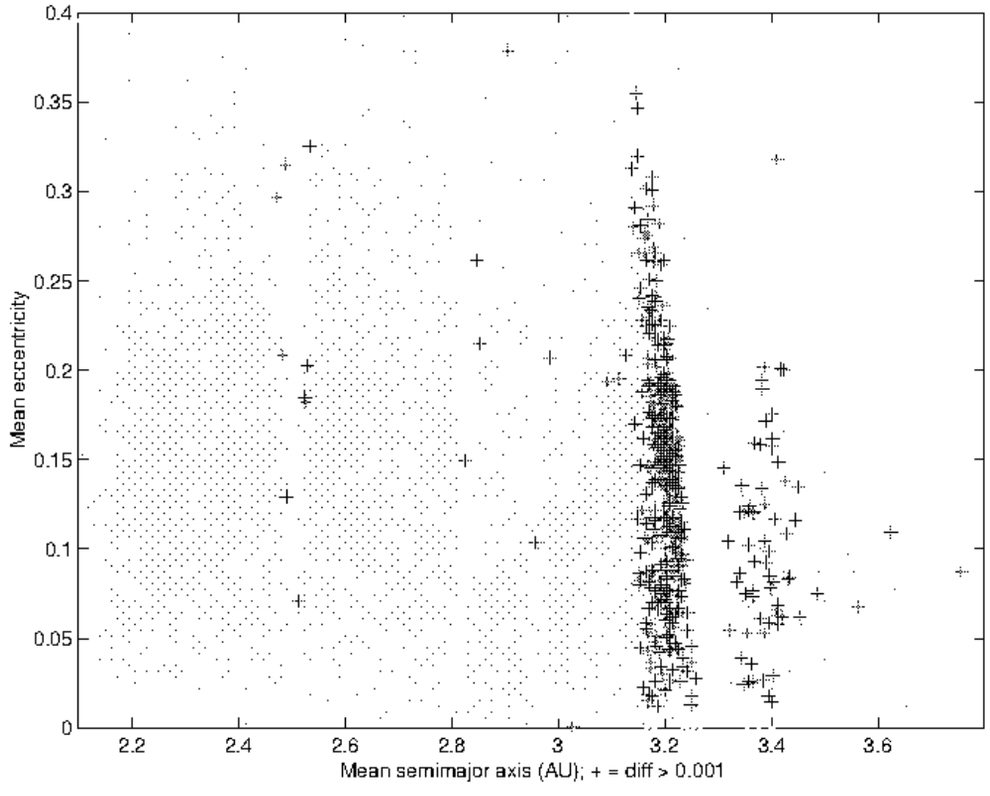


Figure 2.9: The mean semimajor axis and mean eccentricity computed for 13, 345 asteroids. The crosses indicate cases in which the difference between the mean a computed by a first order theory and the one computed with an iterative theory exceeds 0.001 au. These cases are concentrated near the 2/1, 3/2 and 5/2 resonances with Jupiter, and in the outer part of the main belt, where there are many mean motion resonances. [problem:color]

Bibliography

[Arnold 1963a] Arnold, V.I. (1963), Proof of a Theorem by A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian, *Russian Mathematical Surveys* **18**, 13–40.

[Arnold 1963b] Arnold, V.I. (1963), Small denominators and problems of stability of motion in classical and celestial mechanics, *Russian Mathematical Surveys* **18**, 85–191.

[Arnold 1976] Arnold, V. (1976). *Mathematical Methods of Classical Mechanics*, Springer, Berlin.

[Arnold and Avez, 1968] Arnold, V.I. and Avez, A. (1968), *Ergodic problems of classical mechanics*, W. A. Benjamin Inc., New York, Amsterdam.

[Benettin et al., 1984] Benettin, G., Galgani, L. and Giorgilli, A. (1984), A proof of Kolmogorov theorem on invariant tori using canonical transformations without inversion, *Nuovo Cimento* **79b**, 201–223.

[Broucke and Cefola, 1972] Broucke, R. A. and Cefola, P. J. (1972). On the equinoctial orbit elements, *CMDA* **5**, 303–310.

[Brouwer and Clemence 1961] Brouwer, D. and Clemence, G. M. (1961). *Methods of Celestial Mechanics* (Academic Press, New York, London).

[Carpino, 1987] Carpino, M., Milani, A. and Nobili,A.M. (1987), Long-term numerical integrations and synthetic theories for the motion of the outer planets. *Astron. Astrophys.* **181**, 182–194.

[Ellis and Murray 2000] Ellis, K. M. and Murray C. D. (2000). The disturbing function in Solar System dynamics, *Icarus* **147**, 129–144.

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

[Ferraz-Mello 1981] Ferraz-Mello, S. (1981), Estimation of periods from unequally spaced observations, *Astron. J.*, **86**, 619–624.

[Ferraz-Mello 1994] Ferraz-Mello, S. (1994), The convergence domain of the Laplacian expansion of the disturbing function *CMDA* **58**, 37–52.

[Ferraz-Mello 1995] Ferraz-Mello, S. (1995), On the convergence of the disturbing function. *From Newton to Chaos: Modern Techniques for Understanding and Coping with Chaos in N-Body Dynamical Systems*, (A.E. Roy, B. Steves, Eds.), NATO ASI series, pp. 97–98.

[Grönwall 1919] Grönwall, T. H. (1919), Note on the derivatives with respect to a parameter of the solutions of a system of differential equations, *Ann. of Math.*, **20** (2): 292–296.

[Hartmann, 1964] Hartmann, P. (1964). *Ordinary Differential Equations* (J. Whiley and Sons, Hoboken, NJ).

[Hirayama, 1918] Hirayama, K. (1918), *Astron.J.*, **31**, 185–188.

[Hirayama, 1922] Hirayama, K. (1926), *Japanese J. Astron. Geo.*, **1**, 55–93.

[Hori, 1966] Hori, G.: 1966, Theory of General Perturbations with Unspecified Canonical Variables *Publ. Astron. Soc. Japan* **18**, 287–296.

[Knežević 1986] Knežević, Z. (1986), Secular variations of major planets' orbital elements. *Celestial Mechanics* **38**, 123–138.

[Knežević 1989] Knežević, Z. (1989), Asteroid long-periodic perturbations - The second order Hamiltonian. *Celest. Mech.* **46**, 147–158.

[Knežević 1993a] Knežević, Z. (1993a), Minor planet short periodic perturbations - The indirect part of the disturbing function, *CMDA* **55**, 387–404.

[Knežević 1993b] Knežević, Z. (1993b), "Generic term" technique in computation of asteroid perturbations, *Publ. Obs. Astron. Belgrade* **44**, 19–21.

[Knežević and Ćatović 1994] Knežević, Z. and Ćatović, Z. (1994) On the accuracy of Laplace and Leverrier coefficients, *Bull. Acad. Serbe Sci. Arts* **107**, Sci. math. N^o 19), 1–17.

[Knežević & Milani, 2000] Knežević, Z. and Milani, A. (2000), Synthetic proper elements for outer main belt asteroids. *Celest. Mech. Dyn. Astron.* **78**, 17–46.

