6 Perturbation theory

Perturbation theory is an efficient tool for investigating the dynamics of nearly–integrable Hamiltonian systems. The restricted three–body problem is the prototype of a nearly–integrable mechanical system (Section 6.1); the integrable part is given by the two–body approximation, while the perturbation is due to the gravitational influence of the other primary. A typical example is represented by the motion of an asteroid under the gravitational attraction of the Sun and Jupiter. The mass of the asteroid is so small, that one can assume that the primaries move on Keplerian orbits. The dynamics of the asteroid is essentially driven by the Sun and it is perturbed by Jupiter, where the Jupiter–Sun mass–ratio is observed to be about 10^{-3}. The solution of the restricted three–body problem can be investigated through perturbation theories, which were developed in the 18th and 19th centuries; they are used nowadays in many contexts of Celestial Mechanics, from ephemeris computations to astrodynamics.
Perturbation theory in Celestial Mechanics is based on the implementation of a canonical transformation, which allows us to find the solution of a nearly–integrable system within a better degree of approximation [66]. We review classical perturbation theory (Section 6.2), as well as in the presence of a resonance relation (Section 6.3) and in the context of degenerate systems (Section 6.4). We discuss also the Birkhoff normal form (Section 6.5) around equilibrium positions and around closed trajectories; we conclude with some results concerning the averaging theorem (Section 6.6).

6.1 Nearly–integrable Hamiltonian systems

Let us consider an n–dimensional Hamiltonian system described in terms of a set of conjugated action–angle variables $(\underline{I}, \underline{\varphi})$ with $\underline{I} \in V$, V being an open set of $\mathbf{R}^n$, and $\underline{\varphi} \in \mathbf{T}^n$. A nearly–integrable Hamiltonian function $\mathcal{H}(\underline{I}, \underline{\varphi})$ can be written in the form

$$\mathcal{H}(\underline{I}, \underline{\varphi}) = h(\underline{I}) + \varepsilon f(\underline{I}, \underline{\varphi}) \ , \tag{6.1}$$

where h and f are analytic functions called, respectively, the unperturbed (or integrable) Hamiltonian and the perturbing function, while ε is a small parameter measuring the strength of the perturbation. Indeed, for $\varepsilon = 0$ the Hamiltonian function reduces to

$$\mathcal{H}(\underline{I}, \underline{\varphi}) = h(\underline{I}) \ .$$

The associated Hamilton's equations are simply

$$\begin{aligned}\dot{\underline{I}} &= \underline{0} \\ \dot{\underline{\varphi}} &= \underline{\omega}(\underline{I}) \ ,\end{aligned} \tag{6.2}$$

where we have introduced the *frequency vector*

$$\underline{\omega}(\underline{I}) \equiv \frac{\partial h(\underline{I})}{\partial \underline{I}} \ .$$

Equations (6.2) can be trivially integrated as

$$\begin{aligned}\underline{I}(t) &= \underline{I}(0) \\ \underline{\varphi}(t) &= \underline{\omega}(\underline{I}(0))t + \underline{\varphi}(0) \ ,\end{aligned}$$

thus showing that the actions are constants, while the angle variables vary linearly with the time. For $\varepsilon \neq 0$ the equations of motion

$$\begin{aligned}\dot{\underline{I}} &= -\varepsilon \frac{\partial f}{\partial \underline{\varphi}}(\underline{I}, \underline{\varphi}) \\ \dot{\underline{\varphi}} &= \underline{\omega}(\underline{I}) + \varepsilon \frac{\partial f}{\partial \underline{I}}(\underline{I}, \underline{\varphi})\end{aligned}$$

might no longer be integrable and chaotic motions could appear.

6.2 Classical perturbation theory

The aim of *classical perturbation theory* is to construct a canonical transformation, which allows us to push the perturbation to higher orders in the perturbing parameter. With reference to the Hamiltonian (6.1), we introduce a canonical change of variables $\mathcal{C} : (\underline{I}, \underline{\varphi}) \to (\underline{I}', \underline{\varphi}')$, such that (6.1) in the transformed variables takes the form

$$\mathcal{H}'(\underline{I}', \underline{\varphi}') = \mathcal{H} \circ \mathcal{C}(\underline{I}, \underline{\varphi}) \equiv h'(\underline{I}') + \varepsilon^2 f'(\underline{I}', \underline{\varphi}') \ , \tag{6.3}$$

where h' and f' denote, respectively, the new unperturbed Hamiltonian and the new perturbing function. The result is obtained through the following steps: define a suitable canonical transformation close to the identity, perform a Taylor series expansion in the perturbing parameter, require that the change of variables removes the dependence on the angles up to second–order terms; finally an expansion in Fourier series allows us to construct the explicit form of the canonical transformation. Let us describe in detail this procedure.

Define a change of variables through a close–to–identity generating function of the form $\underline{I}' \cdot \underline{\varphi} + \varepsilon \Phi(\underline{I}', \underline{\varphi})$ providing

$$\begin{aligned}\underline{I} &= \underline{I}' + \varepsilon \frac{\partial \Phi(\underline{I}', \underline{\varphi})}{\partial \underline{\varphi}} \\ \underline{\varphi}' &= \underline{\varphi} + \varepsilon \frac{\partial \Phi(\underline{I}', \underline{\varphi})}{\partial \underline{I}'} \ ,\end{aligned} \tag{6.4}$$

where $\Phi = \Phi(\underline{I}', \underline{\varphi})$ is an unknown function, which is determined in order that (6.1) be transformed to (6.3). Let us split the perturbing function as

$$f(\underline{I}, \underline{\varphi}) = \overline{f}(\underline{I}) + \tilde{f}(\underline{I}, \underline{\varphi}) \ ,$$

where $\overline{f}(\underline{I})$ is the average over the angle variables and $\tilde{f}(\underline{I}, \underline{\varphi})$ is the remainder function defined as $\tilde{f}(\underline{I}, \underline{\varphi}) \equiv f(\underline{I}, \underline{\varphi}) - \overline{f}(\underline{I})$. Inserting (6.4) into (6.1) and expanding in Taylor series around $\varepsilon = 0$ up to the second order, one gets

$$\begin{aligned} & h\left(\underline{I}' + \varepsilon \frac{\partial \Phi(\underline{I}', \underline{\varphi})}{\partial \underline{\varphi}}\right) + \varepsilon f\left(\underline{I}' + \varepsilon \frac{\partial \Phi(\underline{I}', \underline{\varphi})}{\partial \underline{\varphi}}, \varphi\right) \\ & = \ h(\underline{I}') + \underline{\omega}(\underline{I}') \cdot \varepsilon \frac{\partial \Phi(\underline{I}', \underline{\varphi})}{\partial \underline{\varphi}} + \varepsilon \overline{f}(\underline{I}') + \varepsilon \tilde{f}(\underline{I}', \underline{\varphi}) + O(\varepsilon^2) \ . \end{aligned}$$

The transformed Hamiltonian is integrable up to the second order in ε provided that the function Φ satisfies:

$$\underline{\omega}(\underline{I}') \cdot \frac{\partial \Phi(\underline{I}', \underline{\varphi})}{\partial \underline{\varphi}} + \tilde{f}(\underline{I}', \underline{\varphi}) = \underline{0} \ . \tag{6.5}$$

The new unperturbed Hamiltonian becomes

$$h'(\underline{I}') \ = \ h(\underline{I}') + \varepsilon \overline{f}(\underline{I}') \ ,$$

which provides a better integrable approximation with respect to that associated to (6.1). An explicit expression of the generating function is obtained solving (6.5). To this end, let us expand Φ and $\tilde{f}$ in Fourier series as

$$\begin{aligned} \Phi(\underline{I}', \underline{\varphi}) \ &= \sum_{\underline{m} \in \mathbf{Z}^n \setminus \{\underline{0}\}} \hat{\Phi}_{\underline{m}}(\underline{I}') \ e^{i \underline{m} \cdot \underline{\varphi}} \ , \\ \tilde{f}(\underline{I}', \underline{\varphi}) \ &= \sum_{\underline{m} \in \mathcal{I}} \hat{f}_{\underline{m}}(\underline{I}') \ e^{i \underline{m} \cdot \underline{\varphi}} \ , \end{aligned} \tag{6.6}$$

where $\mathcal{I}$ denotes a suitable set of integer vectors defining the Fourier indexes of $\tilde{f}$. Inserting (6.6) in (6.5) one obtains

$$i \sum_{\underline{m} \in \mathbf{Z}^n \setminus \{\underline{0}\}} \underline{\omega}(\underline{I}') \cdot \underline{m} \ \hat{\Phi}_{\underline{m}}(\underline{I}') \ e^{i \underline{m} \cdot \underline{\varphi}} = - \sum_{\underline{m} \in \mathcal{I}} \hat{f}_{\underline{m}}(\underline{I}') \ e^{i \underline{m} \cdot \underline{\varphi}} \ ,$$

which provides

$$\hat{\Phi}_{\underline{m}}(\underline{I}') = - \frac{\hat{f}_{\underline{m}}(\underline{I}')}{i \ \underline{\omega}(\underline{I}') \cdot \underline{m}} \ . \tag{6.7}$$

