PERTURBATION METHODS

IN CELESTIAL MECHANICS

by Andrea Milani

Dipartimento di Matematica, Università di Pisa and
Department of Astronomy, Cornell University

E–mail *TWIN2@ICNUCEVM.BITNET*

CHAPTER 1:

CANONICAL TRANSFORMATIONS

This chapter discusses dynamical equations of the Hamilton type:

$$\frac{dq}{dt} = \frac{\partial H}{\partial p} \quad ; \quad \frac{dp}{dt} = -\frac{\partial H}{\partial q} \tag{1.1}$$

(where q, p are N–vectors) and the way they can be transformed by coordinate changes which preserve the hamiltonian character of the equations. No specific application is presented in this chapter; the theory is developed only to the extent which will later be needed to discuss the perturbation theories actually used in Celestial Mechanics, and computational tools are readied for such theories.

1.1 LOCAL CANONICAL TRANSFORMATIONS

A coordinate transformation

$$\Phi : (q, p) \longmapsto (x, y) \tag{1.2}$$

is canonical if the equations of motion (1.1), as seen in the new coordinate system (x, y), are again hamiltonian with as hamiltonian $K = H \circ \Phi^{-1}$, the function with the same values in corresponding points, that is:

$$\frac{dx}{dt} = \frac{\partial K}{\partial y} \quad ; \quad \frac{dy}{dt} = -\frac{\partial K}{\partial x} \tag{1.3}$$

$$K(\Phi(q, p)) = H(q, p) \tag{1.4}$$

The condition for a map to be locally canonical, in the neigbourhood of some point, can be simply stated in terms of the matrix of partial derivatives:

$$D\Phi = \frac{\partial(x, y)}{\partial(q, p)} \tag{1.5}$$

To simplify the formulas, we assemble the coordinates q and the momenta p in a single $2N$–vector $r = (q, p)$ and rewrite the Hamilton equations:

$$\frac{dr}{dt} = J \nabla H^T \tag{1.6}$$

with J a $2N \times 2N$ matrix:

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{1.7}$$

and $\nabla H = \partial H/\partial(q,p)$ the gradient of the Hamiltonian; ∇^T is the gradient transposed, that is used as a column vector. The time derivative of $z = (x,y)$ can be computed by the chain rule:

$$\frac{dz}{dt} = D\Phi \, \frac{dr}{dt} = D\Phi \cdot J \, \nabla H^T \tag{1.8}$$

The gradient can also be computed by the chain rule from (1.4):

$$\nabla H = \nabla K \, D\Phi \tag{1.9}$$

and upon substitution into (1.8):

$$\frac{dz}{dt} = \left[D\Phi \cdot J \cdot D\Phi^T \right] \nabla K^T \tag{1.10}$$

which is equivalent to the new Hamilton equations $dz/dt = J\nabla K^T$ provided:

$$D\Phi \cdot J \cdot D\Phi^T = J \tag{1.11}$$

If (1.11) holds in every point, the map Φ is *locally canonical*. Since J is non–singular, (1.11) also implies that $det(D\Phi) = \pm 1$. In the case of a 2×2 matrix, $det(D\Phi) = 1$ is equivalent to (1.11); for larger matrices, this is not the case.

It is easy to check wether condition (1.11) is verified for any given Φ, but it is not easy to find non–trivial canonical maps, even locally. A standard procedure is the use of a *generating function* $S = S(q,y)$ to define a map in an implicit way:

$$p = \frac{\partial S}{\partial q}(q,y) \quad ; \quad x = \frac{\partial S}{\partial y}(q,y) \tag{1.12}$$

If the implicit definition (1.12) can be untangled in a well defined map (1.2), at least locally, then such a map is locally canonical (this can be shown by the direct computation of (1.11) by means of the implicit function theorems, and also proven indirectly by using variational principles; see Arnold, 1976, chap. 9). However all the topological difficolties are somewhat hidden in the implict map (1.12) and emerge in the attempt to make it explicit, so this method is often of little help in understanding the global properties of the coordinate change.

A further problem arises when we need to consider a time dependent canonical transformation. It is indeed possible to develop a separate formalism of *contact transformations* (see Abraham and Marsden, 1967; Arnold, 1976, App. 4), but this is not really needed for the applications to Celestial Mechanics. The simplest way is the so–called *homogenous formalism*, by which the time is considered as an extra coordinate $t = q_0$ and a new momentum p_0 is added. Then the Hamilton equations with a (possibly time dependent) Hamiltonian $H = H(q, t, p)$ are handled as time–independent Hamilton equations with the homogeneous Hamiltonian $\mathcal{H} = H(q, q_0, p) + p_0$ with the equations (1.1) plus the extra ones:

$$\frac{dq_0}{dt} = \frac{\partial \mathcal{H}}{\partial p_0} = 1 \quad ; \quad \frac{dp_0}{dt} = -\frac{\partial \mathcal{H}}{\partial q_0} = -\frac{\partial H}{\partial t} \tag{1.13}$$

where the last equation shows that the meaning of p_0 is simply $p_0 = -H$. Then a canonical transformation can be defined by a time–dependent generating function:

$$S = S(q, t, p) \quad ; \quad x = \frac{\partial S}{\partial y} \quad ; \quad p = \frac{\partial S}{\partial q} \tag{1.14}$$

just by thinking to the corresponding homogeneous generating function:

$$\mathcal{S} = \mathcal{S}(q, q_0, y, y_0) = S(q, q_0, p) + y_0 q_0 \tag{1.15}$$

constructed in such a way that $x_0 = \partial \mathcal{S}/\partial y_0 = q_0 = t$; the new Hamiltonian is $-y_0$, related to $H = -p_0$ by:

$$p_0 = \frac{\partial \mathcal{S}}{\partial q_0} = y_0 + \frac{\partial S}{\partial t} \tag{1.16}$$

that is the new Hamiltonian is $K = H - \partial S/\partial t$. The homogeneous formalism can also be used to allow for changes in the independent variable, which is often useful.

1.2 RUDIMENTS OF TOPOLOGY

To understand the properties of global canonical maps we need to borrow some concepts from differential topology. Two spaces are said to be *topologically equivalent* if there is a map between them which is one–to–one and differentiable, with the inverse map differentiable as well; such a map is called a *diffeomorphism*. Spaces with quite different metric properties can be equivalent, an example being the equivalence of the half line $\mathcal{R}^+$ with the real line $\mathcal{R}$. There are however non equivalent spaces, such as the real line $\mathcal{R}$ and the circle S^1 (e.g. it can be shown that the latter is compact and the former is not).

Another concept from differential topology widely used in the theory of Hamiltonian system is that of a *local diffeomorphism*, a map which is locally a diffeomorphism in the neigbourhood of each point; by the inverse function theorem, this is equivalent to the matrix of partial derivatives of the map being non–singular at each point. It must be stressed that a local diffeomorphism can exist even between spaces which are not toplogically equivalent, the simplest example being a map $\mathcal{R} \longrightarrow S^1$ such as the one which is constructed whenever angles are measured with numbers, e.g. in radians. The notion of an *angle variable*, which is simply a copy of the map $\mathcal{R} \longrightarrow \mathcal{R}/(2\pi\mathcal{Z}) \simeq S^1$ in which two values of the angular measure are said to define the same angle if they differ by an integer multiple of 2π, is central to the study of integrable dynamical systems.

A third useful concept is that of a diffeomorphism *isotopic to the identity*, or *trivial*; it refers to a map of a space onto itself which not only is a diffeomorphism, but also can be obtained by differentiable deformation from the identity map. By no means all diffeomorphisms are trivial, the simplest example being the map $\theta \longmapsto -\theta$ of the real line onto itself (if θ is regarded as an angle variable, the same formula defines a diffeomorphism of S^1 which is also non trivial).

So far we have only used examples with 1–dimensional spaces, and of course the variety of the possible topological structures increases dramatically with the dimension of the space. However, for the purpose of the study of integrable Hamiltonian systems and their small perturbations we can limit ourselves to the consideration of spaces which are either of dimension 1 or cartesian products of 1–dimensional spaces. Since the cartesian operator preserves topological equivalence, that is $A \simeq B$ and $C \simeq D$ implies $A \times C \simeq B \times D$ (with $\simeq$ used as a symbol for topological equivalence), the classification of such products of simple spaces is easy: e.g. in dimension 2 we can have only the plane $\mathcal{R} \times \mathcal{R} = \mathcal{R}^2$, the cylinder $\mathcal{R} \times S^1$ and the *torus* $S^1 \times S^1 = T^2$; in dimension N we may encounter the *N–torus* $S^1 \times T^{N-1} = T^N$, etc..

However, even simple product spaces such as tori can have subtle topological properties. Let us consider as an example –which is chosen because it will be needed later in the study of resonances– the classification of the non–trivial diffeomorphisms of an N–torus. Let us assume the torus T^N to be defined by N angle variables $\theta = (\theta_1, \theta_2, \ldots, \theta_N)$; when does a linear transformation of $\mathcal{R}^N$:

$$\phi = A\theta \tag{1.17}$$

define a map of the torus onto itself? The answer is that this requires the coefficients of the matrix A to be integers, and can be understood by the simple example of the map $\theta_1 \longmapsto \frac{1}{2}\theta_1$, which results in $\theta_1 = 0$ and $\theta_1 = 2\pi$, the same angle, that is the same point on the circle, mapped into two different angles 0 and π. The same map is also a diffeomorphism of T^N if the inverse matrix A^{-1} exists and has integer entries as well; this is possible if and only if A is an integer matrix with $det\, A = \pm 1$. Matrices with integer coefficients and determinant $+1$ are called *unimodular*.

Unimodular transformations have surprising geometric properties; the most important for our purposes is that no unimodular transformation (different form the

identity) of a torus is trivial, that is no one can be obtained from another one by a continous deformation. Thus there are infinite different ways to parametrize a torus which are not reducible to one another by means of a sequence of small changes, for every N–torus with $N > 1$.

Unimodular transformations can be, to same extent, taylor–made to perform specific adaptations to our computing needs. Let us take the simple example of an ordinary 2–torus: can we find a unimodular 2×2 matrix A such that one of the new angles ϕ defined by (1.17) is a given combination of θ_1 and θ_2:

$$\phi_1 = a\theta_1 - b\theta_2 \tag{1.18}$$

with a, b integers? The second component $\phi_2 = c\theta_1 + d\theta_2$ forms a unimodular transformation together with (1.18) provided that the integers c, d are such that $ad + bc = 1$; since we know from elementary algebra that the equation $ad + bc = MCD(a, b)$ has always a solution with integers c, d, the only constraint in the choice of one of the new angles by an equation like (1.18) is that $MCD(a, b) = 1$. We shall see later how such a choice can allow an important step forward in the theory of resonances.

1.3 GLOBAL CANONICAL MAPS

A *global canonical transformation* can be defined as a diffeomorphism which is also locally canonical, that is fulfills (1.11) in every point. However, this definition is too restrictive. Canonical maps are often used for the very purpose of changing the topology, and in this case they must either be singular somewhere, or not be globally one–to–one, or both.

The simplest examples of global canonical maps can be obtained with the generating function formalism. Let us suppose we have chosen a transformation of the coordinates q: $x = f(q)$ and wish to find a canonical transformation which extends this map to the momenta: this can be done by means of the generating function:

$$S = y \cdot f(q) \quad ; \quad x = \frac{\partial S}{\partial y} = f(q) \quad ; \quad p = \frac{\partial S}{\partial q} = Df(q)^T\, y \tag{1.19}$$

That is, if f is a local diffeomorphism ($det\, Df \neq 0$):

$$x = f(q) \quad ; \quad y = \left[Df(q)^T\right]^{-1} p \tag{1.20}$$

and the momenta are transformed in a *covariant* way. However (1.20) is really a global canonical transformation only if f is itself a diffeomorfism. This is always the

case when $f(q) = B\,q$ is a linear map (with $det\,B \neq 0$):

$$S = y \cdot B\,q \quad ; \quad x = B\,q \quad ; \quad y = \left[B^{-1}\right]^T p \tag{1.21}$$

Another interesting example is:

$$S = y \cdot R_{\omega t}\,q \quad ; \quad x = R_{\omega t}\,q \quad ; \quad y = R_{\omega t}\,p \tag{1.22}$$

where $R_{\omega t}$ is the rotation by an angle $|\omega|t$ aroung the axis defined by the vector ω of $\mathcal{R}^3$. This case must be handled by means of the time dependent (or homogeneous) formalism, and the new Hamiltonian can be computed by means of the derivative:

$$\frac{\partial R_{\omega t}}{\partial t}\,q = R_{\omega t}\,(\omega \times q) \tag{1.23}$$

where the vector product is the standard way to represent an infinitesimal rotation. By using (1.23), and taking into account the rotational invariance of scalar and vector products, the new Hamiltonian is found to be:

$$K = H - \frac{\partial S}{\partial t} = H - \omega \cdot (q \times p) \tag{1.24}$$

Equations (1.22) and (1.24) allow to derive the equations for the circular restricted 3–body problem in the rotating frame (in which the primaries are fixed); the last of (1.22) shows that the momenta in the rotating frame are not the velocities in the rotating frame but the rotated velocities in the inertial frame.

