Chapter 2
Averaging and Normalization in Celestial Mechanics

Near-integrable Hamiltonian systems (i.e., the systems composed by an integrable part and a small perturbation) are a common model currently used to study the dynamics of celestial bodies. This is the study of such systems that Poincaré called the "general problem of dynamics" (Poincaré 1899). Usually the so-called *perturbation approach* is used to transfer the perturbation to higher orders of the Hamiltonian expansion, by means of a canonical transformation close to identity. Then, neglecting the remainder, one can use the truncated Hamiltonian for the analysis of the effects characterizing the system dynamics over long time intervals. Since in Celestial Mechanics these effects are called secular, the construction of a suitable transformation can be interpreted as a development of a secular theory of the dynamical phenomena under consideration.

In this book, mostly the restricted three-body problem (R3BP) is considered, in which a massless particle moves under gravitational attraction of two bodies (the primaries),[1] orbiting around their barycenter. There are several cases when the motion of the massless particle (the tertiary) can be treated as a slightly perturbed Keplerian motion: (1) the distance between the tertiary and one of the primaries is always much less than the distance between the primaries; (2) the distance between the tertiary and the primaries' barycenter is always much greater than the distance between the primaries; (3) the mass of the primary is much greater than the mass of the secondary.

To write down the Hamiltonian of R3BP in the form convenient for application of the perturbation technique, one needs to introduce the *Delaunay variables*. Then the Hamiltonian becomes a sum of an integrable part, corresponding to the Keplerian

[1] The three bodies are called the primary, the secondary, and the tertiary in accord (usually) with the hierarchy of their masses. However, other ways of enumeration are also possible, e.g., according to the hierarchy of the geometric configuration. This is usually clear from the context. The two "primaries" comprise the primary and the secondary.

I. I. Shevchenko, *The Lidov-Kozai Effect – Applications in Exoplanet Research and Dynamical Astronomy*, Astrophysics and Space Science Library 441,
DOI 10.1007/978-3-319-43522-0_2

motion, and a small perturbation. A consequent formal *averaging* of the perturbation over the *fast variables* allows one to construct a first-order theory describing the secular effect in the particle dynamics outside resonances with the primaries.

Basic concepts of the perturbation approach in Hamiltonian dynamics are discussed in many monographs and reviews; see, in particular, Poincaré (1899, 1905), Murray and Dermott (1999), Morbidelli (2002), Arnold et al. (2006), and Ferraz-Mello (2007). These sources are used for the material presentation in this section.

2.1 The Hamiltonian Formalism

In the subsequent analytical treatments, we usually use the equations of motion in the Hamiltonian formulation. Therefore, let us recall how basic problems of celestial mechanics can be expressed in the Hamiltonian "language".

A first-order system of $2n$ ordinary differential equations (ODEs)[2]

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{F}(\mathbf{x}) \tag{2.1}$$

(where $\mathbf{x}$ is a vector variable, $\mathbf{F}(\mathbf{x})$ an arbitrary vector function) is said to have a Hamiltonian form, if there exists a scalar function $\mathcal{H}(\mathbf{x})$, called the *Hamiltonian*, such that the system can be represented in the form

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial \mathcal{H}(\mathbf{x})}{\partial q_i}\,, \quad \frac{\mathrm{d}q_i}{\mathrm{d}t} = +\frac{\partial \mathcal{H}(\mathbf{x})}{\partial p_i}\,, \quad (i = 1, \ldots, n)\,, \tag{2.2}$$

or, equivalently, in the vector form,

$$\dot{\mathbf{p}} = -\nabla_{\mathbf{q}} \mathcal{H}(\mathbf{x})\,, \quad \dot{\mathbf{q}} = +\nabla_{\mathbf{p}} \mathcal{H}(\mathbf{x})\,, \tag{2.3}$$

where $\mathbf{x} = (\mathbf{q}, \mathbf{p}) = (q_1, \ldots, q_n, p_1, \ldots, p_n)$; t is an independent variable (time). The variables $q_1, \ldots, q_n$ and $p_1, \ldots, p_n$ are called the canonical *coordinates* and *momenta* of the system, respectively; n is called the *number of degrees of freedom* of the system. The *canonical variables* $\mathbf{q}$ and $\mathbf{p}$ are called to be *conjugate* to each other.

The advantage of the Hamiltonian formalism is that it allows one to perform all analytic operations with the scalar Hamiltonian function, instead of analyzing the whole set of equations of motion. Many equations in mechanics (including celestial mechanics) and physics can be set in Hamiltonian form. In particular, the equations of motion in a potential $U(\mathbf{r})$ (where $\mathbf{r} \equiv (x, y, z)$ is the position vector of a particle

[2]Note that in this book all vector quantities are set in bold font.

in a Cartesian frame), namely,

$$\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = -\nabla_{\mathbf{r}}\, U(\mathbf{r}), \tag{2.4}$$

where the gradient operator $\nabla_{\mathbf{r}} \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$, can be written down in form (2.2), setting

$$\mathbf{q} = \mathbf{r}, \quad \mathbf{p} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}, \quad \mathcal{H} = \frac{\|\mathbf{p}\|^2}{2} + U(\mathbf{q}). \tag{2.5}$$

