2

Angle–Action Variables. Separable Systems

2.1 Periodic Motions

The trajectories of systems with one degree of freedom are the curves $H(q_1, p_1) = E$. As shown in Sect. 1.8, the equations of the motion are given by

$$t + \alpha_1 = q_1^* = \frac{\partial S}{\partial E} \stackrel{\text{def}}{=} F_{11}(q_1), \tag{2.1}$$

where α_1 is a constant and $S = S(q_1, E)$ is the solution of the Hamilton–Jacobi equation.

In the particular examples given in Sect. 1.9, we have found two kinds of periodic solutions:

- *Circulatory Motions.* Motions occurring when the variable q_1 is defined on a circle (for instance, q_1 is an angle defined from 0 to 2π) and is always increasing or decreasing (see Fig. 2.1a). The periodicity of the motion is due to the angular nature of q_1. The phase space of this system is a cylinder and the circulations are solutions closing on themselves after a complete tour encircling the cylinder.
- *Oscillatory Motions* or *Librations.* Motions occurring when the variable q_1 oscillates periodically between two boundaries a and b (see Fig. 2.1b). The variable q_1 may be either an angle (as in the pendulum) or a length (as in the harmonic oscillator). Accordingly, the phase space is either a cylinder (if $q_1 \in \mathbf{S}$) or a plane (if $q_1 \in \mathbf{R}$). Librations are closed curves with the particular property, in the case $q_1 \in \mathbf{S}$, that they close on themselves without encircling the cylinder.

It is not difficult to see that all bounded solutions of a Hamiltonian system with one degree of freedom are either periodic or asymptotic to an unstable equilibrium point. It is enough to remember that, since the phase flow preserves volumes in phase space (see [5], Chap. 1, Sect. 3.6), the only ordinary singular points allowed in the two-dimensional phase space of a Hamiltonian

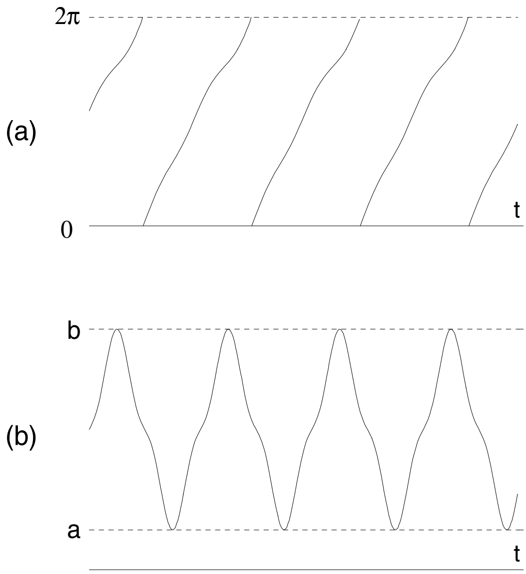

Fig. 2.1. Functions $q_1(t)$: (a) circulation; (b) libration

system are centers and saddle points. All bounded curves in this space that do not start or end in a saddle point correspond to a periodic motion.

2.1.1 Angle–Action Variables[1]

The equations resulting from the transformation $(q_1, p_1) \Rightarrow (q_1^*, p_1^*)$ are

$$\begin{aligned} q_1^* &= t + \alpha_1 \\ p_1^* &= \beta_1 = E, \end{aligned} \tag{2.2}$$

where α_1 and β_1 are constants. The phase space (q_1^*, p_1^*) is either a plane or a cylinder as discussed above. The phase trajectories are the lines $p_1^* = \beta_1$ and the phase velocity is $\dot{q}_1^* = 1$ on all trajectories. There are no explicit constraints imposed on α_1, β_1, which, however, exist and may be found by the analysis of $S(q_1, E)$. For instance, in the harmonic oscillator (Sect. 1.9.2), the solutions exist only in the domain formed by the upper half-plane $E/m \geq 0$. Another property not appearing in the functional expression of the Hamiltonian $H^* = p_1^*$ is the possible periodicity of the solutions (or of one set of solutions). For instance, in the harmonic oscillator, all solutions for $E/m > 0$ are periodic with period $T = 2\pi/\sqrt{k}$, that is, $q_1^* \in \mathbf{R}/T\mathbf{Z}$.

---

[1] Throughout this book, the order coordinate–momentum is adopted. Thus, we shall refer to these variables as “angle–action” variables, instead of “action–angle” as usually done everywhere.

To correct this lack of topological information on the motion in the phase space q_1^*, p_1^*, we introduce, in the case of periodic motions, a new angular variable $w_1 \in \mathbf{S}^1$. By definition, it increases 2π when q_1 performs a complete circulation or libration. From (2.1) we have

$$\begin{aligned} q_1^* &= t + \alpha_1 &&= F_{11}(q_1) \\ q_1^* + T &= t + \alpha_1 + T &&= F_{11}(q_1 + \oint \mathrm{d}q_1), \end{aligned} \tag{2.3}$$

where $\oint \mathrm{d}q_1$ means a complete circulation or libration of q_1. Then, in order to have, instead of q_1^*, a uniformized variable, it is enough to define[2]

$$w_1 = 2\pi \frac{t + \alpha_1}{T} = 2\pi \frac{q_1^*}{T}. \tag{2.4}$$

Obviously, the period T is the same for all initial conditions on a periodic orbit, but it is worth keeping in mind that it is not the same for all periodic solutions of a given system.