[Knežević *et al.* 1995] Knežević, Z., Froeschlé, Ch., Lemaitre, A., Milani, A. and Morbidelli, A. (1995) Comparison between two theories of asteroid proper elements, *Astron. Astrophys.* **293**, 605–612.

[Kolmogorov, 1954] Kolmogorov, A.N. (1954), The conservation of conditionally periodic motions with a small variation in the Hamiltonian *Dokl. Akad. Nauk. SSSR* textbf98, 527–530.

[Laskar 1989] Laskar, J. (1989), Les variables de Poincaré et le développement de la fonction perturbatrice, *Notes scientifiques et techniques du Bureau des longitudes* **S026** pp. 22–58.

[Lemaitre and Morbidelli 1994] Lemaitre, A. and Morbidelli, A. (1994) Proper elements for highly inclined asteroidal orbits, *CMDA* **60**, 29-56.

[Le Verrier 1855] Le Verrier, U.J.-J. (1855). Dveloppment de la fonction qui sert de base au calcul des perturbations des mouvements des plantes, *Ann. Obs. Paris, Mm.*, **1**, pp. 258-331.

[Le Verrier 1855] Le Verrier, U.J.-J. (1856). Dveloppment de la fonction qui sert de base au calcul des perturbations des mouvements des plantes, *Ann. Obs. Paris, Mm.*, **2**, pp. 1-20.

[Milani and Nobili 1984] Milani, A. and Nobili, A.M. (1984), Resonant structure of the outer asteroid belt, *CMDA* **34**, 343–355.

[Milani and Knežević, 1990] Milani, A., and Knežević, Z. (1990), Secular perturbation theory and computation of asteroid proper elements, *CMDA* **49**, 347–411.

[Milani and Knežević, 1992] Milani, A., and Knežević, Z. (1992), Asteroid proper elements and secular resonances. *Icarus* **98**, 211–232.

[Milani and Knežević, 1994] Milani, A., and Knežević, Z. (1994), Asteroid proper elements and the Dynamical Structure of the Asteroid Main Belt. *Icarus* **107**, 219–254.

[Milani and Knežević, 1999] Milani, A & Z. Knežević: 1999, Asteroid mean elements: higher order and iterative theories, *CMDA* **71**, 55-78.

[Milani *et al.* 1987a] Milani, A., Nobili A. M. and Farinella, P. (1987). *Non gravitational perturbations and satellite geodesy*, Adam Hilger, Liverpool.

[Milani *et al.* 1987b] Milani, A., Nobili, A.M. and Carpino, M. (1987b), Secular variations of the semimajor axes - Theory and experiments, *Astron. Astrophys.* **172**, 265-279.

[Milani *et al.* 2007] Milani, A., Gronchi, G. F. and Knežević, Z. (2007). New Definition of Discovery for Solar System Objects, *Earth Moon Planets* **100**, 83–116.

[Milani and Gronchi 2010] Milani, A., and Gronchi, G. F. (2010). *Theory of Orbit Determination*, Cambridge University Press, Cambridge UK.

[Milani *et al.* 2010a] Milani, A., Knežević, Z., Novaković, B. and Cellino, A. (2010) Dynamics of the Hungaria asteroids. *Icarus*, **207**, 769–794.

[Milani et al., 2010b] Milani, A., Tommei, G., Vokhroulický, D., La Torre, E., and Cicaló, S. (2010), Relativistic models for the BepiColombo radioscience experiment, in *Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis*, (S. Klioner, P. K. Seidelmann and M. Soffel, eds.), Cambridge University Press, pp. 356–365.

[Milani et al., 2014] Milani A., Cellino, A., Knežević, Z., Novaković, B., Spoto, F. and Paolicchi, P. (2014), Asteroid families classification: Exploiting very large datasets. *Icarus* **239**, 46–73.

[Milani et al., 2016] Milani, A., Spoto, F. Knežević, Z., Novaković, B. and Tsirvoulis, G. (2016), Families classification including multiopposition asteroids. in: *Proceedings IAU Symposium 318: Asteroids: New Observations New Models*, (S. R. Chesley, A. Morbidelli, R. Jedicke and D. Farnocchia, Eds.), Cambridge Univ. Press, pp. 28–45.

[Morbidelli, 1993] Morbidelli, A. (1993), Asteroid secular resonant proper elements. *Icarus* **105**, 48-66.

[Morbidelli, 2002] Morbidelli, A. (2002), *Modern Celestial Mechanics: Aspects of Solar System Dynamics*, Taylor and Francis.

[Murray 1985] Murray C. D. (1985). A note on Le Verrier's expansion of the disturbing function, *CMDA* **36**, 163–164.

[Nobili et al., 1989] Nobili, A.M., Milani, A. and Carpino, M. (1989), Fundamental frequencies and small divisors in the orbits of the outer planets, *Astronomy and Astrophysics* **210**, 313–336.

[Poincaré 1892] Poincaré, H. (1892), *Les Méthodes Nouvelles de la Mécanique Céleste, Tome I*, Gauthier–Villars, Paris; reprinted by Librairie A. Blanchard, Paris, 1987. [English translation: *New Methods of Celestial Mechanics*, American Institute of Physics, 1993]

[Poincaré 1893] Poincaré, H.: 1893, *Les Méthodes Nouvelles de la Mécanique Céleste, Tome II*, Gauthier–Villars, Paris; reprinted by Librairie A. Blanchard, Paris, 1987. [English translation: *New Methods of Celestial Mechanics*, American Institute of Physics, 1993]

[Sokolov, V. G. 1989] Sokolov, V. G. (1989), New Relations for Laplace Coefficients *Astron. Zh.* **66**, 502–510 (in Russian).

[Sundman 1901] Sundman, K. F. (1901), Über der Störungen der Kleinen Planeten, *Akad. Abhandl.*, K.Alexanders Univ., Helsingfors.

[Sundman 1916] Sundman, K. F. (1916), Sur les conditions nécessaires et suffisantes puor la convergence du développement de la function perturbatrice dans le mouvement plan, *Öfversigt Finska Vetenskaps-Soc. Förh.*, **58** A(24).

[Tisserand 1889] Tisserand F. (1889). *Traite de Mechanique Celeste,* Tome 1 (Gauthier-Villars, Paris).

[Tupikova *et al.* 1999] Tupikova, I., Soffel, M. and Klioner, S. (1999). On the classical expansion of the perturbing function in individual orbital elements, *CMDA* **74**, 147–152.

[Vokrouhlický *et al.* 2000] Vokrouhlický, D, Milani A. and Chesley, S. R. (2000). Yarkovsky effect on small Near Earth asteroids: mathematical formulation and examples, *Icarus* **148**, 118–138.

[Williams 1969] Williams, J. G. (1969) Secular perturbations in the Solar System, Ph.D. Thesis, Univ. California, Los Angeles.

[Yuasa 1973] Yuasa M. (1973). Theory of secular perturbations of asteroids including terms of higher orders and higher degrees, *Publ. Astron. Soc. Japan* **25**, 399–445.