Using (6.6) and (6.7), the generating function is given by

$$\Phi(\underline{I}', \underline{\varphi}) = i \sum_{\underline{m} \in \mathcal{I}} \frac{\hat{f}_{\underline{m}}(\underline{I}')}{\underline{\omega}(\underline{I}') \cdot \underline{m}} \ e^{i \underline{m} \cdot \underline{\varphi}} \ . \tag{6.8}$$

The algorithm described above is constructive in the sense that it provides an explicit expression for the generating function and for the transformed Hamiltonian. We stress that (6.8) is well defined unless there exists an integer vector $\underline{m} \in \mathcal{I}$ such that

$$\underline{\omega}(\underline{I}') \cdot \underline{m} = 0 \ . \tag{6.9}$$

On the other hand if, for a given value of the actions, $\underline{\omega} = \underline{\omega}(\underline{I})$ is rationally independent (which means that (6.9) is satisfied only for $\underline{m} = \underline{0}$), then there do not appear zero divisors in (6.8), though the divisors can become arbitrarily small with a proper choice of the vector $\underline{m}$. For this reason, terms of the form $\underline{\omega}(\underline{I}') \cdot \underline{m}$ are called *small divisors* and they can prevent the implementation of perturbation theory.

6.2.1 An example

We apply classical perturbation theory to the two–dimensional Hamiltonian function

$$\mathcal{H}(I_1, I_2, \varphi_1, \varphi_2) = \frac{I_1^2}{2} + \frac{I_2^2}{2} + \varepsilon\Big[\cos(\varphi_1 + \varphi_2) + 2\cos(\varphi_1 - \varphi_2)\Big] \ ,$$

which can be shortly written as

$$\mathcal{H}(I_1, I_2, \varphi_1, \varphi_2) = h(I_1, I_2) + \varepsilon f(\varphi_1, \varphi_2), \tag{6.10}$$

where

$$h(I_1, I_2) = \frac{I_1^2}{2} + \frac{I_2^2}{2}$$

and

$$f(\varphi_1, \varphi_2) = \cos(\varphi_1 + \varphi_2) + 2\cos(\varphi_1 - \varphi_2).$$

Let us perform the change of coordinates

$$\begin{aligned}
I_1 &= I_1' + \varepsilon \frac{\partial \Phi}{\partial \varphi_1}(I_1', I_2', \varphi_1, \varphi_2) \\
I_2 &= I_2' + \varepsilon \frac{\partial \Phi}{\partial \varphi_2}(I_1', I_2', \varphi_1, \varphi_2) \\
\varphi_1' &= \varphi_1 + \varepsilon \frac{\partial \Phi}{\partial I_1'}(I_1', I_2', \varphi_1, \varphi_2) \\
\varphi_2' &= \varphi_2 + \varepsilon \frac{\partial \Phi}{\partial I_2'}(I_1', I_2', \varphi_1, \varphi_2) \ .
\end{aligned}$$

Expanding the Hamiltonian (6.10) in Taylor series up to the second order, one obtains:

$$\begin{aligned}
&h\left(I_1' + \varepsilon \frac{\partial \Phi}{\partial \varphi_1}, I_2' + \varepsilon \frac{\partial \Phi}{\partial \varphi_2}\right) + \varepsilon f(\varphi_1, \varphi_2) \\
&= h(I_1', I_2') + \varepsilon \frac{\partial h}{\partial I_1}(I_1', I_2') \frac{\partial \Phi}{\partial \varphi_1} + \varepsilon \frac{\partial h}{\partial I_2}(I_1', I_2') \frac{\partial \Phi}{\partial \varphi_2} + \varepsilon f(\varphi_1, \varphi_2) + O(\varepsilon^2) \ ,
\end{aligned}$$

where

$$\frac{\partial h}{\partial I_1}(I_1', I_2') = I_1' \equiv \omega_1 \ , \qquad \frac{\partial h}{\partial I_2}(I_1', I_2') = I_2' \equiv \omega_2 .$$

The first–order terms in ε must be zero; this yields the generating function as the solution of the equation

$$\omega_1 \frac{\partial \Phi}{\partial \varphi_1} + \omega_2 \frac{\partial \Phi}{\partial \varphi_2} = -f(\varphi_1, \varphi_2) \ .$$

Expanding in Fourier series and taking into account the explicit form of the perturbation, one obtains

$$\sum_{m,n} i(\omega_1 m + \omega_2 n)\Phi_{m,n}(I_1', I_2') e^{i(m\varphi_1 + n\varphi_2)} = -\Big[\cos(\varphi_1 + \varphi_2) + 2\cos(\varphi_1 - \varphi_2)\Big] \ .$$

Using the relations $\cos(k_1\varphi_1 + k_2\varphi_2) = \frac{1}{2}(e^{i(k_1\varphi_1 + k_2\varphi_2)} + e^{-i(k_1\varphi_1 + k_2\varphi_2)})$ for some integers k_1, k_2, and equating the coefficients with the same Fourier indexes, one gets:

$$\begin{aligned}
\Phi_{1,1} &= -\frac{1}{2i(\omega_1 + \omega_2)} \ , \qquad & \Phi_{-1,-1} &= \frac{1}{2i(\omega_1 + \omega_2)} \ , \\
\Phi_{1,-1} &= -\frac{1}{i(\omega_1 - \omega_2)} \ , \qquad & \Phi_{-1,1} &= -\frac{1}{i(-\omega_1 + \omega_2)} \ .
\end{aligned}$$

Casting together the above terms, the generating function is given by

$$\begin{aligned}
\Phi(I_1', I_2', \varphi_1, \varphi_2) &= -\frac{1}{\omega_1 + \omega_2}\left(\frac{e^{i(\varphi_1+\varphi_2)} - e^{-i(\varphi_1+\varphi_2)}}{2i}\right) \\
&- \frac{2}{\omega_1 - \omega_2}\left(\frac{e^{i(\varphi_1-\varphi_2)} - e^{-i(\varphi_1-\varphi_2)}}{2i}\right) ,
\end{aligned}$$

namely

$$\Phi(I_1', I_2', \varphi_1, \varphi_2) = -\frac{1}{\omega_1 + \omega_2}\sin(\varphi_1 + \varphi_2) - \frac{2}{\omega_1 - \omega_2}\sin(\varphi_1 - \varphi_2) \ .$$

Notice that the generating function is not defined when there appear the following zero divisors:

$$\omega_1 \pm \omega_2 = 0 \ , \qquad \text{namely} \qquad I_1' = \pm I_2' \ .$$

The new unperturbed Hamiltonian coincides with the old unperturbed Hamiltonian (expressed in the new set of variables), since the average of the perturbing function is zero:

$$h'(I_1', I_2') = \frac{I_1'^2}{2} + \frac{I_2'^2}{2} \ .$$

6.2.2 Computation of the precession of the perihelion

A straightforward application of classical perturbation theory allows us to compute the amount of the precession of the perihelion. A first–order computation is obtained starting with the restricted, planar, circular three–body model. In particular, we identify the three bodies $\mathcal{P}_0$, $\mathcal{P}_1$ and $\mathcal{P}_2$ with the Sun, Mercury and Jupiter. In terms of the Delaunay action–angle variables, the perturbing function can be expanded as in (4.7). The perturbing parameter ε represents the Jupiter–Sun mass ratio. We implement a first–order perturbation theory, which provides a new integrable Hamiltonian function of the form

$$h'(L',G') = -\frac{1}{2L'^2} \; - \; G' \; + \; \varepsilon \; R_{00}(L',G') \; ,$$

where $R_{00}(L,G) = -\frac{L^4}{4}(1+\frac{9}{16}L^4+\frac{3}{2}\mathrm{e}^2)+O(\mathrm{e}^3)$. Hamilton's equations yield

$$\dot{g}' = \frac{\partial h'(L',G')}{\partial G'} = -1 + \varepsilon \frac{\partial R_{00}(L',G')}{\partial G'} \; .$$

Recall that $g = g_0 - t$, being g_0 the argument of the perihelion. Neglecting terms of order $O(\mathrm{e}^3)$ in R_{00}, one gets that to the lowest order the argument of perihelion g_0 varies as

$$\dot{g}_0 = \varepsilon \; \frac{\partial R_{00}(L',G')}{\partial G'} = \frac{3}{4}\varepsilon L'^2 G' \; .$$

Notice that up to the first order in ε one has $L' = L$, $G' = G$. Taking $\varepsilon = 9.54 \cdot 10^{-4}$ (the actual value of the Jupiter–Sun mass ratio), $a = 0.0744$ (setting to one the Jupiter–Sun distance) and $\mathrm{e} = 0.2056$, one obtains

$$\dot{g}_0 = 155.25 \; \frac{\text{arcsecond}}{\text{century}} \; ,$$

which represents the contribution due to Jupiter to the precession of the perihelion of Mercury. A more refined value is obtained taking into account higher–order terms in the eccentricity.