The momentum conjugate to w_1 may be easily obtained in terms of q_1, p_1. Let $\tilde{S}(q_1, J_1)$ be the Jacobian generating function of the canonical transformation $\tilde{\phi} : (q_1, p_1) \Rightarrow (w_1, J_1)$. Hence,

$$w_1 = \frac{\partial \tilde{S}}{\partial J_1} \qquad p_1 = \frac{\partial \tilde{S}}{\partial q_1}. \tag{2.5}$$

The following chain of calculations is simple and just uses elementary calculus:

$$\frac{\mathrm{d}w_1}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \tilde{S}}{\partial J_1}\right) = \frac{\partial}{\partial q_1}\left(\frac{\partial \tilde{S}}{\partial J_1}\right)\dot{q}_1 = \frac{\partial^2 \tilde{S}}{\partial J_1 \partial q_1}\,\dot{q}_1$$

and

$$2\pi = \int_t^{t+T} \frac{\mathrm{d}w_1}{\mathrm{d}t}\,\mathrm{d}t = \oint \frac{\partial^2 \tilde{S}}{\partial J_1 \partial q_1}\,\mathrm{d}q_1 = \frac{\partial}{\partial J_1}\oint \frac{\partial \tilde{S}}{\partial q_1}\,\mathrm{d}q_1 = \frac{\partial}{\partial J_1}\oint p_1\,\mathrm{d}q_1.$$

Hence

$$J_1 = \frac{1}{2\pi}\oint p_1\,\mathrm{d}q_1, \tag{2.6}$$

except for an arbitrary integration constant (of the integration in J_1). The quantity J_1 has the dimension of angular momentum or action and is an invariant of the motion (see Sect. 1.2.2). It is equal to the variation of the action when the solution performs a complete circulation or libration. Because of this property, it was called *modulus of periodicity of the action* [93] or *modulus of variation of the action* [11]. Since it gives the area delimited by the trajectory in the phase plane, it was also called *phase integral.* The adoption of these variables in the old Quantum Theory was first proposed by Sommerfeld.

[2] We adopted $\oint \mathrm{d}w_1 = 2\pi$. In many classical texts, $\oint \mathrm{d}w_1 = 1$.

The conjugate variables w_1, J_1 were called *angle* and *action* variables, a denomination that became standard after its adoption in Born's *Atommechanik* [12]. This denomination is the one currently used. The corresponding canonical equations are

$$\dot{w}_1 = \frac{\partial E}{\partial J_1} = \frac{2\pi}{T}, \qquad \dot{J}_1 = -\frac{\partial E}{\partial w_1} = 0. \tag{2.7}$$

Finally, let us note that, when the periodic motion is a libration, the quantity defined by (2.6) is singular when $J_1 = 0$. Indeed, the integral gives the area enclosed by the libration orbit and the singularity $J_1 = 0$ is a consequence of the fact that the direction of the motion in the phase space (q_1, p_1) cannot be reversed. Examples and consequences of this singularity in Celestial Mechanics will be extensively considered in Chap. 7

2.1.2 The Sign of the Action

We shall emphasize that the result of the operation defining the actions may be either positive or negative. To avoid any ambiguity, it is enough to write the definition of the action as

$$J_1 = \frac{1}{2\pi}\int_t^{t+T} p_1\dot{q}_1\,\mathrm{d}t. \tag{2.8}$$

For instance, in the simple pendulum solutions, J_1 is positive if $m > 0$ or negative if $m < 0$ (see Fig. B.1). We recall that w_1 is, by definition, always such that $\dot{w}_1 > 0$.

Exercise 2.1.1 (Angle–Action Variables of the Harmonic Oscillator).

1. Show that the angle–action variables of the harmonic oscillator defined by $U = \frac{k}{2}q_1^2 \;\; (k > 0)$ are

$$w_1 \;=\; \arcsin\sqrt{\frac{mk}{2E}}q_1 = \sqrt{k}(t + \alpha_1), \tag{2.9}$$

$$J_1 = \frac{E}{\sqrt{k}}. \tag{2.10}$$

   α_1 is a constant. *Hint*: $H = \dfrac{p_1^2}{2m} + \dfrac{km}{2}q_1^2$.

2. Show that

$$p_1 = \sqrt{2mE}\cos w_1. \tag{2.11}$$

2.2 Direct Construction of Angle–Action Variables

It is possible to rearrange the theory to directly obtain the angle–action variables. We may start from

$$J_1 = \frac{1}{2\pi} \oint p_1(q_1, E)\, \mathrm{d}q_1, \tag{2.12}$$

where $p_1(q_1, E)$ is obtained from the inversion of the energy integral $E = E(q_1, p_1)$. If the given Hamiltonian is quadratic in p_1, like in Sect. 1.9, this is an Abelian integral whose solution may benefit from some usual transformations and, when necessary, the use of the theory of residues[3]. The other basic equations are

$$\tilde{S} = \int p_1(q_1, E)\, \mathrm{d}q_1 \tag{2.13}$$

and

$$w_1 = \frac{\partial \tilde{S}}{\partial J_1} = \int \frac{\partial}{\partial J_1} p_1(q_1, E)\, \mathrm{d}q_1. \tag{2.14}$$

This step depends on the algebraic inversion of the solution of (2.12) to obtain $E = E(J_1)$. Another possibility is to take, instead of (2.14),

$$w_1 = \frac{\partial \tilde{S}}{\partial E} \left(\frac{\mathrm{d}J_1}{\mathrm{d}E} \right)^{-1} = \left(\frac{\mathrm{d}J_1}{\mathrm{d}E} \right)^{-1} \int \frac{\partial}{\partial E} p_1(q_1, E)\, \mathrm{d}q_1. \tag{2.15}$$

The replacement of $\mathrm{d}E/\mathrm{d}J_1$ by $(\mathrm{d}J_1/\mathrm{d}E)^{-1}$, which may be directly obtained from (2.12) without the need of any algebraic inversion, is always possible as long as $E(J_1)$ is a monotonic function.