The more interesting cases are when the topology actually changes. The most obvious example is the change to polar coordinates:

$$S = y_1 r \cos\theta + y_2 r \sin\theta \tag{1.25}$$

which is of course locally canonical only for $r > 0$, that is for (r, θ) in $\mathcal{R}^+ \times \mathcal{R}$ and with values (x_1, x_2) in $\mathcal{R}^2 - (0,0) \simeq \mathcal{R}^+ \times S^1$. Since the two spaces are not topologically equivalent, the map cannot be one–to–one and indeed θ is an angle variable.

Another very useful non regular canonical map is the transformation to *Poincaré variables*, the canonical analogue of the polar to cartesian transformation when a coordinate x and its conjugate momentum y are regarded as the couple of cartesian coordinates:

$$x = \sqrt{2p}\,\cos(-q) \quad ; \quad y = \sqrt{2p}\,\sin(-q) \tag{1.26}$$

It is easy to directly check that $det\,\partial(x,y)/\partial(q,p) = 1$, with the minus sign in front of the coordinate q playing a surprisingly essential rôle. (1.26) could be locally

defined by a generating function $S = \frac{1}{2}y^2\, cotg\, q$, but there is no way to have a single generating function defining the map globally, precisely because the topology does change. This canonical map is used to remove the singularity of the angle q for $p = 0$, and viceversa to introduce the angle variable when it is needed.

With canonical maps as a tool, integrable hamiltonian systems can be explicitly solved by reduction to a trivial form; this is performed by a sequence of canonical maps, some of which typically do change the topology. The simplest examples can be found with quadratic Hamiltonians: let us assume the Hamiltonian function is of the form:

$$H = \frac{1}{2}(x \cdot A\, x + y \cdot A\, y) \tag{1.27}$$

This particular form arises from symmetry properties of the Hamiltonian, as we shall see in Section 4.1. The symmetric matrix $A = A^T$ has real eigenvalues $\nu_k,\ k = 1, \ldots, N$ and can be diagonalised by means of an orthogonal matrix $B = (B^{-1})^T$:

$$B^T\, A\, B = diag[\nu_1, \nu_2, \ldots, \nu_N] \tag{1.28}$$

and by the canonical linear change $x = Bw,\ y = Bz$ the Hamiltonian is reduced to the form:

$$H = \frac{1}{2}\sum_{k=1}^{N} \nu_k\, (w_k^2 + z_k^2) \tag{1.29}$$

which represents a set of N uncoupled linear oscillators with frequencies ν_k. Each oscillator problem can be solved by introducing angle variables with the inverse of (1.26):

$$w_k = \sqrt{2\Theta_k}\, \cos(-\theta_k) \quad ; \quad z_k = \sqrt{2\Theta_k}\, \sin(-\theta_k) \tag{1.30}$$

Then the Hamiltonian is reduced to the form:

$$H = \sum_{k=1}^{N} \nu_k\, \Theta_k \tag{1.31}$$

with the *angle variables* θ_k changing linearly with time, each one with its own frequency ν_k, and the *action variables* Θ_k constants of the motion. This is the simplest example of an integrable Hamiltonian system, and it already shows most qualitative features of such systems: the level manifolds with all the Θ_k constant are spanned by N angle variables, hence they are N–tori T^N; most of the phase space is topologically equivalent to $T^N \times (R^+)^N$, apart from the subsets with some $\Theta_k = 0$ where lower dimensional tori occur.

Non trivial canonical transformations can be used to somewhat lessen the difficulties arising from simple singularities such as the lower dimensional tori of the

above example. Suppose we wish to remove the singularity arising for $\Theta_1 = 0$, where θ_1 is undefined. Then the unimodular matrix :

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{1.32}$$

can be used to combine the angles θ_1 and θ_2:

$$\phi_1 = \theta_1 + \theta_2 \quad ; \quad \phi_2 = \theta_2 \tag{1.33}$$

and by (1.21) the matrix $(A^{-1})^T$ gives the corresponding change in the action variables:

$$\Phi_1 = \Theta_1 \quad ; \quad \Phi_2 = -\Theta_1 + \Theta_2 \tag{1.34}$$

The advantage of the new coordinates is in that ϕ_1 can be defined as θ_2 for $\Theta_1 = 0$, and this definition is regular for $\Theta_1 \longmapsto 0$, that is the map from the action–angle variables (θ, Θ) to the Poincaré type variables (w, z) is smooth around $(w_1, z_1) = (0, 0)$. This simple trick is standard in Celestial Mecahnics, e.g. when the longitudes λ, ϖ, Ω are used instead of the Delaunay variables ℓ, ω, Ω.

Another use of non trivial global canonical transformations is to highlight resonances. Suppose that in the example (1.29)–(1.31) above two of the frequencies are in a rational ratio, e.g. :

$$\frac{\nu_1}{\nu_2} = \frac{b}{a} \tag{1.35}$$

Then a unimodular transformation with first component (1.18) can be constructed (since we can assume $MCD(a, b) = 1$ anyway), with matrix:

$$A = \begin{pmatrix} a & -b \\ c & d \end{pmatrix} \; ; \qquad det\, A = 1 \tag{1.36}$$

The actions are changed by the matrix $(A^{-1})^T$:

$$\Phi_1 = d\Theta_1 - c\Theta_2 \quad ; \quad \Phi_2 = b\Theta_1 + a\Theta_2 \tag{1.37}$$

and the Hamiltonian, containing Θ_1 and Θ_2 only in the combination $\nu_1\,\Theta_1 + \nu_2\,\Theta_2 = \nu_1\Phi_2/b$, does not depend upon Φ_1, thus explicitly showing that ϕ_1 is constant.

To summarize our view of canonical transformations, canonical maps which are far from the identity (either because they are non trivial, or because they have singularities, or because they are not one–to–one) are mostly used to set up the appropriate topology of the phase space and a qualitatively suitable parametrisation. Fine tuning of the canonical coordinate system can then be performed by near–identity transformations.

CHAPTER 2:

SMALL TRANSFORMATIONS, SMALL PERTURBATIONS

This chapter discusses canonical transformations which are *near identiy*, that is each point in the phase space is displaced only by a small amount, and their use to solve problems with *small perturbations*, that is with Hamiltonians containing a *small parameter* ε:

$$H = H_0 + \varepsilon H_1 + \varepsilon^2 H_2 + \dots \tag{2.1}$$

In principle a small transformation could be defined by a generating function close to the one of the identity transformation:

$$S = q \cdot y + \varepsilon S_1(q,y) + \varepsilon^2 S_2(q,y) + \dots \tag{2.2}$$

and this formalism has often been successfully used; however another formalism leads to easier computations, especially when the theory needs to be computed to higher order. It is based upon *canonical flows*, that is transformations F^s which form a (local) one–parameter group , with $F^0 = Identity$ and F^s a global canonical map, for s small enough; moreover, $F^s \circ F^z = F^{s+z}$ whenever defined.

2.1 INTEGRAL FLOWS AND VARIATIONAL EQUATIONS

For a given Hamiltonian $\chi(q,p)$, with Hamilton equations for $r = (q,p)$:

$$\frac{dr}{ds} = J \cdot \nabla\chi(r) \tag{2.3}$$

the solutions of all the initial conditions problems can be put together in the *integral flow*, that is the map:

$$F_\chi : (s,r) \longmapsto F^s_\chi(r) \tag{2.4}$$

such that –for fixed r_0– $F^s_\chi(r_0)$ is the solution of (2.3) with initial conditions $r = r_0$ at $s = 0$; that is, F^s_χ satisfies the initial conditions problem:

$$\frac{\partial}{\partial s} F^s_\chi(r) = J \cdot \nabla\chi(F^s_\chi(r)) \quad ; \quad F^0_\chi(r) = r \tag{2.5}$$

The existence and uniqueness theorem for the initial conditions problem ensures that F^s_χ is a local one–parameter group. Moreover the regularity theorem for solutions of ordinary differential equations ensures that F is at least as smooth as χ is, and in all

the variables. Let us now consider F^s_χ, for a fixed value of the independent variable s, as a map of the phase space into itself, sending each initial condition onto the state of the corresponding solution after "time" s has elapsed. Such a map is also differentiable; let the matrix of partial derivatives be A^s:

$$A^s(r) = \frac{\partial}{\partial r} F^s_\chi(r) \tag{2.6}$$

Then $A^s(r)$ satisfies another differential equation, the *variational equation*, which can be obtained by taking the derivatives with respect to the initial conditions r from both sides of (2.5):

$$\frac{\partial}{\partial r}\frac{\partial}{\partial s} F^s_\chi(r) = \frac{\partial}{\partial s}\frac{\partial}{\partial r} F^s_\chi(r) = \frac{\partial}{\partial s} A^s(r)$$

where use was made of the possibility of exchanging the order of the derivatives for a smooth map, and of the definition (2.6);

$$\frac{\partial}{\partial r} J \cdot \nabla\chi(F^s_\chi(r)) = J \cdot \nabla\nabla\chi \cdot \frac{\partial}{\partial r} F^s_\chi(r) = J \cdot \nabla\nabla\chi \cdot A^s(r)$$

where the chain rule for derivatives has been used, and $\nabla\nabla\chi$ is the symmetric matrix of the second derivatives, computed along the orbit starting at r. Thus $A^s(r)$, for a fixed r, is the solution of the initial condition problem for a linear time–dependent equation:

$$\frac{\partial}{\partial s} A^s(r) = J \cdot \nabla\nabla\chi \cdot A^s(r) \quad ; \quad A^0(r) = I \tag{2.7}$$

The *variational equations* (2.7) are themselves Hamiltonian: $A^s(r)$ is the matrix solution of the linear equations defined by the quadratic and time–dependent Hamiltonian $\frac{1}{2} v \cdot \nabla\nabla\chi(F^s_\chi(r))v$. The main property of the solution of (2.7) is that it always satisfies the condition (1.11): let

$$C^s = A^s \cdot J \cdot [A^s]^T$$

then C is solution of an initial condition problem:

$$\frac{\partial}{\partial s} C^s = (J \cdot \nabla\nabla\chi \cdot A^s) \cdot J \cdot [A^s]^T = J \cdot \nabla\nabla\chi \cdot C^s - C^s \cdot \nabla\nabla\chi \cdot J$$

$$C^0 = A^0 \cdot J \cdot [A^0]^T = J$$

which has the trivial solution $C^s = J$; by uniqueness, A^s always fulfills (1.11) and F^s_χ is locally canonical. It can be shown that F^s_χ is actually globally canonical provided it is globally defined. Moreover, the transformations defined in this way are topologically trivial, since they are deformations of the identity corresponding to a time zero flow.

2.2 LIE SERIES

Thus any function $\chi = \chi(q,p)$ defines a canonical flow; once a value of the associated independent variable s is chosen, a canonical tranformation is uniquely defined. However, to explicitly compute the transformation we have to solve the differential equations (2.3)–(2.5). The simplest procedure to compute some approximation to this solution is the Taylor–Mac Laurin formula at $s = 0$:

$$r' = F_\chi^s(r) = r + s\left[\frac{\partial}{\partial s}F_\chi^s(r)\right]\Big|_{s=0} + \frac{s^2}{2}\left[\frac{\partial^2}{\partial s^2}F_\chi^s(r)\right]\Big|_{s=0} + \dots \tag{2.8}$$

A similar formula can be used to compute the transform of any function of the canonical variables $g = g(q,p)$; let us denote the map between the function spaces by T_χ^s:

$$T_\chi^s : g \longmapsto g' = g \circ \left[F_\chi^s\right]^{-1}$$
$$(T_\chi^s g)(r') = g \circ \left[F_\chi^s\right]^{-1}(r') = g(F_\chi^{-s}(r')) \tag{2.9}$$

Two important remarks on equation (2.9). The transformation of a function is performed by composition with the inverse map; this is the same rule used to transform the Hamilton functions in a canonical map (see Section 1.1). Since the initial and the final state are exchanged when the sign of the independent variable is changed –or, equivalently, when the sign of the function χ is changed– the inverse map can be obtained by a change in sign; this also follows from the group property of the integral flow: $F_\chi^s \circ F_\chi^{-s} = F_\chi^0 = Identity$. The Taylor formula for T_χ^s can be computed in essentially the same way:

$$\begin{aligned} g'(r') &= g(F_\chi^{-s}(r')) = \\ &= g(r') + s\left[\frac{\partial}{\partial s}g(F_\chi^{-s}(r))\right]\Big|_{s=0} + \frac{s^2}{2}\left[\frac{\partial^2}{\partial s^2}g(F_\chi^{-s}(r))\right]\Big|_{s=0} + \dots \end{aligned} \tag{2.10}$$

However, (2.10) does not look like an easy to use formula. It can became easy to use provided two conditions are met: 1) a simple method is available to compute the derivatives, including the higher ones, with all the chain rules; 2) the series converges rapidly, so that not too many terms have to be considered for an acceptable accuracy.