2.2 The Two-Body Problem in a Hamiltonian Form

Consider first a Hamiltonian formulation of the two-body problem. Recall that, in this problem, two point masses m_0 and m_1 move under mutual gravitational attraction. The Newtonian equations of motion in an orthogonal frame are given by

$$\frac{\mathrm{d}^2\mathbf{u}_0}{\mathrm{d}t^2} = \frac{\mathcal{G}m_1}{\|\mathbf{u}_1 - \mathbf{u}_0\|^3}(\mathbf{u}_1 - \mathbf{u}_0)\,, \quad \frac{\mathrm{d}^2\mathbf{u}_1}{\mathrm{d}t^2} = \frac{\mathcal{G}m_0}{\|\mathbf{u}_0 - \mathbf{u}_1\|^3}(\mathbf{u}_0 - \mathbf{u}_1)\,, \tag{2.6}$$

where $\mathbf{u}_0$ and $\mathbf{u}_1$ are the bodies' positions in an inertial frame, $\mathcal{G}$ is the gravitational constant; $\|\mathbf{x}\| \equiv (x_1^2 + x_2^2 + x_3^2)^{1/2}$ is the length (norm) of vector $\mathbf{x}$ (see, e.g., Szebehely 1967).

If a frame is centered on the center of mass of the system, then the frame is called *barycentric*. Such a frame is inertial. If a frame is centered on the primary, it is called *heliocentric*, or *geocentric*, or *selenocentric*, etc., depending on the primary body. Such a frame is non-inertial. The relative position of bodies is $\mathbf{r} = \mathbf{u}_1 - \mathbf{u}_0$; the barycenter is located at $\mathbf{s} = (m_0\mathbf{u}_0 + m_1\mathbf{u}_1)/(m_0 + m_1)$.

Thus, $\mathbf{r}$ is the position vector of m_1 with respect to m_0, and $\mathbf{s}$ is the absolute position vector of the barycenter. From Equations (2.6) one has

$$\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = -\frac{\mathcal{G}(m_0 + m_1)}{\|\mathbf{r}\|^3}\mathbf{r}\,, \tag{2.7}$$

$$\frac{\mathrm{d}^2\mathbf{s}}{\mathrm{d}t^2} = 0\,. \tag{2.8}$$

Equation (2.7) describes the relative motion of bodies, and equation (2.8) means simply that the barycenter moves inertially.

Equations (2.7) are straightforwardly reducible to a Hamiltonian form with the Hamiltonian

$$\mathcal{H}_{\text{Kepler}} = \frac{\|\mathbf{p}\|^2}{2} - \frac{\mathcal{G}(m_0 + m_1)}{\|\mathbf{q}\|}, \tag{2.9}$$

where canonical coordinates $\mathbf{q}$ are the Cartesian primary-centric coordinates, and canonical momenta $\mathbf{p} \equiv \frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t}$.

The Hamiltonian equations of motion are

$$\dot{\mathbf{q}} = +\nabla_{\mathbf{p}}\mathcal{H}_{\text{Kepler}}, \qquad \dot{\mathbf{p}} = -\nabla_{\mathbf{q}}\mathcal{H}_{\text{Kepler}}, \tag{2.10}$$

where the gradient operators

$$\nabla_{\mathbf{p}} \equiv \left(\frac{\partial}{\partial p_x}, \frac{\partial}{\partial p_y}, \frac{\partial}{\partial p_z}\right), \qquad \nabla_{\mathbf{q}} \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right). \tag{2.11}$$

The Hamiltonian $\mathcal{H}_{\text{Kepler}}$ is nothing but the total energy of the system, equal to the sum of its kinetic and potential energies. It does not depend on time explicitly; such types of Hamiltonian systems are called *autonomous*. It is easy to show that, for such systems, $\mathcal{H}$ is conserved.

2.3 An *N*-Body Problem in a Hamiltonian Form

Consider the motion of a massless particle in the gravitation field of the Sun and N planets; thus, there are N perturbers. This is a restricted many-body problem (strictly speaking, a restricted $N + 2$-body problem, because the total number of bodies is equal to $N + 2$). If the perturbers are set to move in fixed orbits, the Newtonian equations of the particle's motion, written down in a primary-centric (heliocentric) Cartesian frame, are given by

$$\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = -\frac{\mathcal{G}m_0}{\|\mathbf{r}\|^3}\mathbf{r} + \sum_{i=1}^{N} \mathcal{G}m_i \left(\frac{\mathbf{r}_i - \mathbf{r}}{\|\mathbf{r}_i - \mathbf{r}\|^3} - \frac{\mathbf{r}_i}{\|\mathbf{r}_i\|^3}\right) \tag{2.12}$$

(see, e.g., Morbidelli 2002; Murray and Dermott 1999), where $\mathbf{r}$ and $\mathbf{r}_i$ ($i = 1, \ldots, N$) are the primary-centric positions of the particle and N gravitating perturbers (with masses m_i), respectively. Note that the first term in the parentheses (under the sum) corresponds to an obvious *direct* perturbation from planet m_i, and the second term to a less obvious *indirect* perturbation from the same planet. The *indirect* perturbation arises due to the gravitational effect of the planet on the Sun, implying a shift in the position of the system's center of mass.