NOTA 1

ANTONIO GIORGILLI

« Appunti di Meccanica Celeste »

www.mat.unimi.it/users/antonio/meccel/

7

IL PROBLEMA RISTRETTO DEI TRE CORPI

Il problema che intendo trattare qui si riallaccia idealmente alla discussione del capitolo 2. Avendo studiato in dettaglio la dinamica di un sistema a due corpi sembrerebbe naturale, volendo procedere passo passo senza gettarsi immediatamente nello studio del problema planetario in forma generale, tentare anzitutto di studiare un problema a tre corpi. A tale conclusione, dopotutto, era già arrivato Lagrange, che pur avendo ottenuto i brillanti risultati sulle perturbazioni secolari che ho esposto nel capitolo 6, non aveva rinunciato all'idea di cercare soluzioni in forma generale, e aveva appunto iniziato con uno studio accurato del problema dei tre corpi, proseguendo delle ricerche già iniziate da Eulero.

Lasciando la discussione del problema generale pe il capitolo successivo, mi occuperò qui del *problema ristretto dei tre corpi*. Si tratta di una versione per così dire semplificata del problema generale. È appena il caso di sottolineare che l'aggettivo "semplificata" non è da intendersi nel senso di "semplice": è vero che il numero di gradi di libertà viene ridotto in modo drastico, ma ciò serve solo a ridurre la complessità del calcolo, e non le reali difficoltà che restano tutte ben presenti.

Il problema si enuncia come segue.

> *Due punti materiali, detti corpi primari, si muovono nello spazio su un'orbita Kepleriana (ellittica o circolare). Un terzo punto P di massa trascurabile rispetto ai primi due, detto planetoide, si muove sotto l'azione della forza Newtoniana esercitata dai primari, senza influenzarne il movimento. Si chiede di studiare la dinamica del punto P.*

In altre parole, si suppone che i primari si comportino come un sistema a due corpi la cui dinamica non viene influenzata dal planetoide. A sua volta, il planetoide si muove sotto l'azione di un ambiente esterno, rappresentato appunto dai primari.

Nell'ambito della Meccanica Celeste il problema dei tre corpi si presenta come lo schema più naturale in cui inquadrare, almeno in prima approssimazione, problemi quali il moto dei pianeti interni all'orbita di Giove (Mercurio, Venere, Terra, Marte) o degli asteroidi quando si tenga conto dell'azione di Giove e del Sole. Si tratta ovvi-

amente di un'approssimazione: si assume che le perturbazioni più consistenti siano dovute al pianeta di maggior massa, ossia Giove, ignorando gli effetti dovuti a Saturno, Urano e Nettuno. Nel caso dei pianeti interni e degli asteroidi della fascia principale ciò può giustificarsi ulteriormente in considerazione del fatto che gli altri pianeti maggiori sono molto più lontani di Giove. Lo schema del modello ristretto può applicarsi anche allo studio della dinamica dei satelliti naturali e artificali.

Si possono considerare diverse varianti. Si distingue anzitutto il caso *circolare*, in cui i due primari ruotano a velocità uniforme rispetto al baricentro comune, dal caso *ellittico*. Si distingue poi il caso *piano*, in cui il planetoide è vincolato a muoversi nel piano dell'orbita dei primari, da quello *spaziale*. Combinando tra loro queste due scelte si hanno quattro casi possibili, e comunque tutti non integrabili (o, per essere pignoli, non integrati).

7.1 L'Hamiltoniana e le equazioni

Iniziamo con lo scrivere le equazioni del problema ristretto circolare in forma Hamiltoniana. A tal fine è particolarmente comodo far uso di un sistema di riferimento in cui i primari occupino delle posizioni fisse. Inoltre è conveniente fissare le unità di misura in modo da minimizzare il numero di costanti.

7.1.1 Scelta delle unità di misura

È uso comune, del resto conveniente, scegliere le unità di misura in modo da ricondurre il problema alla forma più semplice possibile. Come misura di lunghezza si sceglie il semiasse maggiore dell'orbita kepleriana dei primari, che nel caso circolare è la distanza tra i due. Si sceglie poi l'unità di massa pari alla somma delle masse dei primari (quella del planetoide è trascurabile), e si denota con μ la massa di uno dei due primari (solitamente la più piccola), sicché l'altra massa risulta essere $1-\mu$. Infine si pone la costante di gravitazione $\mathcal{G}=1$, il che equivale a fissare l'unità di tempo.

Ricordiamo che per la soluzione Kepleriana del problema dei due corpi vale la relazione tra semiassi e periodi

$$\frac{a^3}{T^2}=\frac{\mathcal{G}(m_1+m_2)}{4\pi^2}\ ,$$

dove m_1 , m_2 sono le masse. Ne segue che nelle unità di misura che abbiamo scelto la soluzione Kepleriana circolare per il moto dei due primari ha periodo $T=2\pi$, e dunque frequenza angolare $\omega=1$.

7.1.2 L'Hamiltoniana e le equazioni canoniche nel caso circolare

Il procedimento tradizionale, e in effetti anche il più comodo, consiste nel considerare un sistema di riferimento solidale coi primari, e quindi in moto rotatorio uniforme attorno al baricentro dei primari. In questo sistema di riferimento si pongono le due masse sull'asse x, con la massa maggiore dal lato positivo, come illustrato in figura 7.1. Dunque, la massa $1-\mu$ occupa la coordinata μ e la massa μ occupa la posizione $-1+\mu$ sull'asse x.

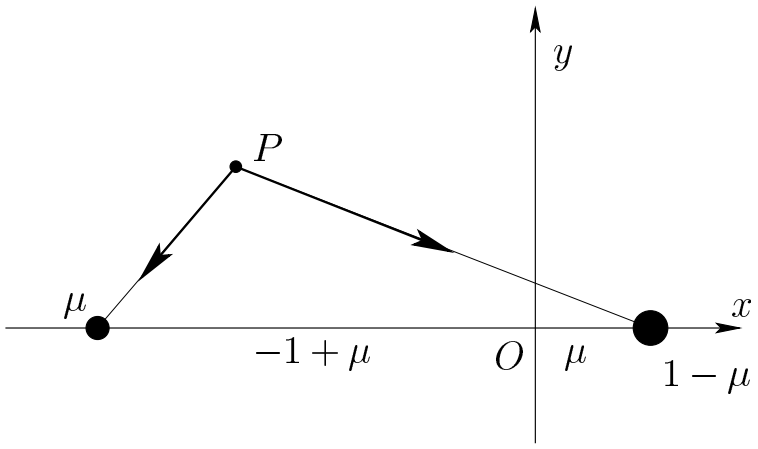


Figura 7.1. Ad illustrazione del problema dei tre corpi nel caso ristretto, circolare e piano.