The first requirement is not too difficult to fulfill, especially for the case we are interested in, that is the canonical flows. The derivative of F_χ^s is provided by equation (2.5), in the form of Hamilton equations. A very compact notation for the total derivative dg/ds of any function g with respect to the independent variable s along the solutions of a set of Hamilton equations is provided by the *Poisson bracket*:

$$\{g,\chi\} = \frac{\partial g}{\partial q}\cdot\frac{\partial \chi}{\partial p} - \frac{\partial g}{\partial p}\cdot\frac{\partial \chi}{\partial q} = \frac{\partial}{\partial s}\left[g(F_\chi^s(r))\right]\Big|_{s=0} = \frac{dg}{ds} \tag{2.11}$$

When (2.11) is substituted into (2.10):

$$g'(r') = g(r') - s\{g,\chi\} + \frac{s^2}{2}\{\{g,\chi\},\chi\} - \frac{s^3}{6}\{\{\{g,\chi\},\chi\},\chi\} + \dots \tag{2.12}$$

where the double Poisson bracket arises from the need to compute the s derivative of the function $\{g,\chi\}$, etc.. The alternate signs account for the composition with the inverse map, and all the Poisson brackets on the right hand side have to be computed in the new coordinates $r' = (q', p')$.

The second requirement is not as trivial to fulfill. For now, we only remark that convergence can occur only if either s is small, or the flow generated by χ is slow –that is, χ is small. It turns out that the two conditions are one and the same: to multiply χ by a small parameter ε is equivalent to multiply s by $1/\varepsilon$, and we can either change χ into $\varepsilon\chi$ and set $s = 1$ or set $s = \varepsilon$ and leave χ as is; there is no use for two parameters. Since the choice is only a matter of taste, we chose to set $s = 1$, so that s disappears from formulas such as (2.12), and simplify the notation (e.g. $F^1_\chi = F_\chi,\ T^1_\chi = T_\chi$). We fulfill the smallness requirement by assuming that χ is of positive order in the *small parameter* ε:

$$\chi = \varepsilon\chi_1 + \varepsilon^2\chi_2 + \dots \tag{2.13}$$

thus the *Lie transform* of a function g is defined by the expansion (formal power series):

$$\begin{aligned} g' =& T_\chi g = g - \{g,\chi\} + \tfrac{1}{2}\{\{g,\chi\},\chi\} + \dots = \\ =& g - \varepsilon\{g,\chi_1\} + \varepsilon^2[-\{g,\chi_2\} + \tfrac{1}{2}\{\{g,\chi_1\},\chi_1\}] + \dots \end{aligned} \tag{2.14}$$

A further development occurs when g is itself expanded in powers of ε:

$$g = g_0 + \varepsilon g_1 + \varepsilon^2 g_2 + \dots \tag{2.15}$$

and the result of (2.14) is reordered by powers of ε:

$$\begin{aligned} g' =& T_\chi g = g_0 + \varepsilon[g_1 - \{g_0,\chi_1\}] + \\ &+ \varepsilon^2[g_2 - \{g_0,\chi_2\} - \{g_1,\chi_1\} + \frac{1}{2}\{\{g_0,\chi_1\},\chi_1\}] + \dots \end{aligned} \tag{2.16}$$

That is, the development of g' is:

$$\begin{aligned} g'_0 &= g_0 \\ g'_1 &= g_1 - \{g_0,\chi_1\} \\ g'_2 &= g_2 - \{g_0,\chi_2\} - \{g_1,\chi_1\} + \tfrac{1}{2}\{\{g_0,\chi_1\},\chi_1\} \\ &\dots \end{aligned} \tag{2.17}$$

It can be seen from (2.17) that the order k part of g, g_k, appears in all the g'_z with $z \geq k$; also the order k part of of χ appears in all the g'_z with $z \geq k$. This relationship can be represented by the *Lie triangle* (Deprit, 1969), and leads to an easy to use recursive formula.

We need to point out that there are many versions of this Lie series algorithm; the one we are going to use here is essentially the one due to Hori (1966). A more general algorithm, allowing for a time–dependent generating function, can also be used (Deprit, 1969). However, the larger group of transformations defined by allowing for a time–dependent generating function is not really needed in the applications to Celestial Mechanics discussed in the next Sections. The Hori method has also the advantage of a much easier computation of the inverse map by means of the same generating function with the opposite sign; this will play an important rôle in simplifying the computations in Sections 3.2 and 4.3.

2.3 ELIMINATION OF THE ANGLES

The main purpose of a Lie series transformation given by formulas such as (2.13)–(2.17) is to "solve" the problem, that is to transform the Hamiltonian (2.1) into a simpler one, whose solution can be somehow explicitly computed. Again we have to remember that this process does not allow to change the topology (not even to perform a non–trivial topological equivalence), thus the coordinate system must already be adapted to the problem in the topological sense. In most cases, the best coordinate system is such that the first approximation problem defined by the order zero Hamitonian H_0 is already solved:

$$H = H_0(p) + \varepsilon H_1(q,p) + \varepsilon^2 H_2(q,p) + \dots \tag{2.18}$$

However, this leaves little choice for the coordinates q_r (conjugate to the momenta p_r, which are integrals of H_0). On the basis of a general result (which we shall discuss later, see Section 3.1) all the q_r can be expected to be angle variables. Then the question is: can we choose χ in such a way that $T_\chi H = H'$ is solved, that is $H' = H'(p')$? We can look at the analogous of the formulas (2.17) as a recursive set of equations, where the H_r are given and the χ_r are to be solved for; the transformed Hamiltonian H' is not known, but the condition is imposed that $H' = H'(p')$:

$$\begin{aligned} H'_0(p') &= H_0(p') \\ H'_1(p') &= H_1(q',p') - \{H_0,\chi_1\}(q',p') \\ H'_2(p') &= H_2(q',p') - \{H_0,\chi_2\} - \{H_1,\chi_1\} + \tfrac{1}{2}\{\{H_0,\chi_1\},\chi_1\} \\ &\dots \end{aligned} \tag{2.19}$$

To write down explicitly the second of the (2.19) as an equation for χ_1, we compute the Poisson bracket $\{H_0, \chi_1\}$ by means of the vector n of the fundamental frequencies appearing in the solution of the integrable approximation H_0:

$$n_r = \frac{\partial H_0}{\partial p_r} \tag{2.20}$$

$$H_1 - \{H_0, \chi_1\} = H_1 + n \cdot \frac{\partial \chi_1}{\partial q'} = H_1'(p') \tag{2.21}$$

Has (2.21) a solution? To answer, we shall use Fourier series expansions; the analytical equivalent of the statement that the q_r are angles is that the Hamiltonian H can be expanded into a (convergent) multiple Fourier series: for the first order (in ε) part H_1:

$$H_1(q, p) = \sum_k H_{1k}(p) \cos(k \cdot q) \tag{2.22}$$

with k a multiindex, that is a vector with integer components. The presence of the cosine terms only arises from the assumption of a discrete symmetry, i.e. H is even, as it is often the case in Celestial Mechanics; of course in general there might be sine terms too.

A formal series solution is then obtained by a simple algorithm, assuming that $\chi(q, p)$ is defined on the same space, thus has a Fourier series expansion; for the first order part $\varepsilon\chi_1$:

$$\chi_1(q, p) = \sum_k \chi_{1k}(p) \sin(k \cdot q) \tag{2.23}$$

where again the presence of the sine terms only arises from discrete symmetry properties, and cosine terms might occur in more general cases. Then (2.21) translates into the following set of equations for the Fourier coefficients:

$$(n \cdot k)\, \chi_{1k} = -H_{1k} \tag{2.24}$$

for all multiindex k; for $k = (0, \ldots, 0)$ there is no constraint on χ_1, and χ_{10} is assigned to guarantee uniqueness: $\chi_{10} = 0$. For $k \neq 0$ the value of the Fourier coefficient χ_{1k} is uniquely determined by (2.24) provided $(n \cdot k) \neq 0$. This *non–resonance condition* will be satisfied in some subset of the phase space; let us for a moment forget about the problem of where our solution will be defined, and define the order one parts of χ, H' by solving (2.24) as if the non–resonance condition was always satisfied:

$$\chi_{1k} = \frac{-H_{1k}}{n \cdot k} \tag{2.25}$$

As for H_1', the $k = 0$ term cannot be removed and there is a unique solution:

$$H_1'(p') = H_{10}(p') \tag{2.26}$$

We can now look at the third of the (2.19), which can be somewhat simplified by using both (2.21) and (2.26):

$$H'_2 = H_2 - \tfrac{1}{2}\{H_1 + H_{10}, \chi_1\} - \{H_0, \chi_2\} \tag{2.27}$$

and is of the same general form; again a unique solution is found by expanding H_2 and χ_2 in Fourier series, imposing $\chi_{20} = 0$, setting H'_2 to be the $k = 0$ Fourier component of whatever is in the right hand side, and solving for the coefficients χ_{2k} by dividing the corresponding Fourier coefficient of $H_2 - 1/2\{H_1 + H_{10}, \chi_1\}$ by the *divisor* $n \cdot k$. The problem of the divisor being somewhere zero gets worse, since χ_2 contains the divisor squared in some denominators; but formally, the second step is not any more difficult than the first. It is easy to show that the same occurs with the higher order parts of χ and H', that is there is a formal series solution for χ such that $T_\chi H = H'(p')$. This algorithm was proposed by Hori (1966); Deprit (1969) proposed an algorithm which is not the same, but gives the same results.

This allows to compute a formal series solution of the original problem (2.18). In the (q', p') coordinate system the problem is solved, because $H' = H'(p')$:

$$\frac{dp'}{dt} = 0 \quad ; \quad \frac{dq'}{dt} = \frac{\partial H'}{\partial p'} = n'(p') \tag{2.28}$$

and the solution is simply given by:

$$p'_r = P_r \quad ; \quad q'_r = n'(P)(t - t_0) + Q_r \tag{2.29}$$

with integration constants P_r, Q_r which we shall call *proper elements* (for historical reasons to be discussed later; see Chapter 4). Then the explicit computation of a solution with given initial condition takes in principle three steps:

[1] Given the values of ε, and the initial conditions of (q, p) at $t = t_0$, use the map F_χ to compute (q', p') at $t = t_0$ and thus (Q, P).

[2] Given P and the expression of $H' = H'_0 + \varepsilon H'_1 + \ldots$ as computed along with χ, find $n'(P)$, thus the solution $(p'(t), q'(t))$ is available for every t.

[3] Use the inverse map $F_\chi^{-1} = F_{-\chi}$ to compute the solution in the (q, p) coordinate system:

$$\begin{aligned} q(q', p') &= T_\chi q = q' - \{q, \chi\} + \tfrac{1}{2}\{\{q, \chi\}, \chi\} \ldots \\ &= q' - \frac{\partial \chi}{\partial p} + \tfrac{1}{2}\{\frac{\partial \chi}{\partial p}, \chi\} + \ldots \\ p(q', p') &= T_\chi p = p' - \{p, \chi\} + \tfrac{1}{2}\{\{p, \chi\}, \chi\} \ldots \\ &= p' + \frac{\partial \chi}{\partial q} - \tfrac{1}{2}\{\frac{\partial \chi}{\partial q}, \chi\} + \ldots \end{aligned} \tag{2.30}$$

The right hand side of (2.30) has to be interpreted as a function originally defined in terms of the variables (q, p) which has however to be evaluated for $q = q', \ p = p'$.

Since the latter are given by (2.29), each q_r circulates with average frequency n'_r (the same as q'_r) but with superimposed oscillations containing all the fundamental frequencies n'_r ; the p_r are approximately constant and close to the corresponding p'_r, with superimposed conditionally periodic oscillations. This qualitative description of the solution (traditionally referred to as the *epicyclic solution*) holds only as long as the singularities of the coordinate system are not encountered, as we shall discuss in Section 4.1. Moreover, the series are handled as if they were convergent, which is a naïve assumption.

The series expansion for both χ and the solution belong to the general class often called *Poisson series*, which are actually double series, Fourier series in some variables (here the q'_r) and power series in others (here ε). The problem of the order in which these series should be sommed has been left open, as well as the convergence and the domain of definition problems. These Poisson series are well defined mathematical objects on which many operations can be performed; unfortunately, the most difficult operation to perform, and even to define, is the *evaluation* of the series, that is to find a real number corresponding to a given set of values for q', p', ε.

2.4 DEGENERACY AND RESONANCES

Even at the formal series solution stage the problems really encountered in Celestial Mechanics are more difficult than the model problem discussed in the previous section. The main difficulty arises from the *degeneracy* of the 2–body problem, i.e. from the simple fact that the perihelia and the nodes are integrals. As a result, the non resonance conditions:

$$\frac{\partial H_0}{\partial p} \cdot k \neq 0 \tag{2.31}$$

cannot be satisfied for every multiindex $k \neq 0$; in a sense, resonance occurs in every point of the phase space.