The right-hand side of Equations (2.12) is representable as $-\nabla_{\mathbf{r}} U(\mathbf{r})$ with

$$U(\mathbf{r}) = -\frac{\mathcal{G}m_0}{\|\mathbf{r}\|} - \mathcal{G}\sum_{i=1}^{N} m_i \left(\frac{1}{\|\mathbf{r} - \mathbf{r}_i\|} - \frac{\mathbf{r}\cdot\mathbf{r}_i}{\|\mathbf{r}_i\|^3} \right) , \tag{2.13}$$

where $\mathbf{r}\cdot\mathbf{r}_i$ is the scalar product of $\mathbf{r}$ and $\mathbf{r}_i$. Consequently, in accord with Equation (2.4), the system's Hamiltonian is written as

$$\mathcal{H} = \frac{\|\mathbf{p}\|^2}{2} - \frac{\mathcal{G}m_0}{\|\mathbf{q}\|} - \mathcal{G}\sum_{i=1}^{N} m_i \left(\frac{1}{\|\mathbf{q} - \mathbf{r}_i\|} - \frac{\mathbf{q}\cdot\mathbf{r}_i}{\|\mathbf{r}_i\|^3} \right) , \tag{2.14}$$

where the canonical conjugate variables are coordinates $\mathbf{q} \equiv \mathbf{r}$ and momenta $\mathbf{p} \equiv \frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t}$.

According to formula (2.14), $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$, where $\mathcal{H}_0 = \frac{\|\mathbf{p}\|^2}{2} - \frac{\mathcal{G}m_0}{\|\mathbf{q}\|}$ coincides with Hamiltonian (2.9) of the two-body problem if one sets m_1 to zero, and term $\mathcal{H}_1$ can be regarded as a perturbation to the two-body problem. With respect to $\mathcal{H}_0$, the relative strength of the perturbation is of the order of ratio of the perturbers' mass to that of the primary.

Here the perturbers are assumed to move in fixed (unperturbed) orbits, i.e., the $\mathbf{r}_i(t)$ are given functions. Thus, $\mathcal{H} = \mathcal{H}(\mathbf{p}, \mathbf{q}, t)$. As it depends explicitly on time, the system is *non-autonomous*.

2.4 The Delaunay Variables

It is often desirable in celestial mechanics to work in canonical variables that are straightforwardly expressible through *osculating*[3] Keplerian orbital elements. (Definitions of the Keplerian orbital elements are given in Sect. 1.1. The Keplerian elements in the two-body problem do not form themselves a canonical set, as it is easy to check.)

What is more, it is often desirable to work with variables that are straightforwardly connected to basic dynamical quantities, such as energy and angular momentum; in order that one could expect the corresponding variables to be approximately conserved when the motion is "perturbed".

A set of such useful canonical variables, which is most popular now, was introduced by Delaunay. In the two-body problem, the *Delaunay variables* are

[3] Approximating the orbital motion of a body at a given instant in a best way (in some sense; see, e.g., Murray and Dermott 1999).

defined as

$$\begin{array}{lll} L = [\mathcal{G}(m_0+m_1)a]^{1/2}, & l = M, \\ G = L(1-e^2)^{1/2}, & g = \omega, \\ H = G\cos i, & h = \Omega, \end{array} \tag{2.15}$$

where L and l, G and g, H and h form three pairs of conjugate *"action-angle" variables*. In what concerns *actions* L, G, and H, the first of them, L, is a function of sole a (semimajor axis) and thus can be expressed through the total energy; G is the module of the reduced (per unit of mass) angular momentum; and H is the reduced angular momentum vector's vertical component. Indeed, the reduced angular momentum of the system is $\mathbf{r} \times \mathrm{d}\mathbf{r}/\mathrm{d}t$. Its norm G and its projection H on the vertical axis z are given by

$$G = [\mathcal{G}(m_0+m_1)a(1-e^2)]^{1/2}, \quad H = G\cos i\,. \tag{2.16}$$

Thus, the *actions* are all conserved in the two-body problem. The conjugate *angles* l, g, and h are just the mean anomaly M, argument of pericenter ω, and longitude of ascending node Ω, respectively. In the two-body problem, g and h are also constants of motion, and l circulates with a constant frequency (the mean motion n).

The angles l, g, and h are poorly defined at the inclinations or eccentricities close to zero; the *modified Delaunay variables*, given by

$$\begin{array}{lll} \Lambda = L = [\mathcal{G}(m_0+m_1)a]^{1/2}, & \lambda = l+g+h = M+\varpi, \\ P = L-G = L[1-(1-e^2)^{1/2}], & p = -g-h = -\varpi, \\ Q = G-H = 2G\sin^2\frac{i}{2}, & q = -h = -\Omega, \end{array} \tag{2.17}$$

evade this disadvantage (e.g., Morbidelli 2002).

When expressed in the Delaunay variables (2.15), the Hamiltonian of the two-body problem is given by

$$\mathcal{H}_{\mathrm{Kepler}} = -\frac{\mathcal{G}^2(m_0+m_1)^2}{2L^2} = -\frac{\mathcal{G}(m_0+m_1)}{2a}. \tag{2.18}$$

As immediately follows from the corresponding Hamiltonian equations of motion, the variables g, h, L, G, and H are conserved. In what concerns l, it varies according to the equation

$$\frac{\mathrm{d}l}{\mathrm{d}t} = +\frac{\partial \mathcal{H}_{\mathrm{Kepler}}}{\partial L} = \frac{\mathcal{G}^2(m_0+m_1)^2}{L^3}. \tag{2.19}$$

2.5 Near-Integrable Systems

If a Hamiltonian function depends merely on canonical momenta, i.e.,

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = \mathcal{H}(\mathbf{p}), \tag{2.20}$$

then the system is integrable and the trajectories are given by simple formulas, namely,

$$\mathbf{p} = \text{ const}, \quad \mathbf{q} = \boldsymbol{\omega}_0 t + \mathbf{q}_{(t=0)}, \tag{2.21}$$

where $\boldsymbol{\omega}_0 = \nabla_{\mathbf{p}} \mathcal{H}$.