Consideriamo anzitutto un sistema di riferimento fisso, e denotiamo con ξ, η, ζ le coordinate cartesiane del planetoide. In un primo momento scriviamo in modo esplicito tutte le costanti, riservandoci di tener conto più avanti della scelta delle unità di misura. L'Hamiltoniana si scrive

$$H(\xi,\eta,\zeta,p_\xi,p_\eta,p_\zeta,t)=\frac{1}{2m}(p_\xi^2+p_\eta^2+p_\zeta^2)+V(\xi,\eta,\zeta,t)\ . \tag{7.1}$$

L'energia potenziale è quella gravitazionale, e sarà la somma dei due termini dovuti all'interazione del planetoide con i primari, ossia

$$V=-\frac{\mathcal{G}(1-\mu)m}{r_1}-\frac{\mathcal{G}\mu m}{r_2} \tag{7.2}$$

dove r_1 è la distanza tra il planetoide e il primario di massa $1-\mu$ e r_2 la distanza dal secondo primario. La dipendenza dal tempo deriva dal movimento dei due primari nel sistema di riferimento assoluto. Qui non serve scrivere l'espressione esplicita della distanza: vedremo subito che nel sistema rotante la forma è alquanto più semplice, e si scrive in modo diretto.

Dette x,y,z le coordinate in un sistema di riferimento rotante uniformemente in senso antiorario intorno all'asse ζ con velocità angolare ω, e scegliendo l'origine del sistema rotante coincidente con quella del sistema fisso (sicché gli assi z e ζ coincidono) si hanno le relazioni geometriche

$$\xi=x\cos\omega t-y\sin\omega t\ ,\quad \eta=x\sin\omega t+y\cos\omega t\ ,\quad \zeta=z\ , \tag{7.3}$$

e le relazioni inverse

$$x=\xi\cos\omega t+\eta\sin\omega t\ ,\quad y=-\xi\sin\omega t+\eta\cos\omega t\ ,\quad z=\zeta\ .$$

Si tratta di una trasformazione puntuale dipendente dal tempo, e possiamo estenderla a trasformazione canonica mediante la funzione generatrice

$$S(p_x,p_y,p_z,\xi,\eta,\zeta)=p_x(\xi\cos\omega t+\eta\sin\omega t)+p_y(-\xi\sin\omega t+\eta\cos\omega t)+p_z\zeta\ ,$$

ma dovremo ricordare che l'Hamiltoniana trasformata è $H + \frac{\partial S}{\partial t}$, come abbiamo visto nella proposizione 3.7. La trasformazione sui momenti si scrive

$$p_\xi = p_x \cos\omega t - p_y \sin\omega t \ , \quad p_\eta = p_x \sin\omega t + p_y \cos\omega t \ , \quad p_\zeta = p_z \ . \tag{7.4}$$

Dobbiamo poi calcolare

$$\frac{\partial S}{\partial t} = \omega p_x(-\xi \sin\omega t + \eta\cos\omega t) - \omega p_y(\xi\cos\omega t + \eta\sin\omega t) = -\omega p_y x + \omega p_x y \ .$$

dove abbiamo sostituito la trasformazione (7.3) sulle coordinate.

Otteniamo dunque l'Hamiltoniana del problema circolare ristretto nel sistema rotante (che con un piccolo abuso di notazione indicheremo ancora con H)

$$\begin{aligned} H(x,y,z,p_x,p_y,p_z) = \frac{1}{2m}(p_x^2+p_y^2+p_z^2) - \omega x p_y + \omega y p_x \\ - \frac{\mathcal{G}(1-\mu)m}{\sqrt{(x-\mu)^2+y^2+z^2}} - \frac{\mathcal{G}\mu m}{\sqrt{(x+1-\mu)^2+y^2+z^2}} \ . \end{aligned}$$

Nella scrittura dell'energia potenziale si è tenuto conto del fatto che nel sistema di riferimento solidale con i primari questi occupano due posizioni fisse, con la massa maggiore $1-\mu$ nel punto $(\mu,0,0)$ e la massa minore μ nel punto $(-1+\mu,0,0)$. Ora teniamo conto delle unità di misura che abbiamo scelto all'inizio, sicché porremo $\mathcal{G}=1$, $\omega = 1$. Infine possiamo eliminare la massa m mediante la trasformazione di scala

$$p_x = m p'_x \ , \quad p_y = m p'_y \ , \quad p_z = m p'_z \ ,$$

e lasciando inalterate le coordinate x, y, z. La trasformazione non è canonica in senso stretto, e occorre dividere per m l'Hamiltoniana trasformata (si veda il paragrafo 3.2.1). Nel nostro caso ciò corrisponde di fatto ad eliminare un fattore m comune a tutti i termini, sicché si ottiene (rimuovendo gli apici) la forma comunemente usata

$$\begin{aligned} H(x,y,z,p_x,p_y,p_z) = \frac{1}{2}(p_x^2+p_y^2+p_z^2) - x p_y + y p_x \\ - \frac{1-\mu}{\sqrt{(x-\mu)^2+y^2+z^2}} - \frac{\mu}{\sqrt{(x+1-\mu)^2+y^2+z^2}} \ . \end{aligned} \tag{7.5}$$

Scriviamo infine le equazioni canoniche

$$\begin{aligned} \dot{x} &= p_x + y \\ \dot{y} &= p_y - x \\ \dot{z} &= p_z \\ \dot{p}_x &= \ \ p_y - \frac{(1-\mu)(x-\mu)}{[(x-\mu)^2+y^2+z^2]^{3/2}} - \frac{\mu(x+1-\mu)}{[(x+1-\mu)^2+y^2+z^2]^{3/2}} \\ \dot{p}_y &= -p_x - \frac{(1-\mu)y}{[(x-\mu)^2+y^2+z^2]^{3/2}} - \frac{\mu y}{[(x+1-\mu)^2+y^2+z^2]^{3/2}} \ . \\ \dot{p}_z &= -\frac{(1-\mu)z}{[(x-\mu)^2+y^2+z^2]^{3/2}} - \frac{\mu z}{[(x+1-\mu)^2+y^2+z^2]^{3/2}} \ . \end{aligned} \tag{7.6}$$

Si vede subito che ponendo come dati iniziali $z(0) = 0$ e $p_z(0) = 0$, con dati arbitrari per la altre variabili, allora la terza e l'ultima equazione hanno la soluzione banale $z(t) = 0$, $p_z(t) = 0$. In tal caso ci si riduce a considerare il problema piano, descritto dalle sole equazioni

$$
\begin{aligned}
\dot{x} &= p_x + y \\
\dot{y} &= p_y - x \\
\dot{p}_x &= \quad p_y - \frac{(1-\mu)(x-\mu)}{\left[(x-\mu)^2+y^2\right]^{3/2}} - \frac{\mu(x+1-\mu)}{\left[(x+1-\mu)^2+y^2\right]^{3/2}} \\
\dot{p}_y &= -p_x - \frac{(1-\mu)y}{\left[(x-\mu)^2+y^2\right]^{3/2}} - \frac{\mu y}{\left[(x+1-\mu)^2+y^2\right]^{3/2}} \ .
\end{aligned}
\tag{7.7}
$$

7.2 Gli equilibri lagrangiani

Vogliamo ora cercare le soluzioni di equilibrio per il sistema (7.6). Dal momento che le equazioni per z, p_z ammettono $z = p_z = 0$ come unico punto di equilibrio potremo restringere la nostra attenzione al problema piano, come descritto dalle equazioni (7.7). Il lettore noterà che gli equilibri sono le posizioni in cui il planetoide sta fermo in un sistema di riferimento che ruota uniformemente. Nel sistema fisso si vedranno i tre corpi girare con velocità uniforme mantenendo la stessa posizione relativa.