However, these tasks are often made very difficult or even impossible to accomplish analytically because of the reasonably complex forms of $p_1(E, q_1)$.

There are ways of overcoming this situation. One of them, used in this book to obtain angle–action variables for the small oscillations of the pendulum (Sect. B.4) and of the Andoyer Hamiltonian (Sect. C.9), is founded on the fact that we are dealing with periodic solutions of the given Hamiltonian system, which may be represented by Fourier series. There are many different ways of calculating these series. In this book, we limit ourselves to the neighborhood of the equilibrium solutions. The solutions of the given system are represented by the series

$$q_1 = a_0 + \sum_{i=1}^{n} a_i \gamma^i,$$

where a_i are undetermined periodic functions of the angle w_1 and γ is a free parameter of the order of the amplitude of the oscillations. ($\gamma = 0$ corresponds to the stable equilibrium solution $q_1 = a_0$.) It is important to keep in mind that we need to construct the whole family of periodic solutions and that $\dot{w}_1$

[3] For some specific examples, see [93], Note 6.

is not the same for all solutions but is itself also a function of the parameter γ. It is assumed to be a power series in γ with undetermined coefficients:

$$\dot{w}_1 = \omega_0 + \sum_{i=1}^{n} o_i \gamma^i .$$

$p_1(w_1)$ is constructed using the equations of the motion or the energy integral. The angle–action variables are w_1 and

$$J_1 \stackrel{\text{def}}{=} \frac{1}{2\pi} \int_0^{2\pi} p_1 \frac{\mathrm{d}q_1}{\mathrm{d}w_1} \, \mathrm{d}w_1 .$$

The order n of the solution may be chosen according to the practical needs of the problem being solved and the means available for the calculation. Existing algebraic manipulators allow high orders to be considered. The practical steps of this construction may be seen in the cases presented in Sects. B.4 and C.9.

A different method is the numerical construction of the angle–action variables [50]. Let $H(q_1, p_1)$ be the Hamiltonian of an autonomous system and

$$\begin{aligned} q_1 &= q_1(q_0, p_0, t) \\ p_1 &= p_1(q_0, p_0, t) \end{aligned} \tag{2.16}$$

its solution for a given initial condition (q_0, p_0) and let $T(q_0, p_0)$ be the period of this solution.

The corresponding angle–action variables are

$$w_1 \stackrel{\text{def}}{=} \frac{2\pi}{T} t \tag{2.17}$$

$$J_1 \stackrel{\text{def}}{=} \frac{1}{2\pi} \int_0^T p_1 \frac{\mathrm{d}q_1}{\mathrm{d}t} \, \mathrm{d}t = -\frac{1}{2\pi} \int_0^T q_1 \frac{\mathrm{d}p_1}{\mathrm{d}t} \, \mathrm{d}t$$

and the inverses of these definitions are

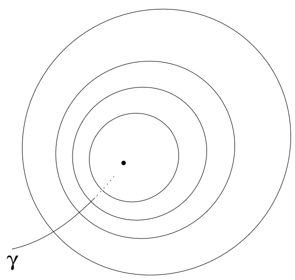

Fig. 2.2. Orbits transverse to a curve γ

$$\begin{aligned} q_0 &= q_0(w_1, J_1) \\ p_0 &= p_0(w_1, J_1). \end{aligned} \tag{2.18}$$

In this technique, all functions are constructed numerically. It is noteworthy that the last inversion may be more economically done when, beforehand, one has constructed the derivatives $\partial J_1/\partial q_0$ and $\partial J_1/\partial p_0$.

A last practical point to be noted is that we need just to numerically integrate from initial conditions lying on a given curve (γ) transverse to the orbits (and passing through the center of the orbits if we intend to include in the study also its immediate neighborhood) (Fig. 2.2). The extension of the solutions of (2.17) to the other points on each orbit is immediate.

The algorithms provided by Mayer's lemma (Sect. 1.10) allow the above construction to be extended to obtain a canonical transformation including other degrees of freedom. (see Sect. 2.4.4)

2.3 Actions in Multiperiodic Systems. Einstein's Theory

Let us consider a conservative Hamiltonian system with N degrees of freedom. It was shown in Sect. 1.2.2 that the action

$$J = \oint \sum_{i=1}^{N} p_i \, \mathrm{d}q_i \tag{2.19}$$

is an invariant of the motion (Helmholtz invariant).

If $S(q, \beta)$ is a solution of the Hamilton–Jacobi equation, then $p_i = \partial S/\partial q_i$,

$$\sum_{i=1}^{N} p_i \, \mathrm{d}q_i \; = \; \mathrm{d}S(q, \beta)$$

is an exact differential and the integral (2.19) has the same value for all closed curves that may be continuously deformed into one another. In particular, for all curves that may be reduced to one point by means of a continuous deformation, we have $J = 0$. When the solutions lie on a multiply connected manifold, there are closed curves that cannot be reduced to one point by continuous deformation (see [42]). This property was used by Einstein [26] to prove that, when the Hamiltonian is integrable, it is possible to construct N independent actions.