6.3 Resonant perturbation theory

Consider the following Hamiltonian system with n degrees of freedom

$$\mathcal{H}(\underline{I},\underline{\varphi}) = h(\underline{I}) + \varepsilon f(\underline{I},\underline{\varphi}) \; , \qquad \underline{I} \in \mathbf{R}^n \; , \; \; \underline{\varphi} \in \mathbf{T}^n$$

and let $\underline{\omega}(\underline{I}) = \frac{\partial h(\underline{I})}{\partial \underline{I}}$ be the frequency vector of the motion. We assume that the frequencies satisfy ℓ resonance relations, with $\ell < n$, of the form

$$\underline{\omega} \cdot \underline{m}_k = 0 \qquad \text{for } k = 1, \ldots, \ell \; ,$$

for some vectors $\underline{m}_1$, ..., $\underline{m}_\ell \in \mathbf{Z}^n$. A *resonant perturbation theory* can be implemented to eliminate the non–resonant terms. More precisely, the aim is to construct a change of variables $\mathcal{C} : (\underline{I}, \underline{\varphi}) \to (\underline{I}', \underline{\varphi}')$ such that the new Hamiltonian takes the form

$$\mathcal{H}'(\underline{I}', \underline{\varphi}') = h'(\underline{I}', \underline{m}_1 \cdot \underline{\varphi}', \ldots, \underline{m}_\ell \cdot \underline{\varphi}') + \varepsilon^2 f'(\underline{I}', \underline{\varphi}') \ , \tag{6.11}$$

where h' depends on $\underline{\varphi}'$ only through the combinations $\underline{m}_k \cdot \underline{\varphi}'$ with $k = 1, \ldots, \ell$. To this end, let us first define the angles

$$\begin{aligned} \vartheta_j &= \underline{m}_j \cdot \underline{\varphi} \ , \qquad j = 1, \ldots, \ell \\ \vartheta_{j'} &= \underline{m}_{j'} \cdot \underline{\varphi} \ , \qquad j' = \ell + 1, \ldots, n \ , \end{aligned}$$

where the first ℓ angle variables are the resonant angles, while the latter $n - \ell$ angles are defined as arbitrary linear combinations with integer coefficients $\underline{m}_{j'}$. The corresponding actions are defined as

$$\begin{aligned} I_j &= \underline{m}_j \cdot \underline{J} \ , \qquad j = 1, \ldots, \ell \\ I_{j'} &= \underline{m}_{j'} \cdot \underline{J} \ , \qquad j' = \ell + 1, \ldots, n \ . \end{aligned}$$

Next we construct a canonical transformation which removes (to higher orders) the dependence on the short–period angles $(\vartheta_{\ell+1}, \ldots, \vartheta_n)$, while the lowest–order Hamiltonian will necessarily depend upon the resonant angles. To this end, let us first decompose the perturbation, expressed in terms of the variables $(\underline{J}, \underline{\vartheta})$, as

$$f(\underline{J}, \underline{\vartheta}) = \overline{f}(\underline{J}) + f_r(\underline{J}, \vartheta_1, \ldots, \vartheta_\ell) + f_n(\underline{J}, \underline{\vartheta}) \ , \tag{6.12}$$

where $\overline{f}(\underline{J})$ is the average of the perturbation over the angles, $f_r(\underline{J}, \vartheta_1, \ldots, \vartheta_\ell)$ is the part depending on the resonant angles and $f_n(\underline{J}, \underline{\vartheta})$ is the non–resonant part. By analogy with classical perturbation theory, we implement a canonical transformation of the form (6.4), such that the new Hamiltonian takes the form (6.11). Using (6.12) and expanding up to the second order in the perturbing parameter, one obtains:

$$\begin{aligned} h\left(\underline{J}' + \varepsilon \frac{\partial \Phi}{\partial \underline{\vartheta}}\right) + \varepsilon f(\underline{J}', \underline{\vartheta}) + O(\varepsilon^2) &= h(\underline{J}') + \varepsilon \sum_{k=1}^{n} \frac{\partial h}{\partial J'_k} \frac{\partial \Phi}{\partial \vartheta_k} \\ &+ \varepsilon \overline{f}(\underline{J}') + \varepsilon f_r(\underline{J}', \vartheta_1, \ldots, \vartheta_\ell) + \varepsilon f_n(\underline{J}', \underline{\vartheta}) + O(\varepsilon^2) \ . \end{aligned}$$

Recalling (6.11) and equating terms of the same orders is ε, one gets that

$$h'(\underline{J}', \vartheta_1, \ldots, \vartheta_\ell) = h(\underline{J}') + \varepsilon \overline{f}(\underline{J}') + \varepsilon f_r(\underline{J}', \vartheta_1, \ldots, \vartheta_\ell) \ , \tag{6.13}$$

provided

$$\sum_{k=1}^{n} \omega'_k \frac{\partial \Phi}{\partial \vartheta_k} = -f_n(\underline{J}', \underline{\vartheta}) \ , \tag{6.14}$$

where $\omega'_k = \omega'_k(\underline{J}') \equiv \frac{\partial h(\underline{J}')}{\partial J_k}$. The solution of (6.14) provides the generating function allowing us to reduce the Hamiltonian to the required form (6.11); moreover, the conjugated action variables, say $J'_{\ell+1}, \ldots, J'_n$, are constants of the motion up to the second order in ε. We remark that using the new frequencies ω'_k, the resonant relations take the form $\omega'_k = 0$ for $k = 1, \ldots, \ell$.

6.3.1 Three–body resonance

As an example of the application of resonant perturbation theory we consider the three–body Hamiltonian (4.2) with the perturbing function expanded as in (4.7). Let $\underline{\omega} \equiv (\omega_\ell, \omega_g)$ be the frequency of motion; we assume that the following resonance relation is satisfied:

$$\omega_\ell + 2\omega_g = 0 \ .$$

Next, we perform the canonical change of variables

$$\begin{aligned} \vartheta_1 &= \ell + 2g \ , \qquad & J_1 &= \frac{1}{2} G \ , \\ \vartheta_2 &= 2\ell \ , \qquad & J_2 &= \frac{1}{2} L - \frac{1}{4} G \ . \end{aligned}$$

In the new coordinates the unperturbed Hamiltonian takes the form

$$h(\underline{J}) \equiv -\frac{\mu^2}{2(J_1 + 2J_2)^2} - 2J_1 \ ,$$

while the perturbing function is given by

$$\begin{aligned} R(J_1, J_2, \vartheta_1, \vartheta_2) &\equiv R_{00}(\underline{J}) + R_{10}(\underline{J}) \cos\left(\frac{1}{2}\vartheta_2\right) + R_{11}(\underline{J}) \cos\left(\frac{1}{2}\vartheta_1 + \frac{1}{4}\vartheta_2\right) \\ &+ R_{12}(\underline{J}) \cos(\vartheta_1) + R_{22}(\underline{J}) \cos\left(\vartheta_1 + \frac{1}{2}\vartheta_2\right) \\ &+ R_{32}(\underline{J}) \cos(\vartheta_1 + \vartheta_2) + R_{33}(\underline{J}) \cos\left(\frac{3}{2}\vartheta_1 + \frac{3}{4}\vartheta_2\right) \\ &+ R_{44}(\underline{J}) \cos(2\vartheta_1 + \vartheta_2) + R_{55}(\underline{J}) \cos\left(\frac{5}{2}\vartheta_1 + \frac{5}{4}\vartheta_2\right) + \ldots \end{aligned}$$

with the coefficients R_{ij} as in (4.8). Let us split the perturbation as $R = \overline{R}(\underline{J}) + R_r(\underline{J}, \vartheta_1) + R_n(\underline{J}, \underline{\vartheta})$, where $\overline{R}(\underline{J})$ is the average over the angles, $R_r(\underline{J}, \vartheta_1) = R_{12}(\underline{J}) \cos(\vartheta_1)$ is the resonant part, while R_n contains all the remaining non–resonant terms. We look for a change of coordinates close to the identity with generating function $\Phi = \Phi(\underline{J}', \underline{\vartheta})$ such that

$$\underline{\omega}'(\underline{J}') \cdot \frac{\partial \Phi(\underline{J}', \underline{\vartheta})}{\partial \underline{\vartheta}} = -R_n(\underline{J}', \underline{\vartheta}) \ ,$$

being $\underline{\omega}'(\underline{J}') \equiv \frac{\partial h(\underline{J}')}{\partial \underline{J}}$. The above expression is well defined since $\underline{\omega}'$ is non–resonant for the Fourier components appearing in R_n. Finally, according to (6.13) the new unperturbed Hamiltonian is given by

$$h'(\underline{J}', \vartheta_1) \equiv h(\underline{J}') + \varepsilon R_{00}(\underline{J}') + \varepsilon R_{12}(\underline{J}') \cos(\vartheta_1) \ .$$