7.2.1 Calcolo degli equilibri relativi

Occorre annullare i secondi membri delle equazioni. Ponendo $\dot{x} = \dot{y} = 0$ otteniamo

$$p_x = -y \ , \quad p_y = x \ .$$

Le altre due equazioni diventano

$$
\begin{aligned}
x - \frac{(1-\mu)(x-\mu)}{r_1^3} - \frac{\mu(x+1-\mu)}{r_2^3} &= 0 \\
y - \frac{(1-\mu)y}{r_1^3} - \frac{\mu y}{r_2^3} &= 0 \ .
\end{aligned}
\tag{7.8}
$$

dove

$$r_1^2 = (x-\mu)^2 + y^2 \ , \quad r_2^2 = (x+1-\mu)^2 + y^2 \tag{7.9}$$

sono le distanze del planetoide dai due primari.

Per il seguito è conveniente anche fattorizzare x e y al numeratore, e riscrivere le equazioni nella forma

$$
\begin{aligned}
x\left(1 - \frac{1-\mu}{r_1^3} - \frac{\mu}{r_2^3}\right) + \mu(1-\mu)\left(\frac{1}{r_1^3} - \frac{1}{r_2^3}\right) &= 0 \\
y\left(1 - \frac{1-\mu}{r_1^3} - \frac{\mu}{r_2^3}\right) &= 0
\end{aligned}
\tag{7.10}
$$

NOTA 2

Consideriamo

$$z^{(0)} = z(s, z_0)$$

$$z^{(\varepsilon)} = z(s, z_0 + \varepsilon v_0)$$

con $|\varepsilon v_0| \ll 1$, soluzioni delle equazioni di Hamilton per χ

$$\frac{dz}{ds} = J \nabla \chi$$

Si ha

$$z^{(\varepsilon)} = z^{(0)} + \overbrace{\frac{\partial z^{(\varepsilon)}}{\partial z} \varepsilon v_0}^{= z^{(1)}} + O(\varepsilon^2)$$

$$\frac{dz^{(\varepsilon)}}{ds} = J \nabla \chi(z^{(0)}) + J \nabla \nabla \chi(z^{(0)}) z^{(1)} + O(\varepsilon^2)$$

Allora si può scrivere

$$\frac{dz^{(1)}}{ds} = J \nabla \nabla \chi(z^{(0)}) z^{(1)}$$

Definisco $v = z^{(\varepsilon)} - z^{(0)}$

Trascurando i termini non lineari in ε pongo $v = z^{(1)}$

Allora ho

$$\frac{dv}{ds} = J \nabla \nabla \chi(z(s, z_0)) v = J \nabla H(v)$$

con $H(v) = \frac{1}{2} v \cdot \nabla \nabla \chi(z(s, z_0)) v$

NOTA 3

ANTONIO GIORGILLI

« Notes on Hamiltonian Dynamical Systems »

$$\det\left(\frac{\partial\Psi_j}{\partial\Phi_k}\right)\neq 0\ . \tag{3.15}$$

Then:

(i) the functions $\Psi_1,\dots,\Psi_n$ form a complete involution system;

(ii) the Liouville coordinates conjugated to Ψ are

$$\beta_j=\sum_{k=1}^{n}\frac{\partial\Phi_k}{\partial\Psi_j}\alpha_k\ , \tag{3.16}$$

where $\alpha_1,\dots,\alpha_n$ are the Liouville coordinates conjugated to Φ.

Proof. (i) The functions Ψ are independent in view of condition (3.15); moreover, they are obviously in involution, being functions only of Φ.
(ii) Recall the extended point transformation, Example 2.25. The generating function is written as

$$S(\Phi,\beta)=\sum_{j=1}^{n}\Psi_j(\Phi)\beta_j\ ,$$

from which (3.16) immediately follows. *Q.E.D.*

3.2 Liouville's Theorem

For a generic system of differential equations on an n-dimensional manifold a complete integration by quadrature can be performed when $n-1$ independent first integrals are known, n being the dimension of the space. Thus, one expects that in the Hamiltonian case, the dimension of phase space being $2n$, one needs $2n-1$ first integrals. However, the canonical structure allows us to perform the complete integration if only n first integrals are known, provided that they fulfil the further condition of being in involution. The proof exploits Liouville's canonical coordinates introduced by Proposition 3.12.

Theorem 3.15: *Assume that an autonomous canonical system with n degrees of freedom and with Hamiltonian $H(q,p)$ possesses n independent first integrals $\{\Phi_1(q,p),\dots,\Phi_n(q,p)\}$ forming a complete involution system. Then the system is integrable by quadratures. More precisely, one can construct the generating function $S(\Phi,q)$ of a canonical transformation $(q,p)=\chi(\alpha,\Phi)$ such that the transformed Hamiltonian depends only on the new momenta $\Phi_1,\dots,\Phi_n$, and the solutions are expressed as*

$$\alpha_j(t)=\alpha_{j,0}+t\left.\frac{\partial H}{\partial\Phi_j}\right|_{(\Phi_{12,0},\dots,\Phi_{n,0})}\ ,\quad j=1,\dots,n\ , \tag{3.17}$$

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

Proof. By Proposition 3.12 there is a canonical transformation $(q,p)=\chi(\alpha,\Phi)$ such that $\Phi_1,\dots,\Phi_n$ are the new momenta. In view of preservation of

(i) Find the cycles γ_j $(j = 1, \ldots, n)$.
(ii) Compute the action variables by quadrature, calculating the integrals (3.51). This can be done by possibly introducing some arbitrary angle variables on the cycles and then integrating over them.
(iii) Apply the algorithm of Liouville to the new involution system $I_1, \ldots, I_n$ in order to find the angle variables $\vartheta_1, \ldots, \vartheta_n$.
(iv) If useful, and if there is any reason to do it, apply any of the transformations of Lemma 3.33 in order to obtain better sets of action-angle variables, depending on the problem at hand.

The hardest part of the algorithm is the first step (i), for it requires an integration of the system via Liouville's algorithm applied to the involution system $\Phi_1, \ldots, \Phi_n$. However, in the most commonly considered examples the first integrals have a nice form, so that the cycles are easily determined.

3.5 The Arnold–Jost Theorem

We turn now to the statement of the of Arnold–Jost Theorem.

Theorem 3.34: *Let the Hamiltonian $H(q,p)$ on the phase space $\mathscr{F}$ possess an involution system $\Phi_1, \ldots, \Phi_n$ of first integrals (so that it is integrable in Liouville's sense). Let $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$ be such that the level surface determined by the equations $\Phi_1(q,p) = c_1, \ldots, \Phi_n(q,p) = c_n$ contains a compact and connected component M_c. Then in a neighbourhood U of M_c there are canonical action-angle coordinates I, ϑ mapping $\mathscr{G} \times \mathbb{T}^n$ to U, where $\mathscr{G} \in \mathbb{R}^n$ is an open set, such that the Hamiltonian depends only on $I_1, \ldots, I_n$, and the corresponding flow is*

$$\vartheta_j(t) = \vartheta_{j,0} + t\omega_j(I_{1,0}, \ldots, I_{n,0}) \ , \quad I_j(t) = I_{j,0} \ , \quad j = 1, \ldots, n \ ,$$

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

The proof is a straightforward application of Proposition 3.31. Just proceed as in the proof of Liouville's theorem, using the actions $I_1, \ldots, I_n$ as first integrals.