The multiperiodic solutions of a conservative integrable Hamiltonian corresponding to N constants β_i form a surface homeomorphic to $\mathbf{T}^N$. An N-dimensional torus is an N times connected surface and then we can find N different closed curves Γ_k that cannot be pairwise deformed into one another or reduced to one point and that may serve to uniquely define N independent actions

$$J_k = \oint_{\Gamma_k} \sum_{i=1}^{N} p_i \, \mathrm{d}q_i. \tag{2.20}$$

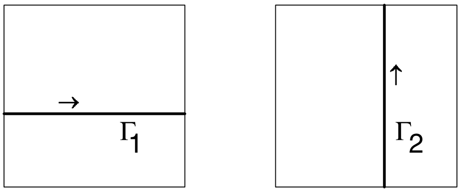

Fig. 2.3. Curves on tori $\mathbf{T}^2$ (the tori are obtained by joining the opposite sides of each square)

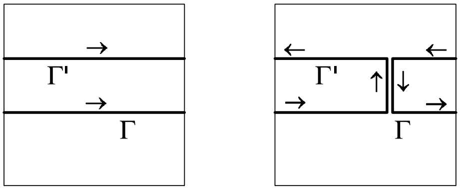

Fig. 2.4. Left: Integration paths Γ and Γ'. Right: Integration path obtained by introducing a cut between them and inverting the direction of Γ'

Let us consider the particular case $N = 2$. In $\mathbf{T}^2$, there are two types of closed curves that cannot be reduced to one point or transformed into one another by continuous deformation. They are shown in Fig. 2.3. All other closed curves on the surface of the torus can, by means of continuous deformations, be reduced to one point or transformed into one or more loops of the curves Γ_1 and Γ_2. To the closed curves Γ_1 and Γ_2, there correspond two independent actions J_1 and J_2.

In order that the definitions of J_k $(k = 1, 2)$ have a meaning, the values of J obtained from all closed curves that can be continuously transformed into one another may be the same. Let Γ and Γ' be two oriented closed curves that may be transformed into one another (Fig. 2.4, left). We may prove that the resulting actions J and J' are such that $J = J'$. To show this, we calculate $J - J'$. First, a cut joining Γ and Γ' is introduced. (The cut is shown in Fig. 2.4, right, as a pair of infinitesimally separated segments.) The resulting path is a curve drawn <u>on</u> the torus without encircling it and which may be reduced to one point; the integral over this path is then equal to zero. If we note that the integrals over the cut are opposite and cancel each other and that the integral over Γ' is done in a direction contrary to that used to define J', it follows that $J - J' = 0$. □

The actions constructed with Einstein's theory may be completed by angles defined by $w_k = \partial\tilde{S}/\partial J_k$ where $\tilde{S}(q, J) = S(q, \beta(J))$ [4].

[4] For a modern and rigorous definition of the angle–action variables of an integrable system, see [4], Sect. 50.

2.4 Separable Multiperiodic Systems

There are no general methods for the solution of the Hamilton–Jacobi equation in the case of more than one degree of freedom. The use of general theories, such as Cauchy characteristics, just recovers the given Hamiltonian system. However, under certain special conditions, for some important problems such as Keplerian motion (or the Rutherford–Bohr atom), a complete solution of the Hamilton–Jacobi equation may be obtained. In these very particular cases, one partial differential equation in N variables can be replaced by N separate ordinary differential equations, one for each variable, and the complete integration of the equation is achieved.

Generally speaking, a problem is said to be *separable* when the corresponding Hamilton–Jacobi equation has a *complete* integral $S(q,\beta)$ which may be separated as

$$S(q,\beta) = S_1(q_1,\beta) + S_2(q_2,\beta) + \cdots + S_N(q_N,\beta), \tag{2.21}$$

where each term $S_k = S_k(q_k,\beta)$ is independent of the q_j $(j \neq k)$.

In this case, the equations of the motion are given by

$$\begin{aligned} t + \alpha_1 = q_1^* &= \frac{\partial S}{\partial \beta_1} = \sum_{k=1}^{N} \frac{\partial S_k(q_k,\beta)}{\partial \beta_1} = \sum_{k=1}^{N} F_{1k}(q_k) \\ \alpha_\varrho = q_\varrho^* &= \frac{\partial S}{\partial \beta_\varrho} = \sum_{k=1}^{N} \frac{\partial S_k(q_k,\beta)}{\partial \beta_\varrho} = \sum_{k=1}^{N} F_{\varrho k}(q_k), \end{aligned} \tag{2.22}$$

$(\varrho = 2,3,\cdots,N)$, where we have introduced the functions

$$F_{jk}(q_k) \stackrel{\text{def}}{=} \frac{\partial S_k(q_k,\beta)}{\partial \beta_j}. \tag{2.23}$$

The other equations, completing the transformation, are

$$p_k = \frac{\partial S}{\partial q_k} = \frac{\partial S_k(q_k,\beta)}{\partial q_k} \qquad (k = 1,\cdots,N). \tag{2.24}$$

Equations (2.24) show that the trajectories projected in the phase subspaces q_k, p_k are mutually independent. The law of motion along the projected trajectories may be obtained by solving the equations of the motion, (2.22), with respect to the q_k. As in Sect. 2.1, these projected periodic motions may be either circulations or librations.