6.4 Degenerate perturbation theory

Consider the Hamiltonian function with n degrees of freedom

$$\mathcal{H}(\underline{I},\underline{\varphi}) = h(I_1,\dots,I_d) + \varepsilon f(\underline{I},\underline{\varphi}) \ , \qquad d<n \ , \tag{6.15}$$

where the unperturbed Hamiltonian depends on a subset of the action variables, being degenerate in $I_{d+1},\dots,I_n$. As in the resonant perturbation theory, we look for a canonical transformation $\mathcal{C} : (\underline{I},\underline{\varphi}) \to (\underline{I}',\underline{\varphi}')$ such that the new Hamiltonian becomes

$$\mathcal{H}'(\underline{I}',\underline{\varphi}') = h'(\underline{I}') + \varepsilon h_1'(\underline{I}',\varphi_{d+1}',\dots,\varphi_n') + \varepsilon^2 f'(\underline{I}',\underline{\varphi}') \ , \tag{6.16}$$

where the term $h' + \varepsilon h_1'$ admits d integrals of motion. Let us split the perturbing function in (6.15) as

$$f(\underline{I},\underline{\varphi}) = \overline{f}(\underline{I}) + f_d(\underline{I},\varphi_{d+1},..,\varphi_n) + f_n(\underline{I},\underline{\varphi}) \ , \tag{6.17}$$

where $\overline{f}$ is the average over the angle variables, f_d is independent of φ_1, ..., φ_d and f_n is the remainder, namely $f_n = f - \overline{f} - f_d$. We want to determine a near–to–identity change of variables of the form (6.4), such that in view of (6.17) the Hamiltonian (6.15) is transformed into (6.16), namely

$$\begin{aligned}
h&\left(I_1' + \varepsilon\frac{\partial\Phi}{\partial\varphi_1},\dots,I_d' + \varepsilon\frac{\partial\Phi}{\partial\varphi_d}\right) + \varepsilon f\left(\underline{I}' + \varepsilon\frac{\partial\Phi}{\partial\underline{\varphi}},\underline{\varphi}\right) \\
&= h(I_1',\dots,I_d') + \varepsilon\sum_{k=1}^d \frac{\partial h}{\partial I_k}\frac{\partial\Phi}{\partial\varphi_k} + \varepsilon\overline{f}(\underline{I}') + \varepsilon f_d(\underline{I}',\varphi_{d+1},\dots,\varphi_n) \\
&\quad + \varepsilon f_n(\underline{I}',\underline{\varphi}) + O(\varepsilon^2) \\
&\equiv h'(\underline{I}') + \varepsilon h_1'(\underline{I}',\varphi_{d+1},\dots,\varphi_n) + O(\varepsilon^2) \ ,
\end{aligned}$$

where

$$\begin{aligned}
h'(\underline{I}') &\equiv h(I_1',\dots,I_d') + \varepsilon\overline{f}(\underline{I}') \\
h_1'(\underline{I}',\varphi_{d+1},\dots,\varphi_n) &\equiv f_d(\underline{I}',\varphi_{d+1},\dots,\varphi_n) \ ,
\end{aligned}$$

provided Φ is determined so that

$$\sum_{k=1}^d \frac{\partial h}{\partial I_k}\frac{\partial\Phi}{\partial\varphi_k} + f_n(\underline{I}',\underline{\varphi}) = 0 \ . \tag{6.18}$$

As in the previous sections, let us expand Φ and f_n in Fourier series as

$$\begin{aligned}
\Phi(\underline{I}',\underline{\varphi}) &= \sum_{\underline{m}\in\mathbf{Z}^n\backslash\{\underline{0}\}} \hat{\Phi}_{\underline{m}}(\underline{I}')\ e^{i\underline{m}\cdot\underline{\varphi}} \\
f_n(\underline{I}',\underline{\varphi}) &= \sum_{\underline{m}\in\mathcal{I}_n} \hat{f}_{n,\underline{m}}(\underline{I}')\ e^{i\underline{m}\cdot\underline{\varphi}} \ ,
\end{aligned}$$

where $\mathcal{I}_n$ denotes a suitable set of integer vectors defining the Fourier indexes of f_n. From (6.18) and setting $\omega_k \equiv \frac{\partial h}{\partial I_k}$, one obtains

$$i \sum_{\underline{m}\in\mathbf{Z}^n\backslash\{\underline{0}\}} \hat{\Phi}_{\underline{m}}(\underline{I}') \sum_{k=1}^{d} m_k\omega_k \ e^{i\underline{m}\cdot\underline{\varphi}} = \ - \sum_{\underline{m}\in\mathcal{I}_n} \hat{f}_{n,\underline{m}}(\underline{I}') \ e^{i\underline{m}\cdot\underline{\varphi}} \ . \tag{6.19}$$

Due to the fact that $\omega_k = 0$ for $k = d+1, \ldots, n$, we obtain that

$$\underline{\omega} \cdot \underline{m} = \sum_{k=1}^{d} m_k\omega_k \ . \tag{6.20}$$

Equation (6.19) yields that the generating function takes the form

$$\Phi(\underline{I}', \underline{\varphi}) = i \sum_{\underline{m}\in\mathcal{I}_n} \frac{\hat{f}_{n,\underline{m}}(\underline{I}')}{\underline{\omega}\cdot\underline{m}} \ e^{i\underline{m}\cdot\underline{\varphi}} \ .$$

The generating function is well defined provided that $\underline{\omega}\cdot\underline{m} \neq 0$ for any $\underline{m} \in \mathcal{I}_n$, which in view of (6.20) is equivalent to requiring that

$$\sum_{k=1}^{d} m_k\omega_k \neq 0 \qquad \text{for } \ \underline{m} \in \mathcal{I}_n \ .$$

6.4.1 The precession of the equinoxes

An application of the degenerate perturbation theory to Celestial Mechanics is offered by the computation of the precession of the equinoxes, namely the constant retrograde precession of the spin–axis provoked by gravitational interactions. In particular, we compute the Earth's equinox precession due to the influence of the Sun and of the Moon. Assume that the Earth $\mathcal{E}$ is an oblate rigid body moving around the (point–mass) Sun $\mathcal{S}$ on a Keplerian orbit with semimajor axis a and eccentricity e; recalling (5.8) and (5.10), in the gyroscopic case $I_1 = I_2$ the Hamiltonian describing the motion of $\mathcal{E}$ around $\mathcal{S}$ is given by

$$\mathcal{H}(L, G, H, \ell, g, h, t) = \frac{G^2}{2I_1} + \frac{I_1 - I_3}{2I_1 I_3} L^2 + \tilde{V}(L, G, H, \ell, g, h, t) \ ,$$

where I_1, I_2, I_3 are the principal moments of inertia and where the perturbation is implicitly defined by

$$\tilde{V} \equiv - \int_{\mathcal{E}} \frac{\mathcal{G} m_{\mathcal{S}} m_{\mathcal{E}}}{|\underline{r}_{\mathcal{E}} + \underline{x}|} \frac{d\underline{x}}{|\mathcal{E}|} \ ,$$

being $\underline{r}_{\mathcal{E}}$ the orbital radius of the Earth and $|\mathcal{E}|$ the volume of $\mathcal{E}$. Setting $r_{\mathcal{E}} = |\underline{r}_{\mathcal{E}}|$ and $x = |\underline{x}|$, we can expand $\tilde{V}$ using the Legendre polynomials as

$$\tilde{V} = -\frac{\mathcal{G} m_{\mathcal{S}} m_{\mathcal{E}}}{r_{\mathcal{E}}} \int_{\mathcal{E}} \frac{d\underline{x}}{|\mathcal{E}|} \left[1 - \frac{\underline{x}\cdot\underline{r}_{\mathcal{E}}}{r_{\mathcal{E}}^2} + \frac{1}{2r_{\mathcal{E}}^2}\left(3\frac{(\underline{x}\cdot\underline{r}_{\mathcal{E}})^2}{r_{\mathcal{E}}^2} - x^2\right)\right] + O\left[\left(\frac{\underline{x}}{\underline{r}_{\mathcal{E}}}\right)^3\right] \ .$$