If a Hamiltonian system is near-integrable, i.e., close (in the sense of a small parameter) to the integrable one, given by Equation (2.20), this means (by definition) that it can be represented in the form

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = \mathcal{H}_0(\mathbf{p}) + \epsilon \mathcal{H}_1(\mathbf{p}, \mathbf{q}) \,, \tag{2.22}$$

where ϵ is the small unitless parameter, $\epsilon \ll 1$.

Assume also that $\nabla_{\mathbf{p}} \mathcal{H}_0 \sim \nabla_{\mathbf{p}} \mathcal{H}_1$ by the order of magnitude. Under such conditions, one can treat $\mathcal{H}_0$ as a so-called *integrable approximation*, and $\mathcal{H}_1$ as a *perturbation* (Arnold et al. 2002, 2006; Morbidelli 2002). Therefore, system (2.22) is called *near-integrable*. Of course, any system that is reducible by a canonical transformation to form (2.22) is also near-integrable.

The solution of a system with Hamiltonian $\mathcal{H}_0$ approximates the solution of the perturbed system *with accuracy* $\sim \epsilon^{\alpha}$, where $\alpha \leq 1$ is some constant whose value depends on the system properties and initial conditions (see Arnold et al. 2002, 2006). This means that the divergence of the approximate solution from the exact one is of the order ϵ^{α} on the unit interval of time. Conversely, on the time interval equal to $\epsilon^{-\alpha}$, the divergence is of order of unity. Therefore, on a long-enough timescale the integrable approximation, given by $\mathcal{H}_0$, fails. Then, to describe the motion correctly, one must take into account the perturbation. *Perturbation theories* and corresponding *perturbative approaches* allow one to do this analytically, or to diminish the amount of the corresponding numerical work substantially.

The Hamiltonians of the restricted and planetary problems both are reducible to form (2.22), where ϵ is of the order of the perturbers' mass. In case of the Solar system, $\epsilon \sim 10^{-3}$, because the ratio of masses of Jupiter and Sun is $\approx 1/1047$. Similar small perturbations are characteristic for many satellite and exoplanetary systems. That is why the perturbative approach is often so useful and fruitful, allowing one to provide an analytical description to many complicated dynamical phenomena.

One of the basic tasks of the perturbative approaches consists in elimination of dependences of a Hamiltonian on *fast* variables, by means of a canonical transformation. As soon as they are eliminated, they become *cyclic* (by definition), and the

conjugate transformed momenta become constant, to the order of transformation. In fact, if all momenta of the system are constant, the solution of the transformed system is given by formulas (2.21), and the solution of the original system can be found by substituting the inverse transformation of variables into these formulas. In celestial mechanics, the role of the fast variables is usually played by the mean longitudes or anomalies. The procedure of elimination of a fast variable is called *averaging* of the system in this variable.

The notion of averaging is related to a (superior) notion of *normalization*. The normalization procedure consists in reducing the Hamiltonian to some "simple" form; e.g., to the *Birkhoff normal form*, in which all dependences on the canonical angles are eliminated in all orders of the Hamiltonian expansion in powers of a small parameter.

To keep the system Hamiltonian, the *normalizing transformation* must be canonical. Usually it is required to find: the normal form of the Hamiltonian up to a specified order; the generating function of the normalizing transformation; the formulae of direct and inverse transformations of the canonical variables. In all applications to the three-body problem, the procedure implies cumbersome analytical calculations. That is why the normalization is often accomplished by means of computer algebra.

2.6 The von Zeipel Method

Among the methods used for averaging and normalization in celestial mechanics, the von Zeipel method (von Zeipel 1916)[4] and the Hori–Deprit method (Deprit 1969; Hori 1966) are most popular. Though less convenient algorithmically, the first of them is an older one and it is analytically straightforward, that is why we describe it first. What is more, it was just the von Zeipel method that was mostly used by researchers for the averaging purposes in developments of the Lidov-Kozai theory.

In fact, the "von Zeipel method" was introduced by Poincaré (1899), who called it the "Lindstedt method". Von Zeipel used it extensively in studies in celestial mechanics, that is why the method is usually called now after his name. The method is described in a number of monographs, in particular, in Giacaglia (1972), Hagihara (1972), Kholshevnikov (1985), Zhuravlev and Klimov (1988), and Marchal (1990). Here we give a brief synopsis, based mostly on reviews in Giacaglia (1972), Zhuravlev and Klimov (1988), and Morbidelli (2002). The designations are generally such as adopted in Morbidelli (2002).