2.4.1 Uniformized Angles. Charlier's Theory

The generalization of the angle variables of Sect. 2.1.1 to N degrees of freedom may be done following the same principle as there. We define a *partial cyclic* *variation* in which the corresponding variable q_i performs a complete circulation or libration while the other variables $q_k (k \neq i)$ are kept unaltered. We then define a set of N angle variables $w_i \in \mathbf{S}$ such that in a partial cyclic variation of q_i, the corresponding w_i increases 2π while the other angles $w_k (k \neq i)$ are not affected. Such angle variables are said to be uniformized.

In a partial cyclic variation of q_i, the functions F_{ji} change while all functions $F_{jk} (k \neq i)$ remain unchanged. Let γ_{ji} be the increment of the functions $F_{ji}(q_i, \beta)$ in a partial cyclic variation of q_i:

$$\gamma_{ji} = F_{ji}(q_i + \oint \mathrm{d}q_i) - F_{ji}(q_i), \tag{2.25}$$

where $\oint \mathrm{d}q_i$ denotes the partial cyclic variation of q_i. It is important to keep in mind that the resulting *repetition numbers* γ_{ji} are not independent of the initial values of the q_i (as T is not independent of the initial q_1 in the case of one degree of freedom).

Proposition 2.4.1 (Charlier [20]**).** *If* $\det(\gamma_{ji}) \neq 0$*, the variables* w_i *defined by the equations*

$$q_j^* = \frac{1}{2\pi} \sum_{\ell=1}^{N} \gamma_{j\ell} w_\ell \tag{2.26}$$

are uniformized angle variables.

Proof. Let us introduce the inverse matrix of (γ_{ji}) and denote its elements by γ_{ji}^{-1}. If $\det(\gamma_{ji}) \neq 0$, (2.26) may be inverted, giving

$$w_k = 2\pi \sum_{j=1}^{N} \gamma_{kj}^{-1} q_j^*. \tag{2.27}$$

In a partial cyclic variation of q_i, the variation of w_k is

$$\delta_i w_k = 2\pi \sum_{j=1}^{N} \gamma_{kj}^{-1} \delta_i q_j^* = 2\pi \delta_{ki} \tag{2.28}$$

(by construction, $\delta_i q_j^* = \gamma_{ji}$). Therefore, in a partial cyclic variation of q_i, w_i increases of 2π while the others w_k $(k \neq i)$ remain unchanged. □

2.4.2 The Actions

The next step is to find the action variables J_k canonically conjugate to the angle variables w_k. To do this, we introduce the Jacobian generating function of the canonical transformation $\tilde{\phi} : (q, p) \Rightarrow (w, J)$, namely $\tilde{S}(q, J)$. We then have

$$w_k = \frac{\partial \tilde{S}}{\partial J_k} \tag{2.29}$$

and

$$\mathrm{d}w_k = \sum_{j=1}^{N} \frac{\partial^2 \tilde{S}}{\partial q_j \partial J_k} \, \mathrm{d}q_j + \sum_{j=1}^{N} \frac{\partial^2 \tilde{S}}{\partial J_j \partial J_k} \, \mathrm{d}J_j . \tag{2.30}$$

In a partial cyclic variation of q_k, w_k increases 2π. Besides, along the given path, $\mathrm{d}q_j = 0$ for $j \neq k$ and $\mathrm{d}J_i = 0$. The above equation then reduces to

$$2\pi = \oint \frac{\partial^2 \tilde{S}}{\partial q_k \partial J_k} \, \mathrm{d}q_k .$$

A trivial calculation, similar to that of Sect. 2.1.1, gives

$$J_k = \frac{1}{2\pi} \oint p_k \, \mathrm{d}q_k \tag{2.31}$$

for every $k \in \{1, \cdots, N\}$.

2.4.3 Algorithms for Construction of the Angles

In practice, we use some straightforward approaches to obtain the angles. The separation of the Hamilton–Jacobi equation leads us to obtain $p_j = p_j(q_j, \beta)$ and the solution

$$S(q, \beta) = \sum_{j=1}^{N} S_j(q_j, \beta) = \sum_{j=1}^{N} \int p_j(q_j, \beta) \, \mathrm{d}q_j . \tag{2.32}$$

We may also solve (2.31) with $p_k = p_k(q_k, \beta)$ to obtain the actions J_k as functions of the constants β_i.

The Jacobian generating function $\tilde{S}(q, J)$ may be obtained, now, from $\tilde{S}(q, J) = S(q, \beta(J))$ and (2.29) gives the angles:

$$w_k = \frac{\partial \tilde{S}}{\partial J_k} = \sum_{i=1}^{N} \frac{\partial S}{\partial \beta_i} \frac{\partial \beta_i}{\partial J_k} = \sum_{i=1}^{N} \frac{\partial \beta_i}{\partial J_k} \sum_{j=1}^{N} \int \frac{\partial p_j}{\partial \beta_i} \, \mathrm{d}q_j . \tag{2.33}$$

These equations are akin to the equations

$$w_k = \frac{\partial}{\partial J_k} \int \sum_{j=1}^{N} \widehat{p}_j(q, J) \, \mathrm{d}q_j , \tag{2.34}$$

which would result if the canonical transformation of Mayer's lemma (Sect. 1.10) were used in this case. The conditions under which that transformation was established (involution of the functions $J_i(q,p)$ and possibility of inversion to obtain the functions $\widehat{p}_i(q, J)$) are satisfied and it can be used. Equation (2.34) transforms itself into (2.33) in the separable case in which every term p_j depends only on the corresponding q_j.