Assume that the Earth rotates around a principal axis, namely that the non–principal rotation angle J is zero or, equivalently, that the actions G and L are equal. Let $\bar{G}$ and $\bar{H}$ be the initial values of G and H at $t = 0$ and let α denotes the angle between $\underline{r}_{\mathcal{E}}$ and $\underline{k}$ ($\underline{k}$ being the vertical axis of the body frame). Retaining only the second order of the development of the perturbing function in terms of the Legendre polynomials, one obtains

$$\tilde{V} = \varepsilon\bar{\omega}\frac{\bar{G}^2}{\bar{H}}\ \frac{(1-\mathrm{e}\cos\lambda_{\mathcal{E}})^3}{(1-\mathrm{e}^2)^3}\ \cos^2\alpha$$

with $\varepsilon = \frac{3}{2}\frac{I_3-I_1}{I_3}$, $\bar{\omega} = \frac{\mathcal{G}m_{\mathcal{S}}}{a^3}\ I_3\ \frac{\bar{H}}{\bar{G}^2}$ and where $\lambda_{\mathcal{E}}$ is the longitude of the Earth. Elementary computations show that

$$\cos\alpha = \sin(\lambda_{\mathcal{E}} - h)\ \sqrt{1-\frac{H^2}{G^2}}\ .$$

Neglecting first–order terms in the orbital eccentricity, we have that $\frac{(1-\mathrm{e}\cos\lambda_{\mathcal{E}})^3}{(1-\mathrm{e}^2)^3} \simeq 1$. A first–order degenerate perturbation theory provides the new unperturbed Hamiltonian in the form (we omit the primes to denote new variables):

$$\mathcal{H}_1(G,H) = \frac{G^2}{2I_3} + \varepsilon\bar{\omega}\ \frac{\bar{G}^2}{\bar{H}}\frac{G^2-H^2}{2G^2}\ .$$

Finally, the average angular velocity of precession is given by

$$\dot{h} = \frac{\partial\mathcal{H}_1(G,H)}{\partial H} = -\varepsilon\bar{\omega}\ \frac{\bar{G}^2}{\bar{H}}\ \frac{H}{G^2}\ .$$

At $t = 0$ it is

$$\dot{h} = -\varepsilon\bar{\omega} = -\varepsilon\omega_y^2\omega_d^{-1}\cos K\ , \tag{6.21}$$

where we used $\bar{\omega} = \omega_y^2\omega_d^{-1}\cos K$ with ω_y being the frequency of revolution and ω_d the frequency of rotation, while K denotes the obliquity.

Astronomical measurements show that $\frac{I_3-I_1}{I_3} \simeq \frac{1}{298.25}$, $K \simeq 23.45^o$. The contribution $\dot{h}^{(S)}$ due to the Sun is thus obtained inserting $\omega_y = 1$ year, $\omega_d = 1$ day in (6.21), yielding $\dot{h}^{(S)} = -2.51857 \cdot 10^{-12}$ rad/sec, which corresponds to a retrograde precessional period of 79 107.9 years. A similar computation shows that the contribution $\dot{h}^{(M)}$ of the Moon amounts to $\dot{h}^{(M)} = -5.49028 \cdot 10^{-12}$ rad/sec, corresponding to a precessional period of 36 289.3 years. The total precessional period is obtained as the sum of $\dot{h}^{(S)}$ and $\dot{h}^{(M)}$, providing an overall retrograde precessional period of 24 877.3 years, in good agreement with the value corresponding to astronomical observations and amounting to 25 700 years for the precession of the equinoxes.

6.5 Birkhoff's normal form

6.5.1 Normal form around an equilibrium position

Assume that the Hamiltonian $\mathcal{H} = \mathcal{H}(\underline{p}, \underline{q})$, $(\underline{p}, \underline{q}) \in \mathbf{R}^{2n}$, admits the origin as a stable equilibrium position; as a consequence, the eigenvalues of the quadratic part are all distinct and purely imaginary. In a neighborhood of the equilibrium position, after a series expansion and eventual diagonalization of the quadratic terms, we can write the Hamiltonian in the form

$$\mathcal{H}(\underline{p}, \underline{q}) = \frac{1}{2} \sum_{j=1}^{n} \omega_j (p_j^2 + q_j^2) + \mathcal{H}_3(\underline{p}, \underline{q}) + \mathcal{H}_4(\underline{p}, \underline{q}) + \dots \ , \tag{6.22}$$

where $\omega_j \in \mathbf{R}$ for $j = 1, \dots, n$ are called the *frequencies* of the motion and the terms $\mathcal{H}_k$ are polynomials of degree k in $\underline{p}$ and $\underline{q}$. The following definitions introduce the resonant relations and the Birkhoff normal form associated to the Hamiltonian (6.22).

Definition. The frequencies $\omega_1, \dots, \omega_n$ are said to satisfy a *resonance relation* of order $K > 0$, if there exists a non–zero integer vector $(k_1, \dots, k_n)$ such that $k_1\omega_1 + \dots + k_n\omega_n = 0$ and $|k_1| + \dots + |k_n| = K$.

Definition. Let K be a positive number; a *Birkhoff normal form* for the Hamiltonian (6.22) is a polynomial of degree K in a set of variables $\underline{P}$, $\underline{Q}$, such that it is a polynomial of degree $[\frac{K}{2}]$ in the quantity $I'_j = \frac{1}{2}(P_j^2 + Q_j^2)$ for $j = 1, \dots, n$.

The construction of the Birkhoff normal form is the content of the following theorem.

Theorem. *Let K be a positive integer; assume that the frequencies $\omega_1, \dots, \omega_n$ do not satisfy any resonance relation of order less than or equal to K. Then, there exists a canonical transformation from $(\underline{p}, \underline{q})$ to $(\underline{P}, \underline{Q})$ such that the Hamiltonian* (6.22) *reduces to a Birkhoff normal form of degree K.*

Proof. Let us introduce action–angle variables $\underline{I} = (I_1, \dots, I_n) \in \mathbf{R}^n$, $\underline{\varphi} = (\varphi_1, \dots, \varphi_n) \in \mathbf{T}^n$, such that

$$\begin{aligned} p_j &= \sqrt{2I_j}\cos\varphi_j \\ q_j &= \sqrt{2I_j}\sin\varphi_j \ , \qquad j = 1, \dots, n \ . \end{aligned} \tag{6.23}$$

Then, the Hamiltonian (6.22) can be written as

$$\mathcal{H}_1(\underline{I}, \underline{\varphi}) = \sum_{j=1}^{n} \omega_j I_j + \mathcal{H}_{1,3}(\underline{I}, \underline{\varphi}) + \mathcal{H}_{1,4}(\underline{I}, \underline{\varphi}) + \dots \ , \tag{6.24}$$

where the terms $\mathcal{H}_{1,k}$ are polynomials of degree $[k/2]$ in $I_1,\ldots,I_n$. Let $\underline{I}'\cdot\underline{\varphi}+\Phi(\underline{I}',\underline{\varphi})$ be the generating function of a canonical transformation close to the identity from $(\underline{I},\underline{\varphi})$ to $(\underline{I}',\underline{\varphi}')$:

$$\begin{aligned}\underline{I} &= \underline{I}' + \frac{\partial\Phi}{\partial\underline{\varphi}}\\ \underline{\varphi}' &= \underline{\varphi} + \frac{\partial\Phi}{\partial\underline{I}'}\ .\end{aligned}\tag{6.25}$$