As a first step, any perturbative treatment of a near-integrable Hamiltonian system aims to find a *close-to-identity* canonical transformation

$$\mathbf{p} = \mathbf{p}^1 + \epsilon \mathbf{f}_1(\mathbf{p}^1, \mathbf{q}^1) , \quad \mathbf{q} = \mathbf{q}^1 + \epsilon \mathbf{g}_1(\mathbf{p}^1, \mathbf{q}^1) , \tag{2.23}$$

[4] The method was originally introduced by Poincaré (1899).

that allows one to transform Hamiltonian (2.22) to

$$\mathcal{K}(\mathbf{p}^1, \mathbf{q}^1) = \mathcal{H}_0(\mathbf{p}^1) + \epsilon\mathcal{K}_1(\mathbf{p}^1) + \epsilon^2\mathcal{K}_2(\mathbf{p}^1, \mathbf{q}^1) \,, \tag{2.24}$$

where $\mathcal{K}_1$ is nothing but the term $\mathcal{H}_1$ averaged over the angles $\mathbf{q}$, and $\mathcal{K}_2$ is a new function of the order not greater than $\mathcal{H}_1$. The Hamiltonian $\mathcal{H}_0 + \epsilon\mathcal{K}_1$ provides an *integrable approximation* (of order ϵ^2) of the original dynamical system.

To achieve a higher level of approximation of the motion, one can iterate, looking at each consecutive step for the canonical transformation

$$\mathbf{p}^{r-1} = \mathbf{p}^r + \epsilon^r \mathbf{f}_r(\mathbf{p}^r, \mathbf{q}^r) \,, \quad \mathbf{q}^{r-1} = \mathbf{q}^r + \epsilon^r \mathbf{g}_r(\mathbf{p}^r, \mathbf{q}^r) \,, \tag{2.25}$$

such that the Hamiltonian obtains the *Birkhoff normal form*

$$\mathcal{K}(\mathbf{p}^r) = \mathcal{H}_0(\mathbf{p}^r) + \epsilon\mathcal{K}_1(\mathbf{p}^r) + \cdots + \epsilon^r\mathcal{K}_r(\mathbf{p}^r) + O(\epsilon^{r+1}) \,. \tag{2.26}$$

Thus, at any consecutive step r of the procedure (at any consecutive order r in ϵ), the system's approximation is integrable, and the solution is given by

$$\mathbf{p}^r = \text{const}, \quad \mathbf{q}^r = \boldsymbol{\omega}^r t + \mathbf{q}^r_{(t=0)} \,, \tag{2.27}$$

where

$$\boldsymbol{\omega}^r = \nabla_{\mathbf{p}^r}[\mathcal{H}_0(\mathbf{p}^r) + \epsilon\mathcal{K}_1(\mathbf{p}^r) + \cdots + \epsilon^r\mathcal{K}_r(\mathbf{p}^r)] \,. \tag{2.28}$$

Constructing the inverse composition of all iterated transformations of the variables, one obtains the solution in the original variables $\mathbf{p}$ and $\mathbf{q}$.

Let us see how these perturbation techniques are realized in the von Zeipel algorithm. Consider Hamiltonian system (2.3) with Hamiltonian (2.22), i.e., a near-integrable original system. Our aim is to transform its Hamiltonian to normal form (2.26). To distinguish the transformed canonical variables and Hamiltonian from the original ones, we designate them by $\mathbf{P}$, $\mathbf{Q}$, and $\mathcal{K}$:

$$\mathcal{K}(\mathbf{P}) = \mathcal{K}_0(\mathbf{P}) + \epsilon\mathcal{K}_1(\mathbf{P}) + \cdots + \epsilon^r\mathcal{K}_r(\mathbf{P}) + O(\epsilon^{r+1}) \,. \tag{2.29}$$

To the second order in ϵ, the sought Hamiltonian is

$$\mathcal{K}(\mathbf{P}) = \mathcal{K}_0(\mathbf{P}) + \epsilon\mathcal{K}_1(\mathbf{P}) + \epsilon^2\mathcal{K}_2(\mathbf{P}) \,. \tag{2.30}$$

In the von Zeipel algorithm, the normalizing canonical transformation is determined by a *generating function* $\mathcal{S}(\mathbf{q}, \mathbf{P}, \epsilon)$:

$$\mathbf{p} = \nabla_{\mathbf{q}}\mathcal{S}(\mathbf{q}, \mathbf{P}, \epsilon) \,, \quad \mathbf{Q} = \nabla_{\mathbf{P}}\mathcal{S}(\mathbf{q}, \mathbf{P}, \epsilon) \,, \tag{2.31}$$

which is sought in the form of a power series in ϵ:

$$\mathcal{S}(\mathbf{q},\mathbf{P},\epsilon) = \mathbf{q}\cdot\mathbf{P} + \epsilon\mathcal{S}_1(\mathbf{q},\mathbf{P}) + \epsilon^2\mathcal{S}_2(\mathbf{q},\mathbf{P}) + \dots \ . \tag{2.32}$$

As needed, the generated transformation is close to identity; at $\epsilon = 0$ it is strictly identical: $\mathbf{p} = \mathbf{P}$, $\mathbf{q} = \mathbf{Q}$.