2.4.4 Angle–Action Variables of $H(q_1, p_1, p_2, \cdots, p_N)$

Let us consider the case of a Hamiltonian having the form $H(q_1, p_1, p_2, \cdots, p_N)$, where the coordinates $q_2, \cdots, q_N$ are ignorable and the momenta $p_2, \cdots, p_N$ are constants. Because of frequent applications, it is worth having the algorithm of the previous section explicitly given in this case.

The Hamiltonian H is reducible to one degree of freedom and the angle–action variables of the reduced Hamiltonian may be obtained with one of the algorithms discussed in Sect. 2.2. The results of the previous section allow the one-degree-of-freedom transformation $(q_1, p_1) \to (w_1, J_1)$, thus obtained, to be embedded into a more general transformation $(q, p) \to (w, J)$ that considers also the remaining degrees of freedom of the given Hamiltonian. To do this, we consider as given the N functions

$$\begin{aligned} J_1 &= f_1(q_1, p) \\ J_\varrho &= p_\varrho \equiv f_\varrho(q, p) \qquad (\varrho = 2, \cdots, N). \end{aligned} \tag{2.35}$$

These functions are pairwise in involution and may be solved for the momenta. Because of the particular form of the functions f_ϱ, the inversion is trivial, giving $p = \widehat{p}(q_1, J)$. The resulting generating function of the transformation (2.35) is simply

$$\sum_{j=1}^{N} \int \widehat{p}_j(q_1, J)\, \mathrm{d}q_j = \int \widehat{p}_1(q_1, J)\, \mathrm{d}q_1 + \sum_{\varrho=2}^{N} J_\varrho q_\varrho. \tag{2.36}$$

We then have

$$\begin{aligned} w_1 &= \Xi_1 \\ w_\varrho &= q_\varrho + \Xi_\rho \qquad (\varrho \geq 2), \end{aligned} \tag{2.37}$$

where

$$\Xi_k = \frac{\partial}{\partial J_k} \int \widehat{p}_1(q_1, J)\, \mathrm{d}q_1 \qquad (k \geq 1). \tag{2.38}$$

We note that (2.36) comes from the integration of an exact differential form in $\mathrm{d}q_j$ and that we may add to the generating function any arbitrary function of J.

The one-degree-of-freedom canonical transformation $(q_1, p_1) \to (w_1, J_1)$ is often given in the inverted form

$$\begin{aligned} q_1 &= Q_1(w_1, J) \\ p_1 &= P_1(w_1, J). \end{aligned} \tag{2.39}$$

In this case, (2.38) may be written

$$\Xi_k = \int \left[\frac{\partial \widehat{p}_1(q_1, J)}{\partial J_k} \right]_{q_1 = Q_1} \frac{\partial Q_1}{\partial w_1}\, \mathrm{d}w_1. \tag{2.40}$$

It is worth emphasizing that the substitution $q_1 = Q_1(w_1, J)$ may be done after the differentiation and that it is no longer possible to permute the differentiation with respect to J_k and the integration.

If the differentials of $\widehat{p}_1(q_1, J)$ and $P_1(w_1, J)$ are compared, we obtain

$$\begin{aligned}\frac{\partial P_1}{\partial J_k} &= \frac{\partial \widehat{p}_1}{\partial J_k} + \frac{\partial \widehat{p}_1}{\partial q_1}\frac{\partial Q_1}{\partial J_k}\\ \frac{\partial P_1}{\partial w_1} &= \frac{\partial \widehat{p}_1}{\partial q_1}\frac{\partial Q_1}{\partial w_1},\end{aligned}$$

which, substituted in (2.40), give the equivalent result

$$\Xi_\varrho = \int_0^{w_1} \left(\frac{\partial Q_1}{\partial w_1}\frac{\partial P_1}{\partial J_\varrho} - \frac{\partial Q_1}{\partial J_\varrho}\frac{\partial P_1}{\partial w_1} \right) \mathrm{d}w_1, \tag{2.41}$$

obtained by Henrard and Lemaitre [50]. We also have the trivial relation $\Xi_1 = w_1$, since the integrand in this case is the one-dimensional Lagrange bracket $[w_1, J_1]$ which is equal to 1 because the given transformation $(q_1, p_1) \rightarrow (w_1, J_1)$ is canonical.

In this section, we have considered a Hamiltonian independent of the coordinates q_ϱ $(\rho = 2, \cdots, N)$. The algorithms derived from Mayer's lemma are valid in more general circumstances, but the results are not angle–action variables of the given Hamiltonian when H depends on the q_ϱ. If, for instance, a general H may be decomposed into two parts: $H = H_a(q_1, p_1) + H_b(q_\varrho, p_\varrho)$, and the formula is used to extend the angle–action variables of H_a, the resulting variables w, J are not angle–action variables of H. It is easy to see that the calculations to obtain the w_ϱ, J_ϱ are the same for any H_b and, thus, we cannot expect that it eliminates the angles from H_b. This comment is somewhat obvious, but is useful to avoid pitfalls.

Exercise 2.4.1. By construction, the functions $q_1(w_1, J)$ and $p_1(w_1, J)$ are 2π-periodic in the angle variable w_1. Under which conditions may we guarantee that the functions Ξ_ϱ are also 2π-periodic in w_1?