Let us decompose Φ as the sum of polynomials

$$\Phi = \Phi_3 + \Phi_4 + \cdots + \Phi_K\ ,$$

where Φ_k, $k=3,\ldots,K$, is a polynomial of order $[\frac{k}{2}]$ in $I_1,\ldots,I_n$. Inserting the first of (6.25) in the Hamiltonian (6.24), one obtains the transformed Hamiltonian

$$\mathcal{H}_2(\underline{I}',\underline{\varphi}) = \underline{\omega}\cdot\underline{I}' + \underline{\omega}\cdot\frac{\partial\Phi}{\partial\underline{\varphi}} + \mathcal{H}_{1,3}\left(\underline{I}' + \frac{\partial\Phi}{\partial\underline{\varphi}},\varphi\right) + \mathcal{H}_{1,4}\left(\underline{I}' + \frac{\partial\Phi}{\partial\underline{\varphi}},\varphi\right) + \ldots \tag{6.26}$$

Let us determine Φ_3 such that the Hamiltonian (6.26) reduces to the Birkhoff normal form up to degree 3. To this end, split $\mathcal{H}_3$ as

$$\mathcal{H}_{1,3}(\underline{I}',\underline{\varphi}) = \bar{\mathcal{H}}_{1,3}(\underline{I}') + \tilde{\mathcal{H}}_{1,3}(\underline{I}',\underline{\varphi})\ ,$$

where $\bar{\mathcal{H}}_{1,3}(\underline{I}')$ is the average of $\mathcal{H}_{1,3}$ over the angles and $\tilde{\mathcal{H}}_{1,3}(\underline{I}',\underline{\varphi})$ is the remainder. Using (6.26) we obtain

$$\mathcal{H}_2(\underline{I}',\underline{\varphi}) = \underline{\omega}\cdot\underline{I}'+\bar{\mathcal{H}}_{1,3}(\underline{I}')+\left[\underline{\omega}\cdot\frac{\partial\Phi_3}{\partial\underline{\varphi}}+\tilde{\mathcal{H}}_{1,3}(\underline{I}',\underline{\varphi})\right]+\underline{\omega}\cdot\frac{\partial\Phi_4}{\partial\underline{\varphi}}+\mathcal{H}_{1,4}\left(\underline{I}'+\frac{\partial\Phi}{\partial\underline{\varphi}},\varphi\right)+\ldots$$

Expanding Φ_3 and $\tilde{\mathcal{H}}_{1,3}$ in Fourier series as

$$\begin{aligned}\Phi_3(\underline{I}',\underline{\varphi}) &= \sum_{\underline{m}\in\mathbf{Z}^n}\hat{\Phi}_{3,\underline{m}}(\underline{I}')\ e^{i\underline{m}\cdot\underline{\varphi}}\\ \tilde{\mathcal{H}}_{1,3}(\underline{I}',\underline{\varphi}) &= \sum_{\underline{m}\in\mathbf{Z}^n\setminus\{\underline{0}\}}\hat{\mathcal{H}}_{1,3,\underline{m}}(\underline{I}')\ e^{i\underline{m}\cdot\underline{\varphi}}\ ,\end{aligned}\tag{6.27}$$

one obtains

$$i\sum_{\underline{m}\in\mathbf{Z}^n}\underline{\omega}\cdot\underline{m}\ \hat{\Phi}_{3,\underline{m}}(\underline{I}')e^{i\underline{m}\cdot\underline{\varphi}} + \sum_{\underline{m}\in\mathbf{Z}^n\setminus\{\underline{0}\}}\hat{\mathcal{H}}_{1,3,\underline{m}}(\underline{I}')\ e^{i\underline{m}\cdot\underline{\varphi}} = 0\ .$$

Therefore the unknown Fourier coefficients of Φ_3 are given by

$$\hat{\Phi}_{3,\underline{m}}(\underline{I}') = -\frac{\hat{\mathcal{H}}_{1,3,\underline{m}}(\underline{I}')}{i\ \underline{\omega}\cdot\underline{m}}\ .\tag{6.28}$$

Casting together (6.27) and (6.28), one obtains

$$\Phi_3(\underline{I}', \underline{\varphi}) = - \sum_{\underline{m} \in \mathbf{Z}^n \setminus \{\underline{0}\}} \frac{\hat{\mathcal{H}}_{1,3,\underline{m}}(\underline{I}')}{i \ \underline{\omega} \cdot \underline{m}} \ e^{i\underline{m} \cdot \underline{\varphi}} \ .$$

Therefore the transformed Hamiltonian depends only on the actions $\underline{I}'$ up to terms of the fourth order, thus yielding a Birkhoff normal form of degree 3; the normalized terms define an integrable system in the set of action–angle variables $(\underline{I}', \underline{\varphi}')$ which provide the set of variables $(\underline{P}, \underline{Q})$ through the transformation $P_j = \sqrt{2I_j'} \cos \varphi_j'$, $Q_j = \sqrt{2I_j'} \sin \varphi_j'$, $j = 1, \ldots, n$. The same procedure applied to higher orders leads to the determination of the generating function associated to the Birkhoff normal form of degree K. $\square$

The Birkhoff normal form can be applied to the resonant case (see [6]) as the classical perturbation theory extends to the resonant perturbation theory. More precisely, recalling the action–angle variables introduced in (6.23), one has the following definition.

Definition. Let $\mathcal{K}$ be a sublattice of $\mathbf{Z}^n$; a *resonant Birkhoff normal form* of degree K for resonances in $\mathcal{K}$ is a polynomial of degree $[\frac{K}{2}]$ in $I_1, \ldots, I_n$, depending on the angles only through combinations of the form $\underline{k} \cdot \underline{\varphi}$ for $\underline{k} \in \mathcal{K}$.

The extension of the Birkhoff normal form to the resonant case is the content of the following theorem.

Theorem. *Let K be a positive integer and let $\mathcal{K}$ be a sublattice of $\mathbf{Z}^n$; assume that the frequencies ω_1, ..., ω_n do not satisfy any resonance relation of order less than or equal to K, except for combinations of the form $\underline{k} \cdot \underline{\varphi}$ for $\underline{k} \in \mathcal{K}$. Then, there exists a canonical transformation such that the Hamiltonian* (6.22) *reduces to a resonant Birkhoff normal form of degree K for resonances in $\mathcal{K}$.*

Remark. The above results extend straightforwardly to mapping systems having the origin as an elliptic stable fixed point, so that all eigenvalues lie on the unitary circle of the complex plane. We briefly quote here the main result, referring to [162] for further details. Let $(p', q') = \mathcal{M}(p, q)$ be a two–dimensional area preserving map with $(p, q) \in \mathbf{R}^2$.

Definition. Let K be a positive number; close to an elliptic fixed point, a Birkhoff normal form of degree K for $\mathcal{M}$ is a polynomial in a set of variables P, Q, which is a polynomial of degree $[\frac{K}{2}] - 1$ in the quantity $I' = \frac{1}{2}(P^2 + Q^2)$.

The Birkhoff normal form for mappings is the content of the following theorem.

Theorem. *If the eigenvalue of the linear part of $\mathcal{M}$ at the elliptic fixed point is not a root of unity of degree less than or equal to K, then there exists a canonical change of variables which reduces the map to a Birkhoff normal form of degree K.*

6.5.2 Normal form around closed trajectories

Let us consider a non–autonomous Hamiltonian system of the form

$$\mathcal{H} = \mathcal{H}(\underline{p}, \underline{q}, t) \ ,$$

where $(\underline{p}, \underline{q}) \in \mathbf{R}^{2n}$ and $\mathcal{H}$ is a 2π–periodic function of the time. Closed trajectories for $\mathcal{H}$ are generally not isolated, but they rather form families. In a neighborhood of a closed trajectory one can reduce the Hamiltonian to the form

$$\mathcal{H}(\underline{p}, \underline{q}, t) = \frac{1}{2} \sum_{j=1}^{n} \omega_j (p_j^2 + q_j^2) + \mathcal{H}_3(\underline{p}, \underline{q}, t) + \mathcal{H}_4(\underline{p}, \underline{q}, t) + \dots \ , \tag{6.29}$$

where $\underline{\omega} = (\omega_1, \dots, \omega_n)$ is the so–called frequency vector. Referring the reader to [6], we introduce the notion of resonance relation and a result on the construction of the Birkhoff normal form for the Hamiltonian (6.29).

Definition. The frequencies $\omega_1, \dots, \omega_n$ are said to satisfy a *resonance relation* of order K, with $K > 0$, if there exists a non–zero integer vector $(k_0, k_1, \dots, k_n)$ such that $k_0 + k_1\omega_1 + \dots + k_n\omega_n = 0$ and $|k_1| + \dots + |k_n| = K$.