To understand what can be done under these circumstances we first describe the elimination procedure in a somewhat more abstract way: given any order zero Hamiltonian H_0, we can define a linear operator L acting on any function $g = g(q, p)$ by:

$$Lg = \{H_0, g\} \tag{2.32}$$

It defines a decomposition of the function space (of the Poisson series) into a direct sum of the kernel (null space) of the operator L and of the image of L:

$$g = \tilde{g} + \overline{g} \quad ; \quad \tilde{g} \in Im\, L \; ; \; \overline{g} \in Ker\, L \tag{2.33}$$

Then the existence of solutions of the recursive equations (2.19) can be discussed by decomposing e.g. $H_1 = \overline{H_1} + \tilde{H_1}$:

$$H'_1 = \overline{H_1} + \tilde{H_1} - L\chi_1 \tag{2.34}$$

has solution with $\chi_1 \in Im\, L$ provided

$$H'_1 = \overline{H_1} \tag{2.35}$$

which is the generalisation of (2.26) when the non–resonance condition is not assumed. The second order equation:

$$H'_2 = H_2 - \tfrac{1}{2}\{H_1 + \overline{H_1}, \chi_1\} - L\chi_2 \tag{2.36}$$

gives the definition of H'_2 and the equation for χ_2 by using the decomposition (2.33):

$$H'_2 = \overline{H_2} - \tfrac{1}{2}\overline{\{\tilde{H}_1, \chi_1\}} \tag{2.37}$$

$$L\chi_2 = \tilde{H}_2 - \{\overline{H_1}, \chi_1\} - \tfrac{1}{2}\{\tilde{H}_1, \chi_1\} + \tfrac{1}{2}\overline{\{\tilde{H}_1, \chi_1\}} \tag{2.38}$$

and the solution $\chi_2 \in Im\, L$ exists and is unique, and so on and so forth. At the end of the infinite recursion on the order, or rather when the process is arrested because the remainder containing ε^r is considered negligible, the Poisson series H' and χ are uniquely determined, with:

$$\chi \in Im\, L \quad ; \quad H' \in Ker\, L \tag{2.39}$$

In other words, the machinery works all the same, but the results are not the same, because H' is not necessarily a trivial Hamiltonian; $H' = H'(p')$ occurs if and only if the non–resonance conditions (2.31) are satisfied for every $k \neq 0$. However, H' is simpler than H in the sense that some of the Fourier components have been removed. To understand which ones we shall compute two useful examples.

Our first example is just the $N+1$–body problem as presented in the canonical coordinates which solve the zero order approximation. The latter is one form or another (depending upon the chosen coordinate system, see Laskar, this volume) of the 2–body problem copied N times. The angular variables $q = (\lambda, \theta)$ are the mean longitudes $\lambda_r,\ r = 1, \ldots, N$ and the longitudes of the perihelia and of the nodes $\theta_r,\ r = 1, \ldots, 2N$. The momenta are Delaunay–type variables $p = (\Lambda, \Theta)$ with the $\Lambda_r,\ r = 1, \ldots, N$ functions of the semimajor axes only and the $\Theta_r,\ r = 1, \ldots, 2N$ related to the angular momenta of the 2–body subsystems, hence to the eccentricities and the inclinations. Then the order zero approximation H_0 is a suitable linear combination of 2–body Hamiltonians:

$$H_0 = H_0(\Lambda) = \sum_r \left(-\frac{K_r}{2\Lambda_r} \right) \tag{2.40}$$

with K_r some coefficients depending only upon the masses (for the meaning of these coefficients, see Message, 1982; Milani and Nobili, 1983). All the derivatives are zero but for the *mean motions*:

$$\frac{\partial H_0}{\partial \Lambda_r} = \frac{K_r}{\Lambda_r^3} = n_r \tag{2.41}$$

The perturbing function is on the contrary dependent upon all the variables:

$$H = H_0(\Lambda) + \varepsilon H_1(\lambda, \theta, \Lambda, \Theta) + \dots \tag{2.42}$$

with the small parameter ε a function of the masses (and possibly of distance scaling parameters). Thus the elimination procedure can be carried out essentially in the same way described in Section 2.3. The effect of the elimination can be easily computed if the further assumption is made that the mean motions fulfill a non–resonance condition:

$$n \cdot k \neq 0 \quad ; \quad for\ every\ k \neq 0 \tag{2.43}$$

The condition (2.43), forbidding resonances in mean motion, is not the same as the general non–resonance condition (2.31) assumed in Section 2.3. As a result, $Ker\, L$ does not only contain the functions of the momenta; all the functions g independent from the longitudes are such that $Lg = 0,\ g = \overline{g}$:

$$Lg = \{H_0, g\} = -n \cdot \frac{\partial g}{\partial \lambda} \tag{2.44}$$

It can be shown, in a suitable function space (e.g. in the space of formal Poisson series), that (2.42) implies that the *long periodic functions* depending only upon $(\theta, \Lambda, \Theta)$ are the ones and the only ones in $Ker\, L$; this arises from the possibility of solving equations for the Fourier coefficients of the form (2.24). On the contrary a *short periodic function* depends upon λ and has all the Fourier coefficients of the arguments $k \cdot \theta$ equal to zero; that is, the decomposition (2.33) can be simply performed by splitting the Fourier series into the terms with and without the λ_r.

Therefore the end product of the elimination process is a new Hamiltonian in the new variables $(\lambda', \theta', \Lambda', \Theta')$:

$$H'(\theta', \Lambda', \Theta') = H_0(\Lambda') + \varepsilon \overline{H_1}(\theta', \Lambda', \Theta') + \varepsilon^2 \left[\overline{H_2} + \overline{\{\tilde{H}_1, \chi_1\}} \right] + \dots \tag{2.45}$$

which defines the *secular perturbations problem* (Message, 1976; 1982; Milani and Nobili, 1987). Is the Hamiltonian (2.45) any better than the original one? The number of variables has essentially been reduced, since the Λ'_r are integrals. Given the initial value of Λ', if a solution is known to the problem given by H' as a function of (θ', Θ') , then λ' can be computed by quadrature, and again the inverse map F_χ^{-1} gives the solution in the original coordinate system as in (2.30). However the secular perturbations problem is not integrable, because it has (in general) more than one angle variable; thus the problem is only displaced. The secular perturbation problem can in turn be attacked with more or less the same method, as we shall discuss in Section 4.2; anyway to represent the solution of the $N+1$–body problem as a Poisson series there is no way to avoid this double computation, one for the elimination of the mean longitudes, one for the solution of the secular problem.

The second example is the *single resonance problem* (see Message, 1988). Let us assume that the order zero Hamiltonian is as in (2.40), and that the non–resonance conditions are satisfied for all k but one, that is there is a multiindex $\tilde{k}$ such that:

$$n \cdot k \neq 0 \ for \ every \ k \neq r\tilde{k} \tag{2.46}$$

The only multiindexes which can possibly generate a null divisor are multiples of the single generator $\tilde{k}$; it follows that the MCD of al the integeres $\tilde{k}_r$ is 1. Then it is possible to find a unimodular transformation of the longitudes such that the first component is:

$$\sigma = \tilde{k} \cdot \lambda \tag{2.47}$$

The construction of a suitable unimodular matrix has been discussed in Section 1.3 for the case of 2 bodies only partecipating in the resonance; it can be shown that the condition for such a matrix A to exist is anyway $MCD(\tilde{k}) = 1$. Let

$$\begin{pmatrix} \sigma \\ \tau \end{pmatrix} = A\lambda \quad ; \quad \begin{pmatrix} \Sigma \\ T \end{pmatrix} = [A^{-1}]^T \Lambda \tag{2.48}$$

be the canonical transformation thus defined. A different elimination process, specially adapted to this case, is obtained by just splitting the set of angle variables in a different way: the *fast angles* are now the τ_r, $r = 1, \ldots, N-1$, while the *slow angles* are $(\sigma, \theta_1, \ldots, \theta_{2N})$. Then the function χ can be recursively computed. Each Fourier coefficient of χ can be solved for in an equation with some divisor $n \cdot k$ which is nonzero because $k \neq r\tilde{k}$, and the final product is an Hamiltonian in the transformed coordinates (primed):

$$H' = H'(\sigma', \theta', \Sigma', T', \Theta') = H_0(\Sigma', T') + \ldots \tag{2.49}$$

which is useful to transform the single resonance problem to a canonical form (see Henrard, this volume). Again the problem has been simplified only in that the momenta T'_r conjugate to the fast angles are integrals, and the fast angles τ_r themselves are cyclic variables to be later computed by quadratures. However to further discuss this problem the setting of this section, based entirely upon formal series, is not enough; order of magnitude consideration must play an essential rôle (see Henrard, this volume; Ferraz–Mello, this volume).

CHAPTER 3:

INVARIANT TORI

In this chapter we discuss a common feature of Hamiltonian systems, namely the existence of invariant manifolds topologically equivalent to an N–dimensional torus. We shall also discuss the related question of the non convergence of the series arising in perturbation theories for an Hamiltonian system perturbed from an integrable one.

3.1 INTEGRABLE SYSTEMS

An Hamiltonian system defined by the Hamiltonian $H = H(q,p)$ on a domain D is said to be integrable if there is a global canonical transformation

$$\begin{aligned} F &: D \longmapsto D' \\ (q,p) &\longmapsto (x,y) \end{aligned} \tag{3.1}$$

such that the transformed Hamiltonian depends only upon the momenta:

$$H(F^{-1}(x,y)) = K(y) \tag{3.2}$$

If the map F^{-1} can be explicitly computed, the integrable system can be solved by transforming into the (q,p) space the trivial solution in the (x,y) space:

$$\begin{cases} y(t) = y(0) \\ x(t) = \nu t + x(0) \end{cases} \qquad \text{where} \quad \nu = \frac{\partial K}{\partial y}(y(0)) \tag{3.3}$$

It must be stressed that the definition requires the canonical map to be global; locally, for a small enough D, every Hamiltonian system is integrable (this follows from the existence of a smooth integral flow which is also canonical, see Section 2.1). Thus the very meaning of the definition depends upon the topology of the definition domain D. As an example, quadratic Hamiltonians are integrable over all of $\mathcal{R}^{2N}$ (see Section 1.3). However, the other case relevant for Celestial Mechanics, namely the 2–body problem, has a different topology. A very interesting result was obtained independently by Arnold and by Jost; it describes sufficient conditions for an Hamiltonian system to be integrable, and it also prescribes the toplogy of D, and this applies (with a small modification) to the 2–body problem as well.

The conditions of Arnold and Jost require the existence of N integrals, where N is the number of degrees of freedom (i.e. the number of components of both q and

p). The integrals are functions $f_1(q,p),\ldots,f_N(q,p)$ defined on D which are constant along the solutions of the Hamiltonian system, parametrised by time t:

$$\frac{df_i}{dt} = \{f_i, H\} = 0 \tag{3.4}$$

Of course one of the f_i, let us say f_1, can coincide with H. The integrals f_i are assumed to be smooth and functionally independent (i.e. their gradients ∇f_i are linearly independent in each point od D); thus for each set of constants $c_i,\ i=1,\ldots,N$ the level set $M(c)$ defined by assigning a value to each integral $f_i(q,p)=c_i,\ i=1,\ldots,N$ is a smooth N–dimensional manifold (if not empty). One further assumption is that the integrals *commute*:

$$\{f_i, f_j\} = 0 \qquad i,j = 1,\ldots,N \tag{3.5}$$

Then the first part of the Arnold–Jost theorem constrains the topology of D:

[1] If a (non–empty) level manifold $M(c)$ is compact, that is it is limited and does not touch the boundary of D, then $M(c)$ is topologically equivalent to an N–torus. If all the level manifolds are compact, then D is toplogically equivalent to the product of an N–torus and some N–manifold W: $D \simeq T^N \times W$. If the level manifold $M(c)$ is not compact, but it is complete (that is, each orbit on $M(c)$ of the system with either H or any of the f_i as Hamiltonian is defined for every time t, $-\infty < t < +\infty$) then $M(c)$ is a generalised *cylinder* $M(c) \simeq T^{N-j} \times \mathcal{R}^j$.

This applies in a straightforward way to the 2–body problem: in polar coordinates in the orbital plane (r,θ,p_r,p_θ) the Hamiltonian is:

$$H = \tfrac{1}{2}(p_r^2 + \frac{p_\theta^2}{r^2}) - \frac{k^2}{r} \tag{3.6}$$

with integrals H and p_θ, and they commute since $\{p_\theta, H\} = -\partial H/\partial\theta = 0$ is automatically satisfied. The only problem can occur where the gradients ∇H and ∇p_θ are parallel; the conditions for this to occur are:

$$p_r = \frac{dr}{dt} = 0 \qquad ; \qquad \frac{k^2}{r^2} = \frac{p_\theta^2}{r^3} = r(\frac{d\theta}{dt})^2 \tag{3.7}$$

which imply a circular orbit. Thus the Arnold and Jost result applies to the entire phase space $\simeq S^1 \times \mathcal{R}^3$ provided the set of initial conditions belonging to circular orbits ($r = p_\theta^2/k^2,\ p_r = 0$) are excluded (this defines the so–called *Delaunay domain*). To find out which of the level manifolds (with fixed values of $H = E$ and $p_\theta = J$) are compact we can solve for p_r from (3.6):

$$p_r^2 = 2E + \frac{2k^2}{r} - \frac{J^2}{r^2} \tag{3.8}$$

and a simple study of the quadratic polynomial in $1/r$ in the right hand side shows that to have a real solution for p_r the value of r is bounded away from both $+\infty$ and 0 (that is from the boundaries of D, corresponding to escape to infinity and collision respectively) if and only if $E < 0$ and $J \neq 0$. Thus all the elliptic orbits with the same E, J, that is with the same a, e, form a 2–torus in phase spac. As for the circular orbits, by using (3.8) and (3.7) we find they are the level manifolds corresponding to couples of values of E, J such that $2EJ^2 = k^4$, consistently with the definition of the eccentricity:

$$e^2 = 1 + \frac{2EJ^2}{k^4}$$

Thus the values of E, J fulfilling $e = 0$ define a level manifold whith a lower dimensionality ($\simeq S^1$). For either $E \geq 0$ (that is, $e \geq 1$) or $J = 0$ the level manifold is not compact, but it is complete, and is a cylinder $S^1 \times \mathcal{R}$; for $1 + 2EJ^2/k^4 < 0$ (3.8) has no real solution p_r and the level manifold is empty.