Evidently, the generating function must satisfy the equation

$$\mathcal{H}(\mathbf{q}, \nabla_{\mathbf{q}}\mathcal{S}, \epsilon) = \mathcal{K}(\mathbf{P}) = \text{const} \ . \tag{2.33}$$

Expanding $\mathcal{H}$, $\mathcal{S}$, and $\mathcal{K}$ in power series in ϵ, one has at each consequent order (0, 1, 2, ...) of ϵ:

$$\mathcal{H}_0(\mathbf{P}) = \mathcal{K}_0(\mathbf{P}) \ , \tag{2.34}$$

$$\nabla_{\mathbf{p}}\mathcal{H}_0\cdot\nabla_{\mathbf{q}}\mathcal{S}_1 + \mathcal{H}_1(\mathbf{q}, \nabla_{\mathbf{q}}\mathcal{S}_0) = \mathcal{K}_1(\mathbf{P}) \ , \tag{2.35}$$

$$\nabla_{\mathbf{p}}\mathcal{H}_0\cdot\nabla_{\mathbf{q}}\mathcal{S}_2 + \nabla_{\mathbf{p}}\mathcal{H}_1\cdot\nabla_{\mathbf{q}}\mathcal{S}_1 + \frac{1}{2}\sum_{i,\,j}\frac{\partial^2\mathcal{H}_0}{\partial p_i\partial p_j}\frac{\partial\mathcal{S}_1}{\partial q_i}\frac{\partial\mathcal{S}_1}{\partial q_j} + \mathcal{H}_2(\mathbf{q}, \nabla_{\mathbf{q}}\mathcal{S}_0) = \mathcal{K}_2(\mathbf{P}) \ , \tag{2.36}$$

$$\dots \ ,$$

where $1 \le i \le n$, $1 \le j \le n$ (and n is the number of degrees of freedom). Substituting $\nabla_{\mathbf{q}}\mathcal{S}_0 = \mathbf{P}$ and $\nabla_{\mathbf{p}}\mathcal{H}_0 = \mathcal{K}(\mathbf{P})$, we arrive at a sequence of linear equations in partial derivatives. Starting with order one in ϵ, this sequence is given by

$$\mathcal{K}(\mathbf{P})\cdot\nabla_{\mathbf{q}}\mathcal{S}_1 + \mathcal{H}_1(\mathbf{q},\mathbf{P}) = \mathcal{K}_1(\mathbf{P}) \ , \tag{2.37}$$

$$\mathcal{K}(\mathbf{P})\cdot\nabla_{\mathbf{q}}\mathcal{S}_2 + \mathcal{F}_2(\mathbf{q},\mathbf{P},\nabla_{\mathbf{q}}\mathcal{S}_1) = \mathcal{K}_2(\mathbf{P}) \ , \tag{2.38}$$

$$\mathcal{K}(\mathbf{P})\cdot\nabla_{\mathbf{q}}\mathcal{S}_3 + \mathcal{F}_3(\mathbf{q},\mathbf{P},\nabla_{\mathbf{q}}\mathcal{S}_1,\nabla_{\mathbf{q}}\mathcal{S}_2) = \mathcal{K}_3(\mathbf{P}) \ , \tag{2.39}$$

$$\dots \ .$$

Taking into account that the original Hamiltonian is periodic in angles $\mathbf{q}$, it is straightforward to choose $\mathcal{K}_1(\mathbf{P})$ equal to $\mathcal{H}_1(\mathbf{q},\mathbf{P})$ averaged over $\mathbf{q}$: $\mathcal{K}_1(\mathbf{P}) = \bar{\mathcal{H}}_1(\mathbf{P})$. Then, one has an equation for $\mathcal{S}_1$:

$$\mathcal{K}(\mathbf{P})\cdot\nabla_{\mathbf{q}}\mathcal{S}_1(\mathbf{q},\mathbf{P}) = \bar{\mathcal{H}}_1(\mathbf{P}) - \mathcal{H}_1(\mathbf{q},\mathbf{P}) \ . \tag{2.40}$$

This equation can be solved with respect to $\mathcal{S}_1$ by expanding $\mathcal{H}_1(\mathbf{q},\mathbf{P})$ in the Fourier series in $\mathbf{q}$ (recall that $\mathcal{H}_1(\mathbf{q},\mathbf{P})$ is periodic in $\mathbf{q}$), and seeking for $\mathcal{S}_1$ also in the form of a Fourier series. (Such a procedure is described in the next section for an analogous situation.)

Substituting the determined $\mathcal{S}_1$ and $\mathcal{K}_2(\mathbf{P}) = \bar{\mathcal{F}}_2(\mathbf{P})$ in Equation (2.38), one arrives at an equation for $\mathcal{S}_2(\mathbf{q}, \mathbf{P})$:

$$\boldsymbol{\omega}(\mathbf{P}) \cdot \nabla_{\mathbf{q}} \mathcal{S}_2(\mathbf{q}, \mathbf{P}) = \bar{\mathcal{F}}_2(\mathbf{P}) - \mathcal{F}_2(\mathbf{q}, \mathbf{P}, \nabla_{\mathbf{q}} \mathcal{S}_1) \,. \tag{2.41}$$

This equation is analogous to (2.40) and is solved in the same manner.

At each higher consecutive order in ϵ, the step algorithm is completely analogous. This is how the von Zeipel algorithm can be summarized.

Note that since the original system is generally non-integrable, the obtained regular solution is strictly *formal*, i.e., the series representing the solution are not obliged to converge. To obtain a solution closest to the true one, the iterations should be terminated at some optimum order r, depending on the system itself and on the value of ϵ.

2.7 The Hori–Deprit Method

At present, the most popular perturbative approach is one based on methods proposed by Hori (1966) and Deprit (1969). They employ the Lie perturbation techniques. This approach provides a number of advantages; in particular, the normalizing transformation is canonical by construction.

In the von Zeipel method, the generating function depends on "old" coordinate and "new" momentum variables. The absence of mixture of similar kind represents one of the principal advantages of the Hori–Deprit method over the von Zeipel method. What is more, the crucial advantage of the Hori–Deprit method consists in its practical recursiveness: it is based on recurrent explicit formulas, which reduce the normalization in every successive order to a standard mathematical procedure.