Theorem. *Assume that K is a positive integer and that the frequencies $\omega_1, \dots, \omega_n$ do not satisfy any resonance relation of order less than or equal to K. Then, there exists a canonical transformation 2π–periodic in time, such that* (6.29) *is reduced to an autonomous Birkhoff normal form of degree K with a time–dependent remainder of order $K + 1$.*

The extension of such result to the resonant case is formulated as follows (see [6]).

Definition. Let $\mathcal{K}$ be a sublattice of $\mathbf{Z}^{n+1}$; a *resonant Birkhoff normal form* of degree K for resonances in $\mathcal{K}$ is a polynomial of degree $[\frac{K}{2}]$ in the actions $I_1, \dots, I_n$, depending on the angles and on the time only through combinations of the form $k_0 t + \underline{k} \cdot \underline{\varphi}$ for $(k_0, \underline{k}) \in \mathcal{K}$.

Theorem. *Let K be a positive integer and let $\mathcal{K}$ be a sublattice of $\mathbf{Z}^{n+1}$; assume that the frequencies $\omega_1, \dots, \omega_n$ do not satisfy any resonance relation of order less than or equal to K, except for combinations of the form $k_0 + \underline{k} \cdot \underline{\varphi}$ for $(k_0, \underline{k}) \in \mathcal{K}$. Then, there exists a canonical transformation reducing the Hamiltonian to a resonant Birkhoff normal form of degree K in $\mathcal{K}$ up to terms of order $K + 1$.*

6.6 The averaging theorem

Consider the n–dimensional nearly–integrable Hamiltonian system

$$\mathcal{H}(\underline{I}, \underline{\varphi}) = h(\underline{I}) + \varepsilon f(\underline{I}, \underline{\varphi}) \ , \qquad \underline{I} \in \mathbf{R}^n \ , \ \underline{\varphi} \in \mathbf{T}^n \ ,$$

with associated Hamilton's equations

$$\begin{aligned} \underline{\dot{I}} &= \varepsilon \underline{F}(\underline{I}, \underline{\varphi}) \\ \underline{\dot{\varphi}} &= \underline{\omega}(\underline{I}) + \varepsilon \underline{G}(\underline{I}, \underline{\varphi}) \ , \end{aligned} \tag{6.30}$$

where $\underline{F}(\underline{I},\underline{\varphi}) \equiv -\frac{\partial f(\underline{I},\underline{\varphi})}{\partial \underline{\varphi}}$, $\underline{\omega}(\underline{I}) \equiv \frac{\partial h(\underline{I})}{\partial \underline{I}}$, $\underline{G}(\underline{I},\underline{\varphi}) = \frac{\partial f(\underline{I},\underline{\varphi})}{\partial \underline{I}}$. Let us decompose $\underline{F}$ as its average plus an oscillating part, say $\underline{F}(\underline{I},\underline{\varphi}) = \overline{\underline{F}}(\underline{I}) + \tilde{\underline{F}}(\underline{I},\underline{\varphi})$, so that we can write (6.30) as

$$\begin{aligned} \dot{\underline{I}} &= \varepsilon\overline{\underline{F}}(\underline{I}) + \varepsilon\tilde{\underline{F}}(\underline{I},\underline{\varphi}) \\ \dot{\underline{\varphi}} &= \underline{\omega}(\underline{I}) + \varepsilon G(\underline{I},\underline{\varphi}) \ . \end{aligned} \tag{6.31}$$

Averaging (6.31) with respect to the angles, we obtain the following differential equations in a new set of coordinates $\underline{J}$:

$$\dot{\underline{J}} = \varepsilon\overline{\underline{F}}(\underline{J}) \ . \tag{6.32}$$

Denoting by $\underline{I}_\varepsilon(t)$ the solution of (6.31) with initial data $\underline{I}_\varepsilon(0)$ and by $\underline{J}_\varepsilon(t)$ the solution of (6.32) with initial data $\underline{J}_\varepsilon(0) = \underline{I}_\varepsilon(0)$, we want to investigate the conditions for which the averaged system is a good approximation of the full system (see for example [83] for applications to Celestial Mechanics). More precisely, we aim to study the conditions for which

$$\lim_{\varepsilon\to 0} |\underline{I}_\varepsilon(t) - \underline{J}_\varepsilon(t)| = 0 \qquad \text{for } t \in \left[0, \frac{1}{\varepsilon}\right] \ . \tag{6.33}$$

We prove such statement in some particular cases. Let us consider first the one–dimensional case described by the Hamiltonian function

$$\mathcal{H}(I,\varphi) = \omega I + \varepsilon f(\varphi) \ ,$$

where ω is a non–zero real number and where the perturbation does not depend on the action. Setting $F(\varphi) = -\frac{df(\varphi)}{d\varphi}$, the equations of motion are given by

$$\begin{aligned} \dot{I} &= \varepsilon F(\varphi) \\ \dot{\varphi} &= \omega \ . \end{aligned} \tag{6.34}$$

In this case (6.33) is guaranteed by the following result.

Proposition. *Let $I_\varepsilon(t)$ and $J_\varepsilon(t)$ denote, respectively, the solutions at time t of* (6.34) *and of the averaged equation with initial conditions, respectively, $I_\varepsilon(0)$ and $J_\varepsilon(0) = I_\varepsilon(0)$. Then, for any $0 \le t \le \frac{1}{\varepsilon}$, one has*

$$\lim_{\varepsilon\to 0} |I_\varepsilon(t) - J_\varepsilon(t)| = 0 \ . \tag{6.35}$$

Proof. Let c be the average of $F(\varphi)$; then $J_\varepsilon(t) = I_\varepsilon(0) + \varepsilon ct$. Defining $\tilde{F}(\varphi) \equiv F(\varphi) - c$, we have

$$\begin{aligned} I_\varepsilon(t) - J_\varepsilon(t) &= I_\varepsilon(0) + \int_0^t \dot{I}_\varepsilon(\tau)d\tau - (I_\varepsilon(0) + \varepsilon ct) \\ &= \varepsilon \int_0^t \tilde{F}(\varphi(0) + \omega\tau)d\tau = \frac{\varepsilon}{\omega} \int_{\varphi(0)}^{\varphi(0)+\omega t} \tilde{F}(\psi)d\psi \ . \end{aligned}$$

If M denotes an upper bound on $\int_{\varphi(0)}^{\varphi(0)+\omega t} \tilde{F}(\psi)d\psi$ for $0 \le t \le \frac{1}{\varepsilon}$, then

$$|I_\varepsilon(t) - J_\varepsilon(t)| \le \frac{\varepsilon}{\omega} M \ ,$$

which yields (6.35). □

For higher–dimensional systems, let us consider the Hamiltonian function with $n > 1$ degrees of freedom:

$$\mathcal{H}(\underline{I}, \underline{\varphi}) = \underline{\omega} \cdot \underline{I} + \varepsilon f(\underline{\varphi}) \ ,$$

where $\underline{I} \in \mathbf{R}^n$, $\underline{\varphi} \in \mathbf{T}^n$, $\underline{\omega} \in \mathbf{R}^n \backslash \{\underline{0}\}$. The equations of motion are

$$\begin{aligned} \dot{\underline{I}} &= \varepsilon \underline{F}(\underline{\varphi}) \\ \dot{\underline{\varphi}} &= \underline{\omega} \ , \end{aligned} \tag{6.36}$$

with $\underline{F}(\underline{\varphi}) \equiv -\frac{\partial f(\underline{\varphi})}{\partial \underline{\varphi}}$. Let $\underline{c}$ be the average of $\underline{F}(\underline{\varphi})$; more precisely, for a suitable sublattice $\mathcal{K}$ of $\mathbf{Z}^n \backslash \{\underline{0}\}$, let

$$\underline{F}(\underline{\varphi}) = \underline{c} + \sum_{\underline{k} \in \mathcal{K}} \hat{\underline{F}}_{\underline{k}} e^{i\underline{k} \cdot \underline{\varphi}} \ .$$

Proposition. *Let $\underline{I}_\varepsilon(t)$ and $\underline{J}_\varepsilon(t)$ denote, respectively, the solutions at time t of* (6.36) *and of the averaged equations with initial conditions, respectively, $\underline{I}_\varepsilon(0)$ and $\underline{J}_\varepsilon(0) = \underline{I}_\varepsilon(0)$. If the set $\mathcal{K}_0 \equiv \{\underline{k} \in \mathcal{K} : \ \underline{k} \cdot \underline{\omega} = 0\}$ is empty, then for any $0 \le t \le \frac{1}{\varepsilon}$, one has*

$$\lim_{\varepsilon \to 0} |\underline{I}_\varepsilon(t) - \underline{J}_\varepsilon(t)| = 0 \ .$$