The extension to three dimensions is not very difficult. The Hamiltonian H itself and the three components of the angular momentum vector are of course integrals, but they do not commute; the Hamiltonian, the component of the angular momentum along a fixed direction, and the length of the angular momentum vector do commute and the Arnold–Jost result applies. The only complication arises from the fact that for orbits in the reference plane the angular momentum scalar coincides with the selected component, thus their gradients are not independent. The level manifolds are mostly T^3, degenerating into T^2 for either circular inclined or elliptic zero inclination orbits, and into S^1 for circular zero inclination. For either parabolic, hyperbolic or collision orbits the level manifolds are mostly $T^2 \times \mathcal{R}$, degenerating into $S^1 \times \mathcal{R}$ for zero inclination.

The proof of the statement [1] is interesting but beyond our scope; it is based on an argument about canonical flows (see Arnold, 1976, Chap. 10; Arnold and Avez, 1968). Namely, each one of the f_i defines a flow on the level manifold M, and this defines a map between the space of the associated independent variables s_i and M which is a local toplogical equivalence. If the manifold M is complete, this map is surjective and this is enough to show that $M \simeq T^{N-j} \times \mathcal{R}^j$.

The second half of the Arnold–Jost theorem defines a canonical coordinate system which actually integrates the system in the sense of (3.2):

[2] If some level manifold $M(c)$ is compact, it has a neighbourhood of the form $D_1 \simeq T^N \times \mathcal{R}^N$ on which a canonical transformation F with the property (3.2) is defined. If all the level manifolds are compact, the transformation F can be defined on all of $D \simeq T^n \times W$.

The new coordinate system is called *angle–action variables*; the angle variables x parametrize each torus and the action variables y are integrals. The proof of this second part of the Arnold–Jost theorem is essentially a global version of Liouville theorem (see Ferraz–Mello, this volume), and can be expressed by means of line integrals of the kind used in Henrard, this volume, for the $N = 1$ case (see

Arnold, 1976; Arnold and Avez, 1968). It is important to stress that the angle–action variables for a given hamitlonian system are by no means unique; given a set of angle–action variables, any unimodular transformation A can be applied to the angle variables provided the actions are transformed by $(A^{-1})^T$. However, it can be shown that under the *non–degeneracy condition* $det(\partial^2 K/\partial y^2) \neq 0$ there are no other sets of angle–action variables than those generated from any one of them by unimodular transformations.

For the 2–body problem (3.6) the angle–action variables are the *Delaunay variables* ℓ, ω, L, G. To remove the indetermination produced by the possibility of applying unimodular transformations we can specify the following: ℓ is a variable making a complete revolution along the orbit, and not changing at all when the points in phase space are subjected to a rotation $\theta \mapsto \theta + const$; ω on the contrary makes a complete revolution when the points in phase space are rotatetd through 2π, and does not change along the orbit. Then ℓ is the mean anomaly, ω the argument of pericenter, $G = p_\theta$ and L is a function of the total energy H such that:

$$K(L, G) = -\frac{k^4}{2L^2} \tag{3.9}$$

3.2 KOLMOGOROV THEOREM

The question arises of whether a slightly perturbed integrable system preserves the property of being integrable, or at least some of the invariant tori. Let us look at the problem in the coordinate system which makes the unperturbed system trivial, namely let us suppose (q, p) are already angle–action variables for all the phase space $D \simeq T^N \times W$: that is, the Hamiltonian H has an "order zero" part H_0 which is integrable, and an "order 1" part which is not, but contains a *small parameter* ϵ as in (2.18):

$$H = H_0(p) + \varepsilon H_1(q, p) + \varepsilon^2 H_2(q, p) + \dots \tag{3.10}$$

The main idea –going back to Linstedt and Poincaré (see Poincaré, 1893)– is to perform a fixed frequency perturbation theory. The best way to understand it is to start from the (false) hypothesis that the system (3.10) is still integrable. Then there would be some angle–action system (x, y) such that $H(q, p) = K(y)$. Let us choose the values of the new actions $y_\circ$ and the corresponding frequencies:

$$\nu_\circ = \frac{\partial K}{\partial y}(y_\circ) \tag{3.11}$$

For fixed $y = y_\circ$ and x variable, since the latter are angle variables, a torus T^N is spanned in the (x, y) space; let us label this torus by the corresponding value of the frequency vector $\nu_\circ$ (this is a correct labelling if we assume that a *non–degeneracy condition* $det(\partial^2 K/\partial y^2) \neq 0$ is satisfied in the (x, y) space as well, thus the map

$y_\circ \mapsto \nu_\circ$ is a local topological equivalence). Then there would be a subset in the (q,p) space which corresponds to a torus in the following sense:

[1] there is an immersion (a map with jacobian of maximum rank) $\Phi : T(\nu_\circ) \mapsto D$ such that $\Phi(T(\nu_\circ)) = T'(\nu_\circ)$ is invariant by the flow of H.

[2] Φ is an *equivariant map* between the flow of K on $T(\nu_\circ)$ and the flow of H restricted to $T'(\nu_\circ)$, that is not only Φ maps orbit onto orbit, but with the same timetable: if $F_H(q,p)$ is the integral flow of H (see 2.1), then provided the initial conditions $(q_\circ, p_\circ)$ are on $T'(\nu_\circ)$:

$$F_H^t(q_\circ, p_\circ) = \Phi(\nu_\circ t + x_\circ, y_\circ) \tag{3.12}$$

for some initial phase vector $x_\circ$. Of course, if such a map Φ was defined not only for $y = y_\circ$ but also for all the y in a neigbourhood, then the system would be integrable thereon. However, we do not ask for a map defined on an open set in the (x,y) space, but only on a "thin" set of dimension N.

To set up an algorithm to compute such a map, we can proceed as in Section 2.3. This time however, we proceed backward; namely we want to add to the integrable Hamiltonian $K(y)$ all the terms depending upon x. That is, we look for a function $\psi(x,y) = \epsilon\psi_1 + \dots$ such that it defines a map $F_\psi : (x,y) \mapsto (q,p)$ and the Hamiltonian is tranformed as follows:

$$\begin{aligned} T_\psi K &= K(p) - \epsilon\{K,\psi_1\} + \epsilon^2[-\{K,\psi_1\} + \tfrac{1}{2}\{\{K,\psi_1\},\psi_1\}] + \dots = \\ &= H = H_\circ + \epsilon H_1 + \epsilon^2 H_2 + \dots \end{aligned} \tag{3.13}$$

Equation (3.13) is identical to (2.14), apart from going the other way round; the main difference arises in the next step, namely the expansion corresponding to formula (2.16) is not performed at this stage, because we do not know a priori the expansion of K in powers of ϵ. We can now set up a recursive system of equations which adds to $K(p)$ the terms depending upon q to reconstruct the original Hamiltonian H: to do this we identify the terms of the some order in ϵ in the two sides of (3.13), and find that there is an obvious mismatch: $H_\circ(p) \neq K(p)$; thus we push forward the discrepancy as an higher order term:

$$K - H_\circ = \epsilon K_1 \tag{3.14}$$

and the order one equation is:

$$K_1 + \{K,\psi_1\} = H_1 \tag{3.15}$$

Now the known function H_1 can be decomposed: $H_1 = \overline{H_1} + \tilde{H}_1$ (as in Section 2.4), with respect to the linear operator $L_\circ$:

$$L_\circ g = \{K,g\} = -\nu_\circ \frac{\partial g}{\partial q}(q,p) \tag{3.16}$$

and the equation (3.15) is decomposed in the component belonging to $Ker\, L_\circ$, which again records a mismatch:

$$K_1 - \overline{H_1} = \epsilon K_2 \tag{3.17}$$

plus an equation in $Im\, L_\circ$:

$$\nu_\circ \frac{\partial \psi_1}{\partial q} = \tilde{H_1} \tag{3.18}$$

The order two equation can be similarly decomposed :

$$K_2 + \tfrac{1}{2}\overline{\{\tilde{H_1}, \psi_1\}} - \overline{H_2} = \epsilon K_3 \tag{3.19}$$

$$\nu_\circ \frac{\partial \psi_2}{\partial q} = \tilde{H_2} + \tfrac{1}{2}\{\tilde{H_1}, \psi_1\} - \tfrac{1}{2}\overline{\{\tilde{H_1}, \psi_1\}} \tag{3.20}$$

and so on. Thus we recursively define $\psi = \tilde{\psi}$ and find the relationship between K and H:

$$K = H_\circ + \epsilon\overline{H_1} + \epsilon^2[\overline{H_2} + \tfrac{1}{2}\overline{\{\tilde{H_1}, \psi_1\}}] + \dots \tag{3.21}$$

which is formally the same found by means of the inverse procedure in Section 2.4. However, the recursive procedure outlined here *is not the same* used in Sections 2.3 and 2.4: $L_\circ$ *is not* L, and the divisors occurring in the computations (3.18), (3.20), etc. are:

$$k \cdot \nu_\circ = k \cdot \frac{\partial K}{\partial y}(y_\circ) = k \cdot \frac{\partial H_\circ}{\partial y} + \epsilon k \cdot \frac{\partial \overline{H_1}}{\partial y} + \dots \tag{3.22}$$

and are different form $k \cdot n = k \cdot \partial H_\circ / \partial p$, unless the point $p_\circ$ to compute the frequencies n is chosen in such a way that $n = \nu_\circ$; this is indeed possible because of the non–degeneracy condition, but (3.22) points out that when this occurs $p_\circ \neq y_\circ$, the difference being of order ϵ.

We can now state the theorem announced by Kolmogorov (1954), and whose proof was published by Arnold (1963): the essential hypothesis is that the divisors $k \cdot \nu_\circ$ occuring in the solutions of the equations such as (3.18), (3.20) are never zero, nor too small. The condition required for the Kolmogorov–Arnold proof is that the divisors cannot go to zero faster than a power of the degree $|k| = \sum |k_i|$ of the multiindex k:

$$|k \cdot \nu_\circ| > \frac{c}{k^r} \tag{3.23}$$

with c, r positive constants. Under this hypothesis the recursive procedure defining the map Φ for the given $\nu_\circ$ can be shown to converge (however, the proof by Arnold used a different formalism; our presentation is more related to Benettin

et al., 1984). However a further condition is necessary to ensure the existence of the torus $T'(\nu_\circ)$, and it is a global condition. If a vector $\nu_\circ$ of frequencies is chosen at random, there is of course no guarantee that a solution (3.12) will exist in the domain D with these frequencies. The non–degeneracy condition ensures the existence of these frequencies locally, that is, if a point $(q_\circ, p_\circ)$ in D is known to have frequencies close to $\nu_\circ$, then a neigbouring point will be on the solution (3.12) with frequencies exactly $\nu_\circ$. But globally we need to ensure that there is some point in D such that the frequencies are at least close. This is obtained by adding a further hypothesis, namely that there is a value $p_\circ$ of the action variables for the unperturbed problem $H_\circ$ such that:

$$n(p_\circ) = \frac{\partial H_\circ}{\partial p}(p_\circ) = \nu_\circ \tag{3.24}$$

Under these hypothesis, namely (3.23), (3.24) and non–degeneracy, for a regular H, there is an $\epsilon_1 > 0$ such that for $0 < \epsilon < \epsilon_1$ the map Φ exists, is regular and has the property (3.12). The exact regularity conditions can change when different demonstration techniques are used, but this of course does not matter for Celestial Mechanics applications where H is anyway real analytic.

It must be stressed that the purpose of the hypothesis (3.24) is to ensure the existence of the image of Φ within D, and it is by no means true that the torus $T'(\nu_\circ)$ must contain any point with $p = p_\circ$: as the computation of the higher order parts of K proceeds (by means of formulas such as (3.13), (3.16)) the value of $y_\circ$ needs to be adjusted to keep the frequencies fixed at $\nu_\circ$. Moreover, p has to be computed as a function of $(x, y_\circ)$ by means of a formula such as (2.30). There is no guarantee that the torus $T'(\nu_\circ)$ will intersect the original unperturbed torus $p = p_\circ$, and it turns out not to be the case in many examples. As discussed by Poincaré (1893), fixed frequencies and fixed actions are incompatible conditions and cannot be satisfied simultaneously.

After the theorem of Kolmogorov was announced, it appeared that the main difficulty in applying it to a realistic problem in Celestial Mechanics had to do with finding a realistic estimate of the limiting value ϵ_1 for the perturbation parameter. Later it was found that this is not the case, as we shall discuss below.

3.3 DEGENERACY AND ARNOLD THEOREM

The first obvious difficulty in applying Kolmogorov theorem to Celestial Mechanics is that one of the hypothesis, namely non–degeneracy, is not true for the N–body problem. Let us suppose the N–body problem is represented in some coordinate system derived from Delaunay elements for each planet (considered as a separate 2–body problem with the Sun); this can be done in a number of ways (see Laskar, this volume), and the result is an order zero Hamiltonian of the form (2.40), which essentially results from a linear combination of 2–body Hamiltonians (3.10) with coefficients depending upon the masses of the planets (and of the Sun). Then:

$$H(\lambda,\theta,\Lambda,\Theta) = H_\circ(\Lambda) + \epsilon H_1(\lambda,\theta,\Lambda,\Theta) + \dots \tag{3.25}$$

and the non–degeneracy condition does not hold because $H_\circ$ depends only upon some of the action variables, that is upon the semimajor axes only. In terms of the global existence condition for a Kolmogorov torus with frequencies $\nu_\circ$, since the unperturbed frequencies for the perihelia and nodes θ are zero, there cannot be an invariant torus $T'(\nu_\circ)$ unless some components of $\nu_\circ$ are close to zero. However, if this global condition is satisfied, the theorem can be used with a transformation of the problem introduced by Arnold (1963).