3.6 Delaunay Variables for the Keplerian Problem

A remarkable application of the Arnold–Jost Theorem is the calculation of action-angle variables for the motion in a central field of force, with particular reference to the case of the Keplerian potential.[12] The latter problem

[12] The action-angle variables for Kepler's problem were discovered by Delaunay. His aim was to replace the orbital elements of a Keplerian orbit, which were used since Lagrange's time in perturbation theory, with an appropriate set of canon-

is known to possess four independent first integrals (see Examples 1.20 and 1.21). A complete involution system of first integrals has been constructed in Example 3.3. Let us recall that in spherical coordinates r, ϑ, φ with the conjugated momenta p_r, p_ϑ, p_φ the functions are

$$J = p_\varphi \ , \quad \Gamma^2 = p_\vartheta^2 + \frac{J^2}{\sin^2 \vartheta} \ , \quad E = \frac{1}{2m}\left(p_r^2 + \frac{\Gamma^2}{r^2}\right) + V(r) \ , \tag{3.52}$$

where m is the mass of the point. In Kepler's case the potential $V(r)$ is

$$V(r) = -\frac{k}{r} \ , \tag{3.53}$$

where k is a positive constant. We also recall the expression of the Hamiltonian

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\vartheta^2}{r^2} + \frac{p_\varphi^2}{r^2 \sin^2 \vartheta}\right) + V(r) \ ; \tag{3.54}$$

this actually coincides with the third integral presented earlier when the explicit expressions of Γ^2 and J are substituted.

3.6.1 Determination of Cycles

We consider the canonical flows generated by the three functions (3.52). The discussion here is quite plain, because each function involves only two conjugated variables. This considerably simplifies the construction of cycles.

The function J is a trivially integrable Hamiltonian: the conjugate variable φ is actually an angle which parameterizes the cycle γ_φ.

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 $V(\vartheta) = J^2/\sin^2 \vartheta$. For $\Gamma^2 > \Gamma^2_{\min} = J^2$ the orbit in the phase plane ϑ, p_ϑ is a closed line, giving the second cycle γ_ϑ (see Fig. 3.10). Note that the construction of the cycles γ_φ and γ_ϑ does not depend on the form of the potential $V(r)$.

The peculiar character of the Keplerian problem shows up when we come to consider the third function. It represents the Hamiltonian of a point moving on the half-line $r > 0$ under the action of the potential

$$V^*(r) = \frac{\Gamma^2}{2mr^2} + V(r) \ .$$

In the Keplerian case, setting $E_{\min} = -mk^2/(2\Gamma^2)$, the motion is bounded for $E_{\min} < E < 0$, while for $E \geq 0$ it is unbounded. In the former case

ical variables. A deduction of Delaunay variables using the Hamilton–Jacobi method is found in Poincaré's treatises [188] and (more detailed) [190]. The calculation in the present notes reflects the exposition in M. Born's book [28].

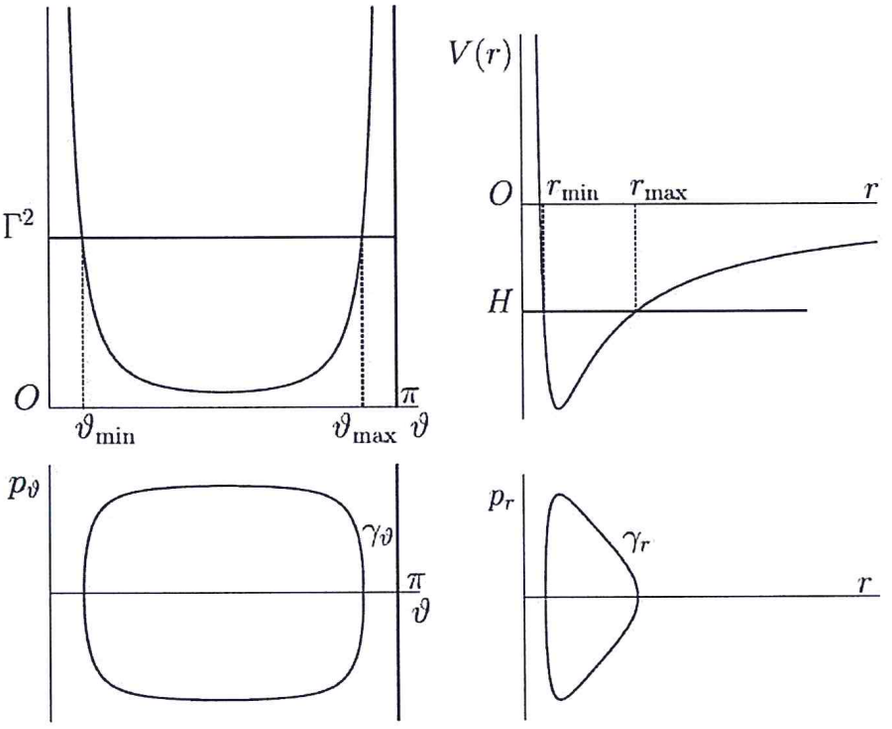


Figure 3.10 The construction of the cycles γ_ϑ and γ_r for the problem of motion in a central field under the Keplerian potential.

the orbit in the phase plane r, p_r is a closed curve, and this gives the third cycle γ_r (see Fig. 3.10). The cycle actually describes the motion of a planet on an elliptic orbit, in full agreement with Kepler's first law. Conversely, no cycle can be found for $E \geq 0$: the radial motion is unbounded, and the invariant surface in phase space for the complete problem is not compact. In the latter case, in agreement with Proposition 3.25, the invariant surface is the product $\mathbb{T}^2 \times \mathbb{R}$, and angular variables can be introduced only for the cycles γ_φ and γ_ϑ. The orbit is either a parabola, for $E = 0$, or a hyperbola, for $E > 0$.

3.6.2 Construction of the Action Variables

Here we restrict our consideration to the Keplerian potential, with the condition $E_{\min} < E < 0$. By inversion of (3.52) we get

$$
\begin{aligned}
p_r &= \left[2m(E - V(r)) - \frac{\Gamma^2}{r^2}\right]^{\frac{1}{2}} , \\
p_\vartheta &= \left(\Gamma^2 - \frac{J^2}{\sin^2\vartheta}\right)^{\frac{1}{2}} , \\
p_\varphi &= J .
\end{aligned}
\tag{3.55}
$$

We should integrate the differential form $p_r\, dr + p_\vartheta\, d\vartheta + p_\varphi\, d\varphi$ over the cycles γ_φ, γ_ϑ and γ_r. This gives the actions I_φ, I_ϑ and I_r as functions of J, Γ and E, namely

$$
\begin{aligned}
I_\varphi &= \frac{1}{2\pi} \oint_{\gamma_\varphi} p_\varphi \mathrm{d}\varphi = J , \\
I_\vartheta &= \frac{1}{2\pi} \oint_{\gamma_\vartheta} p_\vartheta \mathrm{d}\vartheta = \Gamma - |J| , \\
I_r &= \frac{1}{2\pi} \oint_{\gamma_r} p_r dr = -\Gamma + k\sqrt{-\frac{m}{2E}} .
\end{aligned}
\tag{3.56}
$$

The explicit expression as a function of the canonical coordinate is readily found by replacing the expressions of J, Γ, and E in (3.52).