Proof. We can write

$$\begin{aligned} \dot{\underline{I}} &= \varepsilon \underline{c} + \varepsilon \sum_{\underline{k} \in \mathcal{K}} \hat{F}_{\underline{k}} e^{i\underline{k} \cdot \underline{\varphi}(0)} \ e^{i\underline{k} \cdot \underline{\omega} t} \\ &= \varepsilon \underline{c} + \varepsilon \sum_{\underline{k} \in \mathcal{K}_0} \hat{F}_{\underline{k}} e^{i\underline{k} \cdot \underline{\varphi}(0)} + \varepsilon \sum_{\underline{k} \in \mathcal{K} \backslash \mathcal{K}_0} \hat{F}_{\underline{k}} e^{i\underline{k} \cdot \underline{\varphi}(0)} e^{i\underline{k} \cdot \underline{\omega} t} \ , \end{aligned}$$

whose integration yields

$$\underline{I}_\varepsilon(t) - \underline{I}_\varepsilon(0) = \varepsilon \underline{c} t + t\varepsilon \sum_{\underline{k} \in \mathcal{K}_0} \hat{F}_{\underline{k}} e^{i\underline{k} \cdot \underline{\varphi}(0)} + \varepsilon \sum_{\underline{k} \in \mathcal{K} \backslash \mathcal{K}_0} \hat{F}_{\underline{k}} e^{i\underline{k} \cdot \underline{\varphi}(0)} \ \frac{e^{i\underline{k} \cdot \underline{\omega} t} - 1}{i\underline{k} \cdot \underline{\omega}} \ .$$

The sum over $\mathcal{K}_0$ generates secular terms; nevertheless, by assumption the set $\mathcal{K}_0$ is empty. As a consequence, the distance between the complete and averaged solutions becomes:

$$\underline{I}_\varepsilon(t) - \underline{J}_\varepsilon(t) = \varepsilon \sum_{\underline{k} \in \mathcal{K} \backslash \mathcal{K}_0} \hat{F}_{\underline{k}} e^{i\underline{k} \cdot \underline{\varphi}(0)} \ \frac{e^{i\underline{k} \cdot \underline{\omega} t} - 1}{i\underline{k} \cdot \underline{\omega}} \ ,$$

which vanishes as ε tends to zero. □

6.6.1 An example

Let us consider the Hamiltonian function with two degrees of freedom:

$$\begin{aligned}\mathcal{H}(L,G,\ell,g) &= L^2 - G + \varepsilon R(L,G,\ell,g) \\ &= L^2 - G + \varepsilon\Big(R_{00}(L,G) + R_{10}(L,G)\cos 2\ell \\ &+ R_{12}(L,G)\cos(\ell+2g)\Big)\ ,\end{aligned}$$

for some real functions $R_{00}(L,G)$, $R_{10}(L,G)$, $R_{12}(L,G)$. The frequency vector is $(\omega_\ell, \omega_g) = (2L, -1)$; assume that the following resonance condition holds:

$$\omega_\ell + 2\omega_g = 0\ .$$

We perform the symplectic change of variables from (L, G, ℓ, g) to $(I_1, I_2, \vartheta_1, \vartheta_2)$ defined as

$$\begin{aligned}\vartheta_1 &= \ell + 2g\ , & I_1 &= \frac{1}{2}G\ , \\ \vartheta_2 &= 2\ell\ , & I_2 &= \frac{1}{2}L - \frac{1}{4}G\ ;\end{aligned} \tag{6.37}$$

due to the resonance, ϑ_1 is a slow variable, while ϑ_2 is a fast variable. The new Hamiltonian becomes

$$\begin{aligned}\mathcal{H}(I_1,I_2,\vartheta_1,\vartheta_2) &= (I_1+2I_2)^2 - 2I_1 + \varepsilon R(I_1,I_2,\vartheta_1,\vartheta_2) \\ &= (I_1+2I_2)^2 - 2I_1 + \varepsilon\Big(R_{00}(I_1,I_2) \\ &+ R_{10}(I_1,I_2)\cos\vartheta_2 + R_{12}(I_1,I_2)\cos\vartheta_1\Big)\ ,\end{aligned} \tag{6.38}$$

where $R(I_1, I_2, \vartheta_1, \vartheta_2)$ (and its coefficients) is the transformed function of $R(L, G, \ell, g)$ (and of its coefficients). Hamilton's equations are

$$\begin{aligned}\dot I_1 &= \varepsilon R_{12}(I_1,I_2)\sin\vartheta_1 \\ \dot I_2 &= \varepsilon R_{10}(I_1,I_2)\sin\vartheta_2 \\ \dot\vartheta_1 &= 2(I_1+2I_2) - 2 + \varepsilon\frac{\partial R(I_1,I_2,\vartheta_1,\vartheta_2)}{\partial I_1} \\ \dot\vartheta_2 &= 4(I_1+2I_2) + \varepsilon\frac{\partial R(I_1,I_2,\vartheta_1,\vartheta_2)}{\partial I_2}\ .\end{aligned}$$

Averaging over the fast variable ϑ_2 and denoting by $(J_1, J_2, \varphi_1, \varphi_2)$ the averaged variables, one obtains the Hamiltonian

$$\begin{aligned}\overline{\mathcal{H}}(J_1,J_2,\varphi_1,\varphi_2) &= (J_1+2J_2)^2 - 2J_1 + \varepsilon\Big(R_{00}(J_1,J_2) + R_{12}(J_1,J_2)\cos\varphi_1\Big) \\ &= (J_1+2J_2)^2 - 2J_1 + \varepsilon\bar R(J_1,J_2,\varphi_1)\ ,\end{aligned}$$

where $\bar{R}(J_1, J_2, \varphi_1) \equiv R_{00}(J_1, J_2) + R_{12}(J_1, J_2)\cos\varphi_1$. The associated Hamilton's equations are

$$\begin{aligned}
\dot{J}_1 &= \varepsilon R_{12}(J_1, J_2)\sin\varphi_1 \\
\dot{J}_2 &= 0 \\
\dot{\varphi}_1 &= 2(J_1 + 2J_2) - 2 + \varepsilon \frac{\partial \bar{R}(J_1, J_2, \varphi_1)}{\partial J_1} \\
\dot{\varphi}_2 &= 4(J_1 + 2J_2) + \varepsilon \frac{\partial \bar{R}(J_1, J_2, \varphi_1)}{\partial J_2} \ .
\end{aligned}$$

As a special case we set $R_{00}(L, G) = L$, $R_{12}(L, G) = L^2G$, $R_{10}(L, G) = LG^2$; taking $\varepsilon = 0.01$ and setting the initial conditions in the transformed variables (6.37) as $I_1(0) = 0.9$, $I_2(0) = 0.5$, $\vartheta_1(0) = 0$, $\vartheta_2(0) = 0$, one obtains that the difference between the complete and averaged solutions (see Figure 6.1) is $|I_1(t) - J_1(t)| < 0.076$ for any $0 \leq t \leq 100$ in agreement with the averaging results discussed in Section 6.6.

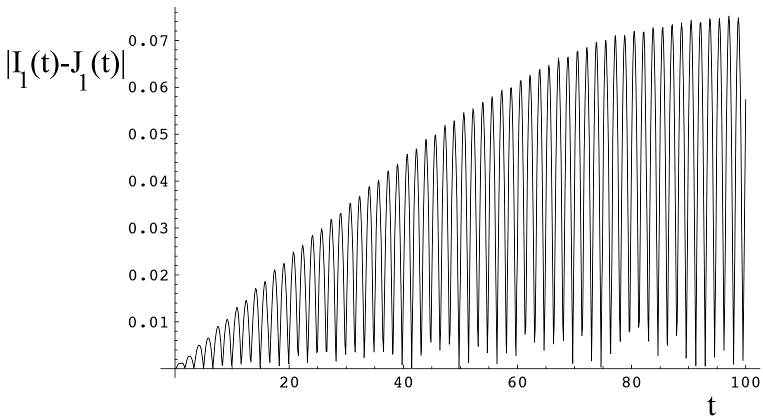

Fig. 6.1. The difference between the complete and averaged solutions associated to (6.38) for the special case $R_{00}(L, G) = L$, $R_{12}(L, G) = L^2G$, $R_{10}(L, G) = LG^2$ with $\varepsilon = 0.01$ and with initial conditions $I_1(0) = 0.9$, $I_2(0) = 0.5$, $\vartheta_1(0) = 0$, $\vartheta_2(0) = 0$.