Exercise 2.4.2. Find the angle variable conjugate to $J_1 = \frac{1}{2}(q_1^2 + p_1^2)$. Check the result with $\{w_1, J_1\} = 1$.

Exercise 2.4.3. Find a set of angle–action variables for the Hamiltonian

$$H = \frac{1}{2}(p_1^2 + p_2^2) + \frac{1}{2}\lambda^2 q_1^2,$$

where $\lambda = \lambda(p_2)$. Check the results with $\{w_1, J_j\} = \delta_{1j}$. *Hint:* See Exercise 2.1.1.

2.4.5 Historical Postscript

The given definitions of angle–action variables follow those found in several classical texts on Celestial Mechanics and on the old Quantum Theory[5]. Angles and actions appeared separately. The *angles* were first introduced by Charlier [20] as a complement to what was then called *Staude–Stäckel theory.* They came as a result of an application of Weierstrass' theory of multiperiodic functions to the solutions of the Hamilton–Jacobi equation of a separable system. The *actions* evolved from the quantity defined in Malpertuis' least action principle (see Sect. 1.2.1), quantized in the theories of Planck and Bohr, to their definition for separable multiperiodic Hamiltonians given by Sommerfeld [92] and Epstein [28]. The introduction of the angles as variables canonically conjugate to the actions through a Jacobian generating function $\tilde{S}(q, J)$ is due to Kramers (cf. [10], Note 24). The definition of the actions of an integrable Hamiltonian system without recourse to the separability hypothesis is due to Einstein [26]. (An alternative construction was presented, at the same time, by Burgers; cf. [12].) The introduction of invariant tori in modern theory is due to Arnold [3]. It is worth mentioning that Einstein's construction of invariant tori is very different from that adopted in the modern theory of Hamiltonian systems. Einstein considered one example (central motions in a plane) and used the fact that, for given β_i, the phase space may be seen as a vector field on a Riemann surface formed by two annular sheets joined by their edges (which is homeomorphic to $\mathbf{T}^2$).

The angle–action variables ℓ, g, h, L, G, H obtained in Sect. 2.7.2 as an application of the Schwarzschild transformation to the angle–action variables of Keplerian motion, were actually discovered by Delaunay a long time before and fully employed in his (canonical) *Théorie de la Lune* [22].

It is worth emphasizing that the introduction of angle–action variables in Delaunay's work, as well as in the work of Sommerfeld and his contemporaries, resulted from specific needs for the actual solution of problems in Astronomy and Physics. The hiatus between the results of old Quantum Theory (before 1920) and modern theories (ca. 1960) has an explanation. The construction of action variables was the central point of the Bohr–Sommerfeld quantum condition. With the foundation of Quantum Mechanics, in the early 1920s, the actions lost their position in center stage. KAM theory has again made angle and action variables central concepts in Physics and Dynamics.

2.5 Simple Separable Systems

We only know some sets of sufficient conditions for separability. Some simple cases are the dynamical systems whose Hamiltonians have special structures, such as

[5] Specifically, we mention (in chronological order) Charlier [20], Schwarzschild [84], Einstein [26], Sommerfeld [93], Born [12], and Boll and Salomon [11].

$$H = G[f_1(q_1, p_1), \cdots, f_N(q_N, p_N)] \tag{2.42}$$

and

$$H = f_1 \{q_1, p_1, f_2 [q_2, p_2, f_3 (\cdots, f_N(q_N, p_N))]\} . \tag{2.43}$$

In the first case, the variables in the expression for the function H are separated, i.e., only one pair of conjugate variables q_i, p_i enters into each function f_i. The Hamilton–Jacobi equation corresponding to this case is

$$G\left[f_1\left(q_1, \frac{\partial S}{\partial q_1}\right), \cdots, f_N\left(q_N, \frac{\partial S}{\partial q_N}\right)\right] = E. \tag{2.44}$$

After the introduction of $S = \sum_{i=1}^{N} S_i(q_i)$, this equation is separated into N equations

$$f_i\left(q_i, \frac{\mathrm{d}S_i}{\mathrm{d}q_i}\right) = \beta_i, \tag{2.45}$$

the integration constants β_i being such that

$$E = G(\beta_1, \cdots, \beta_N). \tag{2.46}$$

In the second case, the variables appear in a hierarchical disposition and the corresponding Hamilton–Jacobi equation,

$$f_1\left\{q_1, \frac{\partial S}{\partial q_1}, f_2\left[q_2, \frac{\partial S}{\partial q_2}, f_3\left(\cdots, f_N\left(q_N, \frac{\partial S}{\partial q_N}\right)\right)\right]\right\} = E, \tag{2.47}$$

after the introduction of $S = \sum_{i=1}^{N} S_i(q_i)$, is separated into N equations

$$\begin{aligned} f_N\left(q_N, \frac{\mathrm{d}S_N}{\mathrm{d}\,q_N}\right) &= \beta_N \\ f_i\left(q_i, \frac{\mathrm{d}S_i}{\mathrm{d}\,q_i}, \beta_{i+1}\right) &= \beta_i \qquad (i = 1, \cdots, N-1), \end{aligned} \tag{2.48}$$

with $\beta_1 = E$. If we assume that $\partial f_i/\partial p_i \neq 0$ for all $i = 1, \cdots, N$, these equations can be solved to give

$$\begin{aligned} \frac{\mathrm{d}S_i}{\mathrm{d}\,q_i} &= G_i(q_i, \beta_{i+1}, \beta_i) \qquad (i = 1, \cdots, N-1) \\ \frac{\mathrm{d}S_N}{\mathrm{d}\,q_N} &= G_N(q_N, \beta_N). \end{aligned} \tag{2.49}$$