Let $\mathbf{q}$ and $\mathbf{p}$ be conjugate canonical coordinates and momenta, and $f(\mathbf{p}, \mathbf{q})$ and $g(\mathbf{p}, \mathbf{q})$ are some functions of them; then the Poisson bracket of f and g is defined as

$$\{f, g\} = \nabla_{\mathbf{q}} f \cdot \nabla_{\mathbf{p}} g - \nabla_{\mathbf{p}} f \cdot \nabla_{\mathbf{q}} g = \sum_{i=1}^{n} \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \,, \tag{2.42}$$

where n is the number of degrees of freedom.

For an arbitrary function $f = f(\mathbf{p}, \mathbf{q})$, where $\mathbf{p}$ and $\mathbf{q}$ are the solutions of a Hamiltonian system with a Hamiltonian $\mathcal{S}$, one has

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \nabla_{\mathbf{q}} f \cdot \dot{\mathbf{q}} + \nabla_{\mathbf{p}} f \cdot \dot{\mathbf{p}} = \{f, \mathcal{S}\} \,. \tag{2.43}$$

Let us expand f in a power series of t in the neighbourhood of $t = 0$:

$$f(t) = f(0) + \sum_{k=1}^{\infty} \frac{t^k}{k!} \cdot \frac{\mathrm{d}^k f}{\mathrm{d}t^k}\bigg|_{t=0} , \tag{2.44}$$

where $f(0) \equiv f(\mathbf{p}(0), \mathbf{q}(0))$ and $\frac{\mathrm{d}^k f}{\mathrm{d}t^k}\Big|_{t=0} \equiv \frac{\mathrm{d}^k f}{\mathrm{d}t^k}(\mathbf{p}(0), \mathbf{q}(0))$. Noting that

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \{f, \mathcal{S}\} , \quad \frac{\mathrm{d}^2 f}{\mathrm{d}t^2} = \left\{ \frac{\mathrm{d}f}{\mathrm{d}t}, \mathcal{S} \right\} = \{\{f, \mathcal{S}\}, \mathcal{S}\} , \quad \ldots , \tag{2.45}$$

and defining the *Lie operators* D^k with generator $\mathcal{S}$ by the recurrent relation

$$D^k f = D(D^{k-1} f) , \quad \text{where } Df = D^1 f = \{f, \mathcal{S}\} , \tag{2.46}$$

one can express $f(t)$ as

$$f(t) = f(0) + \sum_{k=1}^{\infty} \frac{t^k}{k!} \cdot D^k f\big|_{t=0} , \tag{2.47}$$

Series (2.47) is called the *Lie series* of a function f along the flow $\mathcal{S}$. Expansion (2.47) can be interpreted as an operator applied to f. As such, it is henceforth designated $L^t_{\mathcal{S}} f$.

Recall that a transformation of variables is called canonical, if it preserves the Hamiltonian form of equations. It is straightforward to verify (see Morbidelli 2002) that a transformation $(\mathbf{p}, \mathbf{q}) \to (\mathbf{P}, \mathbf{Q})$ of the form

$$\mathbf{p} = \mathbf{P} + \int_0^{\epsilon} \dot{\mathbf{P}} \mathrm{d}t = L^{\epsilon}_{\mathcal{S}} \mathbf{P} , \quad \mathbf{q} = \mathbf{Q} + \int_0^{\epsilon} \dot{\mathbf{Q}} \mathrm{d}t = L^{\epsilon}_{\mathcal{S}} \mathbf{Q} , \tag{2.48}$$

is canonical, if there exists a function $\mathcal{S}(\mathbf{P}, \mathbf{Q}, \epsilon)$ such that $\dot{\mathbf{P}}$ and $\dot{\mathbf{Q}}$ satisfy Hamiltonian equations

$$\dot{\mathbf{P}} = -\partial \mathcal{S} / \partial \mathbf{Q} , \quad \dot{\mathbf{Q}} = \partial \mathcal{S} / \partial \mathbf{P} . \tag{2.49}$$

Function $\mathcal{S}(\mathbf{P}, \mathbf{Q}, \epsilon)$ (where ϵ is a parameter) is called the *generating Hamiltonian* or the *generating function*.

Equations (2.48) represent an outcome of a canonical transformation (defined by the Hamiltonian flow with Hamiltonian $\mathcal{S}$), taken at a "time moment" ϵ. The *Lie perturbation techniques* are based on the transform representation (2.48).

Consider the normalizing transformation in the first order of ϵ. Thus, we define $\mathbf{p}^1 \equiv \mathbf{P}$, $\mathbf{q}^1 \equiv \mathbf{Q}$. As in case of any Lie series of a function, one has for the Hamiltonian:

$$\mathcal{K} = L^{\epsilon}_{\mathcal{S}} \mathcal{H} . \tag{2.50}$$

We seek for a generating function $\mathcal{S}$ such that the Hamiltonian (2.22) is transformed to the Birkhoff normal form in the first order of ϵ, i.e., the dependence on angles is eliminated in the first order.