3.6.3 Delaunay variables

By a straightforward inversion of the third part of (3.56) we calculate the Hamiltonian as

$$
H = -\frac{mk^2}{2\left(I_r + I_\vartheta + |I_\varphi|\right)^2} .
\tag{3.57}
$$

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 coincide, 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, Θ defined by the linear transformation

$$
\begin{aligned}
L &= I_r + I_\vartheta + |I_\varphi| , \\
G &= I_\vartheta + |I_\varphi| , \\
\Theta &= |I_\varphi| .
\end{aligned}
\tag{3.58}
$$

It is immediately clear that G and Θ coincide with Γ and J, respectively. Since the transformation is performed via a unimodular matrix, the corresponding transformation on the angles preserves the periods, as stated by Lemma 3.33. The Hamiltonian in Delaunay's variables takes the form

$$H = -\frac{mk^2}{2L^2} \,. \tag{3.59}$$

Denoting by ℓ, g, h the angles conjugated to the actions G, G, Θ, we can write Hamilton's equations as

$$\dot{\ell} = \frac{mk^2}{L^3} \,, \quad \dot{g} = \dot{h} = \dot{L} = \dot{G} = \dot{\Theta} = 0 \,.$$

Thus, the motion is periodic with a single frequency

$$\omega(L) = \frac{mk^2}{L^3} \,.$$

3.6.4 Construction of the Angle Variables

The canonical transformation should now be completed by constructing the angle variables associated to the actions L, G and Θ. To this end we must first write the generating function

$$\begin{aligned} S &= \int (p_r dr + p_\vartheta d\vartheta + p_\varphi d\varphi) \\ &= \int \sqrt{-\frac{m^2k^2}{L^2} + \frac{2mk}{r} - \frac{G^2}{r^2}}\, dr + \int \sqrt{G^2 - \frac{\Theta^2}{\sin^2\vartheta}}\, d\vartheta + \int \Theta d\varphi \,. \end{aligned}$$

The angle variables are then given by

$$\begin{aligned} \ell &= \frac{\partial S}{\partial L} = \frac{m^2k^2}{L^3} \int \frac{dr}{\sqrt{-\frac{m^2k^2}{L^2} + \frac{2mk}{r} - \frac{G}{r^2}}} \,, \\ g &= \frac{\partial S}{\partial G} = G \int \frac{d\vartheta}{\sqrt{G^2 - \frac{\Theta^2}{\sin^2\vartheta}}} - G \int \frac{dr}{r^2\sqrt{-\frac{m^2k^2}{L^2} + \frac{2mk}{r} - \frac{G^2}{r^2}}} \,, \\ h &= \frac{\partial S}{\partial \Theta} = -\Theta \int \frac{d\vartheta}{\sin^2\vartheta\sqrt{G^2 - \frac{\Theta^2}{\sin^2\vartheta}}} + \int d\varphi \,. \end{aligned}$$

Thus, the calculation of the angle variables is reduced to a quadrature; the actual calculation presents minor differences with respect to the case of the integrals (3.55).

It may also be useful to recall the relation between the Delaunay actions and the so called *orbital elements*. Without entering into the details, here are the relations:

$$L = \sqrt{mka} \,, \quad G = L\sqrt{1-e^2} \,, \quad \Theta = G\cos\iota \,, \tag{3.60}$$

where a is the semimajor axis, e is the eccentricity, and ι is the inclination of the orbital plane. Concerning the conjugated angles: ℓ is the so-called *mean*

anomaly, namely an angle which evolves uniformly, thus averaging in some sense the *true anomaly*, which is the angle giving the actual position of the planet with respect to the Sun; the angles g and h are the *longitude of the perihelion* and the *longitude of the node*, respectively.

3.7 The Linear Chain

In the first half of the eighteenth century, a strong discussion arose about the dynamics of a string with fixed ends, for example, a string of a musical instrument such as a harpsichord. Among the mathematicians involved in the discussion we find Daniel Bernoulli and Jean le Rond D'Alembert. It was in this connection that D'Alembert discovered the string equation known under his name and found the solution which describes the wave propagation along an infinite string. The problem relates to the one of propagation of sound, already raised by Newton ([182], Liber II, sect. VIII). In the case of a string with fixed ends, the problem under discussion was whether the solution could be written as a superposition of sinusoidal stationary waves with shape $\sin \frac{k\pi x}{L}$, where L is the length of the string (see [134], § 25).

Lagrange had written two long memoirs on the subject of propagation of sound [132][133]. Concerning the problem of the string with fixed ends, he exploited Newton's idea of investigating the dynamics of particles. In a further memoir [134] he started his investigations introducing a discrete approximation of the continuous string, representing it as a system of $N+2$ particles on a line, with the two particles at the ends kept fixed and subject to a first-neighbours interaction, as represented in Fig. 3.11. The simplifying hypothesis is assumed that the particles move orthogonally to the rest line. Moreover, the interaction between two neighbouring particles is uniform along the whole chain. Therefore, by symmetry, the equilibrium configuration is the one with the jth particle placed at distance $\frac{jL}{N+1}$ from, say, the left end, where L is the length of the string, and the configuration of the chain is determined by the vertical displacement x_j. Thus, denoting by y_j the canonical momentum conjugated to x_j, the equations may be written as a canonical system with Hamiltonian

$$H(x,y) = \frac{1}{2}\sum_{j=1}^{N} y_j^2 + \sum_{j=0}^{N} V(x_{j+1} - x_j)\ , \quad x_0(t) = x_{N+1}(t) = 0\ . \tag{3.61}$$

The simplest choice, made by Lagrange, consists in assuming an elastic interaction, namely substituting $V(r) = kr^2/2$ in Eq. (3.61), with a constant k which we may set to one, for simplicity. We are thus led to study the quadratic Hamiltonian

ALESSANDRO MORBIDELLI

« Modern Celestial Mechanics »

ii) They will have only one degree of freedom, i.e. $\mathcal{H}(v, x)$. In this case the Hamiltonian system is integrable because it has one constant of motion, that is the Hamiltonian itself. The motion will then evolve along level curves of $\mathcal{H}$ on the two-dimensional phase space (v, x).

iii) They will depend only on one coordinate, i.e. $\mathcal{H}(v_1, \ldots, v_n, x_k)$. In this case the system is integrable because it has n independent constants of motion which are $v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n$ and $\mathcal{H}$. The motion evolves preserving the value of $v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n$, and follows level curves of $\mathcal{H}$ on the plane (v_k, x_k).