Let $\overline{H_1}$ be the part of H_1 not depending upon the mean longitudes λ, and $\tilde{H_1} = H_1 - \overline{H_1}$ as in Section 2.4. Then it is possible to reorder the Hamiltonian in this way:

$$H(\lambda,\theta,\Lambda,\Theta) = [H_\circ(\Lambda) + \epsilon\overline{H_1}(\theta,\Lambda,\Theta)] + \epsilon\tilde{H_1}(\lambda,\theta,\Lambda,\Theta) + \dots \tag{3.26}$$

and then apply the Kolmogorov theorem to an Hamiltonian whose zero order part is the one enclosed in square brackets in (3.26). The non–degeneracy condition, that is the non–degeneracy of the matrix:

$$\frac{\partial^2 H_\circ + \epsilon\overline{H_1}}{\partial(\Lambda,\Theta)^2} = \begin{pmatrix} \frac{\partial^2 H_\circ}{\partial\Lambda^2} + \epsilon\frac{\partial^2 H_\circ}{\partial\Lambda^2} & \epsilon\frac{\partial^2\overline{H_1}}{\partial\Lambda\partial\Theta} \\ \epsilon\frac{\partial^2\overline{H_1}}{\partial\Theta\partial\Lambda} & \epsilon\frac{\partial^2\overline{H_1}}{\partial\Theta^2} \end{pmatrix} \tag{3.27}$$

can be guaranteed for small ϵ if the matrix:

$$D = \frac{\partial^2\overline{H_1}}{\partial\Theta^2} \tag{3.28}$$

is non–degenerate. It turns out that the matrix D plays a very important rôle in the theory of secular perturbations, (see Section 4.3), and explicit computations to show that it is non–degenerate can be performed. Thus it is possible to choose a set of frequencies $\nu_\circ$ –with some fast frequencies $n_\circ$ (corresponding to some set of unperturbed $\Lambda_\circ$) and some slow frequencies corresponding to possible values of $\epsilon\partial\overline{H_1}/\partial\Theta$– fulfilling (3.23), and apply Kolmogorov theorem.

As a result of the appearance of ϵ both in the determinant of the matrix (3.27) and in the perturbation $\epsilon\tilde{H_1}$, to obtain realistic estimates for the limiting value ϵ_1 is even more difficult than it already is in the non–degenerate case. Moreover, since the matrices (3.27) and (3.28) can be explicitly computed only for zero eccentricities and inclinations, Arnold result applies only to orbits with very small eccenticities and inclinations. Nevertheless it was important to show that invariant tori exist in the full N–body problem, even though this cannot be rigorously shown to occur for realistic values of planetary masses, eccentricites and inclinations.

3.4 DIVERGENCE OF THE SERIES

The object of this section will be a "generic" perturbed Hamiltonian, by which we mean $H = H_0 + \epsilon H_1 + \ldots$, the sum of an integrable, non degenerate order zero part:

$$H_0 = H_0(p) \quad ; \quad det\, \frac{\partial^2 H_0}{\partial p^2} \neq 0 \tag{3.29}$$

and a perturbation which contains "all the terms", namely when it is expanded in a Fourier series in the angle variables q:

$$H_1 = \sum_k H_{1k}(p) \cos(k \cdot q) \quad ; \quad H_{1k} \sim \alpha^{|k|} \tag{3.30}$$

where the $\sim$ symbol indicates an asymptotic relationship and $\alpha < 1$. It is essential for what follows that $\sim$ indicates not only an upper bound for the Fourier coefficients –which is needed to ensure the convergence of the Fourier series to a real analytic function– but also a lower bound, that is for large enough $|k|$ the Fourier coefficients H_{1k} are not allowed to be zero.

The procedure described in Section 2.3 allows to define a formal Poisson series:

$$\chi(q,p) = \epsilon\chi_1 + \epsilon^2\chi_2 + \ldots \quad ; \quad \chi_i = \sum_k \chi_{ik} \sin(k \cdot q) \tag{3.31}$$

such that the problem is solved in the formal sense, namely a formal series transformation is defined by χ, such that:

$$F_\chi : (q,p) \mapsto (x,y) \quad ; \quad T_\chi H = H(F_\chi^{-1}(x,y)) = K(y) \tag{3.32}$$

On the other hand, the Kolmogorov theorem of Section 3.2 defines another Poisson series:

$$\psi(x,y) = \epsilon\psi_1 + \epsilon^2\psi_2 + \ldots \quad ; \quad \psi_i = \sum_k \psi_{ik} \sin(k \cdot x) \tag{3.33}$$

by which a map is defined in the other direction:

$$F_\psi : (x,y) \mapsto (q,p) \quad ; \quad T_\psi K = K(F_\psi^{-1}(q,p)) = H(q,p) \tag{3.34}$$

Now the question arises of whether any one of the two formal Poisson series (3.31), (3.33) can be convergent, and where. Kolmogorov theorem appears to give a very asymmetrical answer, namely the series (3.33) for ψ can be convergent on a

torus $y = y_\circ$ in the (x, y) space provided $\nu_\circ = \partial K/\partial y(y_\circ)$ is bounded away from resonances by an inequality like (3.23). On the contrary, the series (3.31) for χ is essentially always divergent. The fundamental asymmetry of the two maps arises from the fact that in one of the two spaces, the one with coordinates (x, y), the dynamical system is assumed to be integrable, actually integrated: $K = K(y)$. On the contrary, the algorithms to generate χ and ψ are the same: at each stage an equation of the form

$$L\chi_i = -n(p)\frac{\partial \chi_i}{\partial q} = g(q,p) \quad ; \quad L_\circ \Psi_i = -\nu(y)\frac{\partial \psi_i}{\partial x} = f(x,y) \tag{3.35}$$

is solved, with g, f functions determined by the previous steps of the recursive procedure. The frequency vectors are $\nu = \partial K/\partial y\,(y)$ and $n = \partial H_0/\partial q\,(p)$ respectively. By Fourier expansion, the coefficients are solutions of equations of the form:

$$\chi_{ik} = -\frac{g_k}{n(p)\cdot k} \quad ; \quad \psi_{ik} = -\frac{f_k}{\nu(y)\cdot k} \tag{3.36}$$

Then the divergence of the series can be described by perfectly symmetric statements:

[1] Let A be a pathwise connected set in the (q, p) space; if the series for χ converges on A, then $p = const$ on A.

[2] Let B be a pathwise connected set in the (x, y) space; if the series for ψ converges on B, then $y = const$ on B.

The proof of either [1] or [2] is very simple: let us assume a point with $p = p_\circ$ and one with $p = p_1$ belong to A; then there is a continous path inside A joining the two; along this path the frequency vector $n(p)$ changes continously from $n(p_\circ)$ to $n(p_1)$, thus there is some multiindex k such that $n(p)\cdot k = 0$ somewhere along the path (actually, there is an infinite number of such k, since rational numbers are dense). Where this occurs, there is a term in the series for χ which is singular and χ cannot be defined; this contradiction proves that $p_\circ = p_1$.

The asymmetry arises when the convergence of the series is used to solve the problem. Namely, χ can be defined on some $p = p_\circ$ torus and ψ can be defined on some $y = y_\circ$ torus; however, $y = y_\circ$ is an invariant subset, thus ψ can be convergent on a full orbit in the (x, y) space, and the map $F_\psi = \Phi$ provides a solution with the property (3.12). On the contrary, χ might well be convergent on some $p = p_\circ$, but this does not provide a solution because the image by F_χ of $p = p_\circ$ does not contain any solution in the (x, y) space.

This leaves us in an embarassing situation, because the algorithm to find a solution of the original problem in the (q, p) space as described in Section 2.3 can never work; namely, even if the series for χ were convergent on a set A including the initial conditions $(q(0), p(0))$, the image in the (x, y) space is such that the solution

of the integrable problem $K(y)$ cannot be computed on $F_\chi(A)$. On the other hand given a frequency $\nu_\circ$ the series for ψ could be convergent on an invariant set $y = y_\circ$, and this does provide a solution through F_ψ; however, we do not know where this solution is going to be at the initial time. Even if the initial condition in the (q, p) space actually belongs to an invariant torus of the kind described by the Kolmogorov theorem, we do not know any constructive and convergent algorithm to compute ν, thus we do not know how to compute ψ and the solution. In particular, F_ψ is not $F_\chi^{-1} = F_{-\chi}$, because they are defined on different sets in the (x, y) space (for a related dicussion, see Milani, 1988).

A further remark is needed on the generality of the negative results such as the one above. The hypothesis that all the Fourier terms in the perturbation H_1 have non zero coefficients for high enough degree $|k|$ has been used to show that whenever a divisor $n \cdot k$ is zero a (non removable) singularity arises in equation (3.36); even if $H_{1k} = 0$, there must be some multiple $\overline{k}$ of k with a non zero coefficient $H_{1\overline{k}}$, and it can be shown that the functions g appearing in the right hand side of (3.35) inherit the same property. The N–body problem however does not have such a property, namely there are *D'Alembert rules* which prescribe some constraints to the multiindexes k for the Hamiltonian to have $H_{1k} \neq 0$; this arises from the fact that the Hamiltonian of the N–body problem admits some integrals such as the total angular momentum (see section 4.1). However the proof of [1] can be completed by showing that some k with $n \cdot k = 0$ must occur even among the k fulfilling the D'Alembert rules (see Poincaré, 1892).

This problem of the divergence of the series has no "perfect" solution, because it arises from intrinsic properties of perturbed Hamiltonian systems; roughly speaking, any "perfect" solution would require integrability of the original problem. However, there can be ways out, and approximate solutions, as we shall see in the next chapter.

CHAPTER 4:

PROPER ELEMENTS

In this Chapter we discuss how to derive proper elements for planetary orbits. As it is already clear from the discussion in Chapter 3, the perfect proper elements, which should be integrals of the motion, do not exist precisely because the N–body problem is not integrable and the series of perturbation theory cannot be convergent on any open set. However, it is possible to perform some finite accuracy computations; they are used in two main applications to the dynamics of the Solar System, namely to define proper elements for the purpose of identifying asteroid families and to compute the secular perturbations of the orbits of the major planets. In the following Sections we shall not give details relative to either the one or the other of these two main cases, but rather discuss a general perturbed Hamiltonian derived from the coupling of a number N of 2–body problems as in (2.40)–(2.42):

$$H = H_0(\Lambda) + \varepsilon H_1(\lambda, \theta, \Lambda, \Theta) \tag{4.1}$$

where the unperturbed part of the Hamiltonian defines the mean motions $n_i = \partial H_0/\partial \Lambda_i$ and the angles λ_i are the mean longitudes, $\theta_i = \omega_i + \Omega_i,\ i = 1, \ldots, N$ and $\theta_j = \Omega_j,\ j = N+1, \ldots, 2N$. The perturbative part H_1 of the Hamiltonians we shall consider have one important property, arising from the symmetry of the problem with respect to rotations around any axis. This symmetry property is most easily described analytically if the set of action–angle variables for the unperturbed system H_0 is chosen in such a way that the total angular momentum along a fixed z axis is the sum of all the actions:

$$J_z = \sum_{i=1}^{N} \Lambda_i + \sum_{i=1}^{2N} \Theta_i \tag{4.2}$$

(this can be achieved by using jacobian coordinates and a suitable combination of rescaling and unimodular transformations similar to (1.32); see Laskar, this volume; Message, 1982). Then the two equivalent properties of invariance of H with respect to rotations around the z axis and of invariance of J_z with respect to the flow of H can be expressed by a zero Poisson bracket:

$$\{J_z, H\} = -\sum_i \frac{\partial H}{\partial \lambda_i} - \sum_i \frac{\partial H}{\partial \theta_i} = 0 \tag{4.3}$$

When the perturbing function is expanded in a Fourier series, the symmetry property (4.3) gives a condition to be satisfied by the coefficients of the series, that is a *D'Alembert rule:*

$$H_1 = \sum_{j,k} H_{1jk} \cos(j \cdot \lambda + k \cdot \theta) \tag{4.4}$$

$$H_{1jk} \neq 0 \; only \; if \sum_i j_i + \sum_i k_i = 0 \tag{4.5}$$

Another property of the set of action–angle variables $(\lambda, \theta, \Lambda, \Theta)$, is that the Hamiltonian is well defined even for $\Theta = 0$, which corresponds to zero eccentricity and inclination. This imposes further D'Alembert rules, namely for $\Theta \longrightarrow 0$ the coefficients $H_{1jk} \longrightarrow 0$, more precisely H_{1jk} at $\Theta = 0$ must have a zero of order at least $|k| = \sum_i |k_i|$ in the square roots of the Θ. In all the applications to Celestial Mechanics the Hamiltonian has the further property of being even in the angles θ, as a result of the invariance of both kinetic and potential energy with respect to a mirror symmetry; this allows to use only cosine terms in the expansions such as (4.4). These properties will have an essential rôle in the theory of secular perturbations.