According to (2.47), to the second order in ϵ, one has

$$\mathcal{K} = \mathcal{H}_0 + \epsilon\mathcal{H}_1 + \epsilon\{\mathcal{H}_0, \mathcal{S}\} + \epsilon^2\{\mathcal{H}_1, \mathcal{S}\} + \frac{\epsilon^2}{2}\{\{\mathcal{H}_0, \mathcal{S}\}, \mathcal{S}\} + O(\epsilon^3)\ . \tag{2.51}$$

It is implied that $\mathcal{H}_0$, $\mathcal{K}$, and $\mathcal{S}$ are all functions of $\mathbf{p}^1$, $\mathbf{q}^1$. Picking out the terms of the first order in ϵ, we equate them to a function of momenta only:

$$\mathcal{H}_1 + \{\mathcal{H}_0, \mathcal{S}\} = \bar{\mathcal{H}}_1\ . \tag{2.52}$$

The generating function $\mathcal{S}(\mathbf{p}^1, \mathbf{q}^1)$ can be found from this equation, the $\bar{\mathcal{H}}_1$ function being subject to the mentioned restriction (the dependence merely on $\mathbf{p}^1$). This equation, called the *homologic equation*, is nothing but a linear equation in partial derivatives, and it is easily solvable. It is sufficient to find any particular solution of this equation.

As described in Morbidelli (2002), the solution can be obtained in the following way. Since $\mathcal{H}$ is periodic in angles $\mathbf{q}^1$, one may expand $\mathcal{H}_1$ in the Fourier series:

$$\mathcal{H}_1(\mathbf{p}^1, \mathbf{q}^1) = \sum_{\mathbf{k}\in\mathbf{Z}^n} c_{\mathbf{k}}(\mathbf{p}^1) \exp\left(\iota\mathbf{k}\cdot\mathbf{q}^1\right)\ , \tag{2.53}$$

where $\iota = \sqrt{-1}$.

The solution of Equation (2.52) is sought also as the Fourier series

$$\mathcal{S}(\mathbf{p}^1, \mathbf{q}^1) = \sum_{\mathbf{k}\in\mathbf{Z}^n} d_{\mathbf{k}}(\mathbf{p}^1) \exp\left(\iota\mathbf{k}\cdot\mathbf{q}^1\right)\ ; \tag{2.54}$$

immediately one has

$$\{\mathcal{H}_0, \mathcal{S}\} = -\iota \sum_{\mathbf{k}\in\mathbf{Z}^n} d_{\mathbf{k}}(\mathbf{p}^1)\mathbf{k}\cdot\boldsymbol{\omega}_0(\mathbf{p}^1) \exp\left(\iota\mathbf{k}\cdot\mathbf{q}^1\right)\ , \tag{2.55}$$

where $\boldsymbol{\omega}_0 = \nabla_{\mathbf{p}^1}\mathcal{H}_0$; and the coefficients of the generating function (2.54) are determined as

$$d_{\mathbf{0}} = 0\ , \quad d_{\mathbf{k}}(\mathbf{p}^1) = -\iota\frac{c_{\mathbf{k}}(\mathbf{p}^1)}{\mathbf{k}\cdot\boldsymbol{\omega}_0(\mathbf{p}^1)} \tag{2.56}$$

at all non-zero $\mathbf{k}$. Besides,

$$\bar{\mathcal{H}}_1(\mathbf{p}^1) = c_0(\mathbf{p}^1)\ . \tag{2.57}$$

In the higher orders in ϵ, the procedure can be accomplished recursively; at each order, an analogous homologic equation is defined and solved (see Giacaglia 1972; Zhuravlev and Klimov 1988). The described procedure is called non-resonant, as it is implied that, in the course of normalization, one does not encounter zero resonant combinations in the denominators in Equations (2.56). For algorithms tackling the resonant situations, see the aforementioned monographs.

By means of the described algorithm the components of the generating function and the normalized Hamiltonian can be successively found, up to the required order of normalization. As a result we obtain the normalized Hamiltonian and the generating function of the normalizing transformation.

The canonical variables which the obtained generating function depends upon are not mixed. This circumstance allows one to calculate the normalizing transformation of the canonical variables as the Lie transformation with the generator equal to the newly found generating function $\mathcal{S}$. This transformation is given by the formulas

$$\mathbf{p} = \mathbf{P} + \sum_{k=1}^{\infty} \frac{\epsilon^k}{k!} D^k \mathbf{P}\Big|_{\epsilon=0} \ , \quad \mathbf{q} = \mathbf{Q} + \sum_{k=1}^{\infty} \frac{\epsilon^k}{k!} D^k \mathbf{Q}\Big|_{\epsilon=0} \ . \tag{2.58}$$

Variables $\mathbf{Q}$ and $\mathbf{P}$ represent the new canonical coordinates and momenta. The inverse transformation is the Lie transformation with the generator $(-S)$:

$$\mathbf{P} = \mathbf{p} + \sum_{k=1}^{\infty} \frac{(-\epsilon)^k}{k!} D^k \mathbf{p}\Big|_{\epsilon=0} \ , \quad \mathbf{Q} = \mathbf{q} + \sum_{k=1}^{\infty} \frac{(-\epsilon)^k}{k!} D^k \mathbf{q}\Big|_{\epsilon=0} \ . \tag{2.59}$$

When calculating the transformations of the canonical variables in practice, it is sufficient to leave the terms up to the order $M-1$ inclusive in the right-hand parts of the formulas, where M is the final order of normalization. Such length of expansions is sufficient to transform the Hamiltonian to the normal form of the given order, or, in the case of inverse transformation, to the initial form.