1.9 Action–angle variables

For integrable Hamiltonian systems, of crucial importance is the *Arnold–Liouville theorem*, an extension of Liouville's theorem (see Section 1.8) found by Arnold (1963a). Arnold proved that, in the hypotheses of Liouville's theorem and if the n-dimensional surface implicitly defined by the constants of motion $\Phi_1, \ldots, \Phi_n$ is *compact*, it is then possible to introduce canonical momenta $\mathbf{p}$ and coordinates $\mathbf{q}$ such that

i) The coordinates $q_1, \ldots, q_n$ are *angles*, cyclically defined on the interval $[0, 2\pi]$, and the canonical transformation from the original momenta and coordinates, i.e. $\mathbf{v}(\mathbf{p}, \mathbf{q}), \mathbf{x}(\mathbf{p}, \mathbf{q})$, is 2π-periodic on the angles $q_1, \ldots, q_n$.

ii) In the new variables, the Hamiltonian is a function of the momenta $\mathbf{p}$ only, i.e. $\mathcal{H} \equiv \mathcal{H}(\mathbf{p})$.

The momenta $\mathbf{p}$ are usually called the *actions* of the system. A set of canonical variables $(\mathbf{p}, \mathbf{q})$, where the coordinates $\mathbf{q}$ are angles will be generically called *action–angle variables*. Although action–angle variables had been previously used by several authors (Epstein, 1916; Sommerfeld, 1922; Born, 1927) on specific problems, the Arnold–Liouville theorem is very important because it shows that basically any integrable Hamiltonian can be written, in suitable action–angle variables, as a function of the sole actions. Therefore, in the light of the Arnold–Liouville theorem, one can generically represent the integrable Hamiltonians by functions $\mathcal{H}(\mathbf{p})$, and work out the general theory of quasi-integrable Hamiltonian dynamics in action–angle variables – as will be done hereafter in this book.

Moreover, the proof of the Arnold–Liouville theorem also provides a constructive recipe for the introduction of action–angle variables in practical cases.

The existence of n constants of motion for an n-degree of freedom Hamiltonian system ensures that the motion evolves on an n-dimensional surface M_Φ embedded in $2n$-dimensional phase space. The fact that $\{\Phi_i, \Phi_j\} = 0$ for $i \neq j$ ensures that the motion can be decomposed in n independent flows, generated by the functions $\Phi_1, \ldots, \Phi_n$, each considered as a one-degree of freedom Hamiltonian. This means that the evolution of the motion at time t, i.e. $\mathbf{v}(t), \mathbf{x}(t)$, can be obtained following the flow of Φ_1 for a time t, from the initial condition $\mathbf{v}(0), \mathbf{x}(0)$ to a point $\mathbf{v}_1, \mathbf{x}_1$, then following the flow of Φ_2 for a time t, from $\mathbf{v}_1, \mathbf{x}_1$ to another point $\mathbf{v}_2, \mathbf{x}_2$, and so on. The final point $\mathbf{v}_n, \mathbf{x}_n$ will coincide with $\mathbf{v}(t), \mathbf{x}(t)$. The condition that the surface M_Φ is compact implies that the individual flows of $\Phi_1, \ldots, \Phi_n$, and hence the global motion, can be decomposed into independent periodic cycles, which we denote by $\gamma_1, \ldots, \gamma_n$. The actions $\mathbf{p}$ are then introduced by

$$p_i = \frac{1}{2\pi} \oint_{\gamma_i} \sum_{j=1}^{n} v_j \mathrm{d}x_j \ . \tag{1.55}$$

Then, writing $\mathbf{v}$ as functions of $\mathbf{p}$ and $\mathbf{x}$, the integral generating function

$$S(\mathbf{p}, \mathbf{x}) \equiv \int \sum_{j=1}^{n} v_j(\mathbf{p}, \mathbf{x}) \mathrm{d}x_j \tag{1.56}$$

is defined, and the new coordinates $\mathbf{q}$ are introduced as

$$q_i = \frac{\partial S}{\partial p_i}(\mathbf{p}, \mathbf{x}) \ . \tag{1.57}$$

The transformation $(\mathbf{v}, \mathbf{x}) \to (\mathbf{p}, \mathbf{q})$ so defined is of the form (1.41) and therefore is, by construction, canonical. One can prove that $q_1, \ldots, q_n$ are angles, namely q_i is increased by 2π when a complete cycle γ_i is followed, and that the Hamiltonian $\mathcal{H}$ is dependent on the actions $\mathbf{p}$ only (see Arnold, 1963a).

1.9.1 Delaunay variables

As an example of the application of the Arnold–Liouville theorem, let's proceed to introduce the action–angle variables for the integrable Hamiltonian of the two-body problem. They will be the variables that we will later use to study the dynamics of the restricted problem and of the planetary problem using Hamiltonian perturbation techniques. We follow here the approach of Born (1927), elaborated for the equivalent problem of classical motion of an electron around the core of the hydrogen atom.

NOTA 4

MOSER & ZEHNDER, « Notes on Dynamical Systems »

Per introdurre le variabili coniugate a C e a D consideriamo i flussi integrali associati ai campi vettoriali hamiltoniani X_C e X_D:

$$\Phi_C^t = \exp(tX_C)$$

$$\Phi_D^t = \exp(tX_D).$$

Cominciamo da Φ_C^t. In termini delle variabili canoniche $(p,q) \in \mathbb{R}^6$, dove $q \in \mathbb{R}^3$ è il vettore posizione calcolato rispetto ad un riferimento inerziale e $p \in \mathbb{R}^3$ rappresenta il vettore dei momenti coniugati alle componenti di q, si ha

$$C = \sqrt{(p_1 q_2 - p_2 q_1)^2 + (p_1 q_3 - p_3 q_1)^2 + (p_2 q_3 - p_3 q_2)^2}.$$

La coppia di vettori indipendenti p, q definisce un piano E che rimane fisso nel tempo. Inoltre $p(t), q(t)$ sono ottenuti da $p(0), q(0)$ attraverso una rotazione su questo piano di un angolo φ. Per dimostrare questa affermazione, applichiamo a (p,q) una rotazione data dalla trasformazione canonica

$$(p,q) \longmapsto (Rp, Rq)$$

in modo da far coincidere il piano orbitale con il piano di riferimento

$$RE = \text{span}(e_1, e_2).$$

Chiamando $(\tilde{p}, \tilde{q})$ le nuove variabili si ha

$$C = \pm (\tilde{q}_1 \tilde{P}_2 - \tilde{q}_2 \tilde{P}_1)$$

che a meno del segno definisce le equazioni differenziali

$$\begin{cases} \dot{\tilde{q}}_1 = -\tilde{q}_2 \\ \dot{\tilde{q}}_2 = \tilde{q}_1 \end{cases} \qquad \begin{cases} \dot{\tilde{P}}_1 = -\tilde{P}_2 \\ \dot{\tilde{P}}_2 = \tilde{P}_1 \end{cases}$$

$$\downarrow \qquad\qquad \downarrow$$

$$\ddot{\tilde{q}}_2 = -\tilde{q}_2 \qquad\qquad \ddot{\tilde{P}}_2 = -\tilde{P}_2$$

Passando ora a Φ_D^t, dalla discussione appena fatta si vede che D rappresenta la hamiltoniana di una rotazione attorno ad e_3.

Possiamo scegliere come coordinata coniugata a C l'argomento del pericentro ω e come coordinata coniugata a D la longitudine del nodo Ω.