4.1 LINEAR SECULAR PERTURBATION THEORY

The first approximation in the theory of secular perturbations is obtained by neglecting all the terms in the perturbative series but he ones of order 1 in the small parameter ε (which is roughly speaking the ratio of the massses of the planets to the mass of the Sun) and degree 1 in the actions Θ (that is, degree 2 in the eccentricities and inclinations of all the orbits). Then the two step procedure outlined in Section 2.4 can be explicitly performed without much difficulties.

The first step is the *elimination of the longitudes* (at order 1 in ε). Only the first order part of the generating function χ is solved for:

$$\chi = \varepsilon \chi_1 \quad ; \quad \chi_{1k} = -\frac{H_{1k}}{n \cdot k} \tag{4.6}$$

and the Lie transform is also computed to order 1 only:

$$T_\chi H = H' = H_0(\Lambda') + \varepsilon \overline{H_1}(\theta', \Lambda', \Theta') \tag{4.7}$$

In the transformed Hamiltonian the longitudes λ' do not appear, and the momenta Λ' (functions of the *proper semimajor axes*) are integrals. To apply the second simplification to the problem, namely to truncate H' to degree 2 in the eccentricities and inclinations, we need to use again the D'Alembert rules. Because of the way it is obtained from H, H' obeys the same D'Alembert rules: it is regular for $\Theta' \longrightarrow 0$ and such that it is invariant with respect to a rotation, that is for a change such that $\theta_i \mapsto \theta_i + \delta$, the same δ for all i. These properties are best expressed by means of a change to Poincaré variables (see (1.26)):

$$x_i = \sqrt{-2\Theta'_i}\cos(\theta'_i) \quad ; \quad y_i = \sqrt{-2\Theta'_i}\sin(\theta'_i) \tag{4.8}$$

The D'Alembert properties can be expressed in the Poincaré variables by stating that H' is a smooth function at $x = y = 0$, thus can be expanded in a series of positive powers of the x_i and y_i, and that H' is invariant with respect to a rotation of all the 2–vectors (x_i, y_i) by the same angle. It is known from elementary geometry that the functions defined on a set of vectors and invariant by rotation must contain only scalar and vector products (that is, distances and cosines and sines of the angles); thus H' must be a series formed only with positive powers of the combinations $x_i x_j + y_i y_j$ (scalar product) and $y_i x_j - x_i y_j$ (vector product). Since H' is even, only the cosine–type expressions $x_i x_j + y_i y_j$ need to be used. Thus when H' is expressed as a sum of homogeneous polynomials:

$$H' = H'_0 + H'_2 + H'_4 + \ldots \tag{4.9}$$

(with H'_r of degree r in the Poincaré variables), the degree 2 part must be of the form:

$$H'_2 = \tfrac{1}{2}\sum_{r,s} A_{rs}(x_r x_s + y_r y_s) \tag{4.10}$$

with the coefficients A_{rs} dependent only upon Λ' (that is, upon the proper semimajor axes). Within this approximation, the secular perturbation equations are linear equations with constant coefficients:

$$\frac{dx_r}{dt} = \sum_s A_{rs} y_s \quad ; \quad \frac{dy_r}{dt} = -\sum_s A_{rs} x_s \tag{4.11}$$

and can be solved in a straightforward way as already discussed in Section 1.3; the solutions contain the *fundamental secular frequencies* $\nu_i, \quad i = 1, \ldots, 2N$ which are the eigenvalues of the symmetric matrix A, and are obtained by means of the rotation matrix $B = (B^{-1})^T$ (which diagonalises A: $B^T A B = diag[\nu_1, \nu_2, \ldots, \nu_{2N}]$) from the proper modes of oscillation $w = B^{-1}x, z = B^{-1}y$. The latter are solutions of uncoupled harmonic oscillator problems, and for each oscillator (w_k, z_k) it is possible to define an action variable Θ^*_k and an angle variable θ^*_k:

$$w_k = \sqrt{-2\Theta^*_k}\cos(\theta^*_k) \quad ; \quad z_k = \sqrt{-2\Theta^*_k}\sin(\theta^*_k) \tag{4.12}$$

such that the solution is given by:

$$\Theta^* = const \quad ; \quad \theta^* = \nu t + \theta^*(0) \tag{4.13}$$

that is the solution in the original Poincaré variables is:

$$\begin{aligned} x_r &= \sum_s B_{rs} \sqrt{-2\Theta_s^*} \cos(\nu_s t + \theta_s^*(0)) \\ y_r &= \sum_s B_{rs} \sqrt{-2\Theta_s^*} \sin(\nu_s t + \theta_s^*(0)) \end{aligned} \tag{4.14}$$

This is the classical representation of the secular perturbations by means of epicycles, that is finite Fourier series with 2 terms per planet. Some caution must be used even in the interpretation of this simple solution. The proper action variables Θ_i^* can be used to define the proper eccentricities and inclinations (by the same formulas relating the ordinary eccentricities and inclinations to the Θ_i); however, these proper elements do not have the same geometrical meaning of the ordinary orbital elements (e.g. the perihelion distance cannot be computed directly from the proper eccentricity).

The geometrical meaning is even more involved for the proper angles θ_i^*: since the matrix A is in general not symmetric, the rotation B is not the identity, thus the angles θ_i^* and θ_i are not measured from the same origin. If the matrix A has a strongly dominating principal diagonal, then B is close to the identity and (provided the proper actions Θ_i^* are not too small) all the angles θ_i circulate with the same frequency as the corresponding proper θ_i^*, that is ν_i. This would be the case, in the planetary secular perturbation problem, if all the planets were very far apart from each other (that is, all the ratios between the semimajor axes of two consecutive planets were very small). However, the real Solar System is a weakly hierarchical system, and the off–diagonal terms in A are significant; as a result, some of the θ_i behave in a qualitatively different way from the corresponding θ_i^*: e.g. the perihelion of Jupiter and the perihelion of Uranus circulate around the Sun with the same mean frequency ν_5 (Milani and Nobli, 1985).

4.2 SECULAR PERTURBATIONS

To improve the computation of a long term solution of an Hamiltonian such as (4.1) we need to take into account both effects containing ε^2 and effects of degree 2 in the action variables Θ' (that is, of degree 4 in the eccentricities and inclinations); which of the two is most important will of course depend upon the problem.

The second order (that is, $O(\varepsilon^2)$) effects can arise in two different ways: in the elimination of the longitudes and in the solution of the secular perturbation equations. The transformation defined by the generating function $\chi = \varepsilon\chi_1 + \varepsilon^2\chi_2 + \ldots$ has many second order terms; the most important effects do not actually arise from the second order part of χ, but from the fact that a Lie transformation is not linear in the generating function. Let us take as an example the variables Λ, that is essentially the semimajor axes: after elimination of the longitudes, the Λ' are constant, and the Λ are given by formulas such as (2.30):

$$\begin{aligned}\Lambda &= \Lambda' - \varepsilon\{\Lambda', \chi_1\} + \tfrac{1}{2}\varepsilon^2\{\{\Lambda', \chi_1\}, \chi_1\} - \varepsilon^2\{\Lambda', \chi_2\} + \ldots = \\ &= \Lambda' + \varepsilon\frac{\partial \chi_1}{\partial \lambda'} - \tfrac{1}{2}\varepsilon^2\{\frac{\partial \chi_1}{\partial \lambda'}, \chi_1\} + \varepsilon^2\frac{\partial \chi_2}{\partial \lambda'} + \ldots\end{aligned} \tag{4.15}$$

To assess the relevance of the different parts of (4.15) we have to take into account that χ is short–periodic, that is $\overline{\chi} = 0$; if we compute the long–periodic part of Λ, all the parts of (4.15) linear in χ do not contribute, while the quadratic parts can give rise to beat terms, because χ_1 and $\partial\chi_1/\partial\lambda'$ contain terms with the same frequency whose Poisson bracket has a long periodic part:

$$\overline{\Lambda} = \Lambda' - \tfrac{1}{2}\varepsilon^2\overline{\{\frac{\partial \chi_1}{\partial \lambda'}, \chi_1\}} + \ldots \tag{4.16}$$

These *secular perturbations of the semimajor axes* contain only the secular frequencies ν_i and are not necessarily very small. Since the Fourier coefficients of χ_1 are derived from equations (4.5), they contain the divisors $\delta = n \cdot j + \nu \cdot k$; the second order part in (4.16) contains one factor of $1/\delta$ from χ_1, one from $\partial\chi_1/\partial\lambda'$ and a further factor $1/\delta$ arising from the derivative of $1/\delta$ with respect to Λ' occurring in the Poisson bracket; thus the order of magnitude is $\varepsilon^2\Theta^d\delta^{-3}$ if the divisor δ has a *D'Alembert characteristic* $d = \sum_i j_i$. If d is small, as in the case in which the ratio of two mean motions can be approximated by a fraction $(m + d)/m$ with small d, even a *shallow resonance* with a δ of the order of $\varepsilon^{1/3}$ can produce a large effect. For details on the calculations of these secular perturbations on the semimajor axes see Milani et al. (1987).

The short periodic terms from $\partial\chi_1/\partial\lambda'$ have order of magnitude $\varepsilon\Theta^{d/2}\delta^{-1}$, and if δ is of the order of $Sqrt\Theta^d\varepsilon$ the second order terms are actually larger, which means that the expansion in powers of ε is not performed in the right way: this *deep resonance* case must be solved in a very different way (see Henrard, this volume and Ferraz–Mello, this volume).

Of course the same effects occur in the variables Θ, that is in the eccentricities and inclinations; however, these variables are of the form $\Theta = \Theta' + O(\varepsilon)$ and the Θ', solution of the secular perturbation equations, are not constant at all, but undergo changes whose order of magnitude depends only upon the size of the proper action variables Θ^* (see (4.14)). This occurs because the derivatives of the Θ' are $O(\varepsilon)$, but the frequencies of the secular oscillations are $\nu = O(\varepsilon)$ and the periods $O(1/\varepsilon)$, thus the amplitude of the oscillations is of order zero in ε. In other words, the amplitude of the oscillations in the eccentricities and the inclinations of some orbit is controlled by the eccentricities and inclinations of the orbits of the other planets, not by the planetary masses. Thus the *free oscillations* described by equations such as (4.14) are more important that the second order effects due to the removal of the longitudes, and even than the first order short periodic effects unless the eccentricites and inclinations are all very small.

The other kind of second order effects can be more important for the eccentricities and inclinations: again since the Lie transform is not linear in χ, beat terms of the second order can appear in the transformed Hamiltonian H';

$$T_\chi H = H' = H_0 + \varepsilon\overline{H_1} - \tfrac{1}{2}\varepsilon^2\overline{\{H_1, \chi_1\}} + \dots \tag{4.17}$$

and the beat terms associated with a divisor δ of characteristic d will have order of magnitude $\varepsilon^2\Theta^d\delta^{-2}$, which is small outside deep resonance. However, the first order terms have order of magnitude $\varepsilon\Theta$, and for $d = 1$ (that is, close to a resonance like 2/1, 3/2, etc.) the ratio of the second order terms to the first order terms is $\varepsilon\delta^{-2}$: even for a very shallow resonance the contribution can be important. These second order corrections to the secular Hamiltonian were first computed by Yuasa (1973) for the asteroid problem and by Bretagnon (1974) for the major planets. There are cases in which the most significant improvement with respect to the classical linear perturbation theory of the first order (in ε) is a *second order linear theory*, that is the coefficents A_{rs} in (4.10) are computed not only from the first order part $\varepsilon\overline{H_1}$ but also including the contribution of the second order beat terms; the corresponding secular frequencies ν can change by a comparatively large amount. This technique is used in Milani and Knežević (1989).

The effects of the terms of degree four in the eccentricities and inclinations (that is, of degree 2 in the Θ') can be computed by a technique of elimination of the angles θ^* based again on the formalism of Lie series (see Yuasa, 1973; Message, 1976; 1982; Milani et al., 1987; for a different formalism, see also Bretagnon, 1974). For this purpose, let us assume that the secular perturbation Hamiltonian H' is expressed as a sum of homogeneous polynomials as in (4.9); each homogeneous component can be expressed as a function of the proper modes (w, z) (since the transformation between (x, y) and (w, z) is linear, the degree does not change), and in turn (w, z) can be expressed as a function of the proper action–angle variables (θ^*, Θ^*):

$$\begin{aligned} H' =& H_0 + H'_2(w, z) + H'_4(z, w) + \dots = \\ =& H'_0 + \nu \cdot \Theta^* + \sum_k c_k(\Theta^*)\cos(\theta^* \cdot k) \end{aligned} \tag{4.18}$$

Careful examination of (4.18) shows that H'_0 does not matter, since it contains neither θ^* nor Θ^*; H'_2 becomes trivial when expressed as a funtion of (θ^*, Θ^*) because these are the action–angle variables used to solve the problem with Hamiltonian H'_2; and the problem is to perform a new change in coordinate such that the angles θ^* are removed from H'_4. We have to stress again that both H'_2 and H'_4 may contain also the $O(\varepsilon^2)$ contribution.

Thus it is enough to find a $\chi' = \chi'_4 + \chi'_6 + \dots$ such that:

$$F'_\chi : (\theta^*, \Theta^*) \longmapsto (\theta^{**}, \Theta^{**}) \quad ; \quad T_{\chi'} H' = H''(\Theta^{**}) \tag{4.19}$$

To degree 4 it is enough to find the part χ'_4 homogeneous of degree 4 (in (w, z)): this because of the rule on the degree of the Poisson bracket (in Poincaré–type variables): $deg(\{f, g\}) = deg(f) + deg(g) - 2$.

$$T_{\chi'_4} H' = H'_2 + [H'_4 - \{H'_2, \chi'_4\}] + \dots \tag{4.20}$$

and the procedure follows the same lines of the one of Section 2.4: we define an operator L_2 and a decomposition of each function into the part with and without the angles θ^* (which we shall denote again with a bar and a tilde, although the meaning of these symbols is not the same as in Chap. 2):

$$L_2 g = \{h'_2, g\} = -\nu \cdot \frac{\partial g}{\partial \theta^*} \quad ; \quad H'_4 = \overline{H'_4} + \tilde{H'_4} \tag{4.21}$$

Then the part of degree 4 of the new hamiltonian H'' and the equations to be solved for χ'_4 are:

$$\begin{aligned} H''_4 &= \overline{H'_4} \quad ; \quad in\ Ker\, L_2 \\ L_2 \chi'_4 &= \tilde{H'_4} \quad ; \quad in\ Im\, L_2 \end{aligned} \tag{4.22}$$

and the new *nonlinear proper elements* $(\theta^{**}, \Theta^{**})$ obtained by the transformation $F_{\chi'}$ are known functions of time:

$$\Theta^{**} = const \quad ; \quad \theta^{**} = \nu^* t + \theta^{**}(0) \tag{4.23}$$

where the new fundamental frequencies ν^* are obtained from the new Hamiltonian H'', or at least from its degree 2 and degree 4 parts which have been explicitly computed:

$$\nu^* = \frac{\partial H''}{\partial \Theta^{**}}(\Theta^{**}) = \frac{\partial (H''_2 + H''_4)}{\partial \Theta^{**}}(\Theta^{**}) = \nu + \frac{\partial H''_4}{\partial \Theta^{**}}(\Theta^{**}) \tag{4.24}$$

At the end of this procedure, we can say that the solution is completely known up to all the effects arising from terms of degree 4 in the eccentricites and inclinations, since the ordinary orbital elements can be computed by reversing all the chain of transformations, that is by computing:

[1] $(\theta^{**}, \Theta^{**}) \longrightarrow (\theta^*, \Theta^*)$ by the Lie series defined by $-\chi'_4$;

[2] $(\theta^*, \Theta^*) \longrightarrow (\theta', \Theta')$ by means of the linear transformation B, that is esssentially by (4.14) and (4.8);

[3] $(\theta', \Theta') \longrightarrow (\theta, \Theta)$ by adding the short periodic perturbations with the Lie series defined by $-\chi$

Since all this sequence is long and tedious, but the algorithm is straightforward, one could think that the problem can be solved up to order 2 and degree 4 (or, for that matter, up to any order and degree such that the expansions and the Poisson brackets can be computed with a reasonable effort, possibly with the use of algebraic manipulation languages). Unfortunately, the real solution cannot be obtained in this conceptually simple way.

4.3 THE PROPER ELEMENTS MAP

To better understand the essential difficulty in the calculation of higher degree proper elements, let us assume that the elimination of the longitudes has been performed successfully (to order ε^2) and that the secular perturbation Hamiltonian has been expanded in powers of the linear proper elements: to be able to drop most of the stars, primes, etc. we change notation and assume the linear proper elements are (ϕ, I) and the secular Hamiltonian is K:

$$K = K(\phi, I) = \sum_k \nu_k I_k + \sum_j K_j^{(4)}(I) \cos(j \cdot \phi) \tag{4.25}$$

where we assume that $K_j^{(4)}$ is homogeneous of degree 2 in the actions I, thus the sum over the multiindex j is extended only to degree 4 (that is, $|j| = \sum |j_k| \leq 4$), and all the higher degree terms are neglected. It is easy (in principle) to extract the part of K which does not depend upon the angles ϕ:

$$K^* = K^*(I) = \nu \cdot I + K_0^{(4)}(I) \tag{4.26}$$

Then the algorithm described in the previous section prescribes to compute a generating function Γ satisfying the equation:

$$\{\nu \cdot I, \Gamma\} = \sum_{j \neq 0} K_j^{(4)}(I) \cos(j \cdot \phi) \tag{4.27}$$

and a simple manipulation of Fourier series gives the fully explicit solution:

$$\Gamma(\phi, I) = -\sum_{j \neq 0} \frac{K_j^{(4)}(I)}{j \cdot \nu} \sin(j \cdot \phi) \tag{4.28}$$

By taking the derivatives of Γ, the transformations to proper elements (ϕ^*, I^*) can be computed: let us indicate this transformation by:

$$\begin{cases} \phi^* = \phi + c(\phi, I, \nu) \\ I^* = I + C(\phi, I, \nu) \end{cases} \tag{4.29}$$

where we have stressed the dependence of the transformation upon the values of the first approximation frequencies, which appear in the generating function. Since combinations of up to four frequencies appear in the generating function as divisors, the correction (c, C) can be very sensitive to the values of the ν; if a resonance $j \cdot \nu = 0$ occurs, a transformation with the required properties cannot be defined. However, the purpose of the transformation (4.29) is to exploit the explicit solution of the problem in the starred variables, which contain the frequencies ν^*:

$$I^* = const \quad ; \quad \nu^* = \frac{\partial K^*(I^*)}{\partial I^*} \quad ; \quad \phi^* = \nu^* t + const \tag{4.30}$$

and of course $\nu^* \neq \nu$, because the degree four part of the Hamiltonian contributes to the frequencies. Actually the rate of change of ν^* as a function of the action variables is measured by the matrix computed in (3.28). Then what happens if there is a resonance $j \cdot \nu^* = 0$?

In this setting it is not difficult to show that the problem is not different from the one we have already discussed in Section 3.4. The transformation (4.29) can be defined, provided $j \cdot \nu \neq 0$ for every j with $0 < |j| \leq 4$, but what it gives is just the first approximation to a series which is divergent; the true solution of the equations of motion is obtained by the map going in the opposite direction, from proper elements (ϕ^*, I^*) such that $\nu^*(I^*)$ is non–resonant in some strong sense to an invariant torus in the (ϕ, I) space: as in Section 3.2, we can explicitly compute the first approximation to this map and find for the generating function Γ^* the solution:

$$\Gamma^*(\phi^*, I^*) = + \sum_{j \neq 0} \frac{K_j^{(4)}(I^*)}{j \cdot \nu^*} \sin(j \cdot \phi^*) \tag{4.31}$$

where two changes only occur with respect to (4.28): the opposite sign and the use of starred variables (arising from the fact that the map has to go in the opposite way, and in the Hori method this is obtained by changing the sign), and the use of the corrected frequencies ν^* in the divisors. Thus the analogous of (4.29) is:

$$\begin{cases} \phi = \phi^* - c(\phi^*, I^*, \nu^*) \\ I = I^* - C(\phi^*, I^*, \nu^*) \end{cases} \tag{4.32}$$

with the same functions (c, C) (they are functions with the same analytical expression, computed on different variables). The same arguments used in Section 3.4 show that (4.32) is the right solution (or at least a first approximation to it), while (4.29) is not. If a specific divisor $j \cdot \nu$ is small, and has opposite sign from $j \cdot \nu^*$, the corresponding term in Γ is large and has the 'wrong' sign, that is the opposite sign with respect to the corresponding term in Γ^*. Although this does not necessarily occur for a given (ϕ, I), it is bound to occur for some choices of the initial conditions (ϕ, I) and of the frequencies ν (that is, of the semimajor axes).

The potential dangers of the approximation involved in using the linear frequencies ν in the solutions of secular perturbation problems have been known since a long time; the unreliability of this approximation became unacceptable only when formulas such as (4.29) have been used to systematically map a large portion of the phase space into the proper elements phase space.

A solution to this difficulty has been proposed only recently (Milani and Knežević, 1989; Knežević and Milani, 1989); the idea is to use (4.32) instead of (4.29) to compute the proper elements (ϕ^*, I^*). Since (4.32) can be explicitly computed as a map $(\phi^*, I^*) \longrightarrow (\phi, I)$ which goes in the other direction, the proper elements have to be found by solving an implicit function problem. The implicit function problem can be written in a fixed point form:

$$\begin{cases} \phi^* = \phi + c(\phi^*, I^*, \nu^*) \\ I^* = I + C(\phi^*, I^*, \nu^*) \\ \nu^* = [\partial K^*/\partial I^*](I^*) \end{cases} \tag{4.33}$$

which can be solved by an iterative scheme as follows: as a first approximation, the difference between proper and non proper elements is neglected:

$$\phi_0^* = \phi \quad ; \quad I_0^* = I \quad ; \quad \nu_0^* = \nu + \frac{\partial K^*}{\partial I^*}(I_0^*) \tag{4.34}$$

then the successive approximations are computed by applying the map defined by the right hand side of (4.33):

$$\begin{cases} \phi_{N+1}^* = \phi + c(\phi_N^*, I_N^*, \nu_N^*) \\ I_{N+1}^* = I + C(\phi_N^*, I_N^*, \nu_N^*) \\ \nu_{N+1}^* = \nu + [\partial K^*/\partial I^*](I_N^*) \end{cases} \tag{4.35}$$

If the iteration scheme (4.35) converges to some limit (ϕ^*, I^*, ν^*), then this limit fulfills equation (4.33) and ν^* is the set of fundamental frequencies to be used in the solution of the secular perturbation equations. Of course this is only a first approximation; in principle, the corrected frequencies ν^* could be used as a first approximation for a further iteration which takes into account the terms of degree 6, and so on.

The study of this kind of "reverse KAM map" has just begun; it has not been proven that this infinite sequence of iterations can actually converge to a KAM torus; we do not know yet how to estimate the error introduced by stopping the procedure at a given degree; even for a fixed degree, we do not know where the iteration scheme (4.35) converges. However, experimental results obtained in the problem of asteroid proper elements seem to indicate that this is a good way to go.

References

Abraham, R. and Marsden, J.D.: 1967, *Foundations of Mecahnics*, Benjamin

Arnold, V.I. :1963, 'Small denominators and problems of stability of motion in classical and celestial mechanics', *Russian Mathematical Surveys* **18**, 85–191

Arnold, V.I. : 1976, *Méthodes Mathématiques de la Mécanique Classique*, MIR, Moscou

Arnold, V.I. and Avez, A.: 1968, *Ergodic problems of classical mechanics*, Benjamin

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

Bretagnon, P.: 1974, 'Termes à longues périodes dans le système solaire', *Astronomy and Astrophysics* **30**, 141-154

Deprit, A.: 1969, 'Canonical transformations depending on a small parameter', *Celestial Mechanics* **1**, 12–30

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

Knežević, Z. and Milani, A.: 1989, 'Asteroid proper elements from an analytical second order theory' in *Asteroids II*, Matthews, M. and Binzel, R. eds., Arizona University Press

Kolmogorov, A.N.: 1954, 'The conservation of conditionally periodic motions with a small variation in the Hamiltonian' *Dokl. Akad. Nauk. SSSR* **98**, 527–530

Message, P.J.: 1976, 'Formal expressions for the motion of N planets in the plane, with the secular variations included, and an extension to Poisson's theorem', in *Long-Time Predictions in Dynamics*, Szebehely and Tapley eds., Reidel Pu. Co., Dordrecht, Holland, 279–293

Message, P.J.: 1982, 'Asymptotic series for planetary motion in periodic terms in three dimensions', *Celestial Mechanics* **26**, 25–39

Message, P.J.: 1988, 'Planetary perturbation theory from Lie series, including resonance and critical arguments' in *Long–Term Dynamical Behaviour of Natural and Artificial N–body systems*, A.E. Roy ed., Kluwer Academic, 47–72

Milani, A.: 1988, 'Secular perturbations of planetary orbits and their representation as series', in *Long–Term Dynamical Behaviour of Natural and Artificial N–body systems*, A.E. Roy ed., Kluwer Academic, 73–108

Milani, A. and Knežević, Z.: 1989,'Asteroid proper elements', in preparation

Milani, A. and Nobili, A.M.: 1983, 'On the stability of hierarchical 4-body systems', *Celestial Mechanics* **31**, 241–291

Milani, A. and Nobili, A.M.: 1985, 'Resonant structure of the outer solar system', *Celestial Mechanics* **35**, 269–287

Milani, A., Nobili, A.M. and Carpino, M.: 1987, 'Secular variations of the semimajor axes: theory and experiments', *Astronomy and Astrophysics* **172**, 265–279

Poincaré H.:1892, *Methodes Nouvelles de la Mechanique Celeste, Vol. I,* Gauthier-Villars, Paris

Poincaré H.:1893, *Methodes Nouvelles de la Mechanique Celeste, Vol. II,* Gauthier-Villars, Paris

Yuasa, M.: 1973, 'Theory of Secular perturbations od Asteroids Including Trems of Higher Order and Higher Degrees', *Pubbl. Astron. Soc. Japan* **25**, 399–445