2.5.1 Example: Central Motions

The classical example of a separable system of this kind is the motion of a particle in a central force field. In spherical coordinates, the total energy of the particle is

$$H = T + mU(r) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta\, \dot{\phi}^2) + mU(r) \tag{2.50}$$

or, introducing the generalized momenta

$$\begin{aligned} p_r &= \frac{\partial T}{\partial \dot{r}} = m\dot{r} \\ p_\theta &= \frac{\partial T}{\partial \dot{\theta}} = mr^2\dot{\theta} \\ p_\phi &= \frac{\partial T}{\partial \dot{\phi}} = mr^2 \sin^2\theta\, \dot{\phi}, \end{aligned} \tag{2.51}$$

we obtain

$$H = \frac{1}{2m}\left[p_r^2 + \frac{1}{r^2}\left(p_\theta^2 + \frac{p_\phi^2}{\sin^2\theta}\right)\right] + mU(r). \tag{2.52}$$

The above Hamiltonian has the special structure of (2.43) with

$$\begin{aligned} f_1 &= \frac{1}{2m}\left(p_r^2 + \frac{f_2}{r^2}\right) + mU(r) = E \\ f_2 &= p_\theta^2 + \frac{f_3}{\sin^2\theta} = \beta_2 \\ f_3 &= p_\phi^2 = \beta_3, \end{aligned} \tag{2.53}$$

and an application of (2.49) gives

$$\begin{aligned} p_r &= \frac{\mathrm{d}S_1}{\mathrm{d}r} = \sqrt{2mE - \frac{\beta_2}{r^2} - 2m^2U(r)} \\ p_\theta &= \frac{\mathrm{d}S_2}{\mathrm{d}\theta} = \sqrt{\beta_2 - \frac{\beta_3}{\sin^2\theta}} \\ p_\phi &= \frac{\mathrm{d}S_3}{\mathrm{d}\phi} = \sqrt{\beta_3}. \end{aligned} \tag{2.54}$$

2.5.2 Angle–Action Variables of Central Motions

Let us calculate the angle–action variables of the central motions, starting with J_ϕ. A short chain of elementary calculations gives

$$J_\phi = \frac{1}{2\pi}\oint \sqrt{\beta_3}\,\mathrm{d}\phi = \frac{1}{2\pi}\int_0^{2\pi} \sqrt{\beta_3}\,\mathrm{d}\phi = \sqrt{\beta_3} = p_\phi. \tag{2.55}$$

The integration of the next one is also elementary, but not as immediate:

$$J_\theta = \frac{1}{2\pi}\oint p_\theta\,\mathrm{d}\theta = \frac{1}{2\pi}\oint \sqrt{\beta_2 - \frac{\beta_3}{\sin^2\theta}}\,\mathrm{d}\theta.$$

It may, however, be avoided by noting that when the plane of the motion is chosen as the fundamental reference plane, we have, in analogy with the previous result,

$$J_\psi = p_\psi = mr^2\dot{\psi}, \tag{2.56}$$

where ψ denotes the longitude reckoned on the plane of motion. Taking into account (2.53) and $\dot{\psi}^2 = \dot{\theta}^2 + \sin^2\theta\,\dot{\phi}^2$, we obtain

$$J_\psi = \sqrt{\beta_2}. \tag{2.57}$$

Comparing, now, the kinetic energies in the two reference systems:

$$p_r\dot{r} + p_\theta\dot{\theta} + p_\phi\dot{\phi} = p_r\dot{r} + p_\psi\dot{\psi}, \tag{2.58}$$

it follows that

$$J_\theta = \frac{1}{2\pi}\oint p_\theta\,\mathrm{d}\theta = \frac{1}{2\pi}\oint (p_\psi\,\mathrm{d}\psi - p_\phi\,\mathrm{d}\phi) = J_\psi - J_\phi \tag{2.59}$$

and then[6]

$$J_\theta = \sqrt{\beta_2} - \sqrt{\beta_3}. \tag{2.60}$$

The radial action

$$J_r = \frac{1}{2\pi}\oint \sqrt{2mE - 2m^2U(r) - \frac{\beta_2}{r^2}}\,\mathrm{d}r \tag{2.61}$$

cannot be calculated now since the potential $U(r)$ has not yet been given (see next section).

To obtain the angle variables w_k, we follow the procedure given in Sect. 2.4.1. We first write

$$S(q,\beta) = \int p_r\,\mathrm{d}r + \int p_\theta\,\mathrm{d}\theta + \int p_\phi\,\mathrm{d}\phi \tag{2.62}$$

and then introduce

$$\begin{aligned} \beta_1 &= E = E(J_r, J_\theta, J_\phi) \\ \beta_2 &= (J_\theta + J_\phi)^2 \\ \beta_3 &= J_\phi^2, \end{aligned} \tag{2.63}$$

where we have to keep in mind that the function $E = E(J_r, J_\theta, J_\phi)$ may be known only when the potential $U(r)$, of the central force, is given. We then have

[6] $J_\psi \geq J_\phi > 0$, $\beta_2 \geq \beta_3 > 0$ and $J_\theta \geq 0$.

$$\begin{aligned} w_r &= \frac{\partial \tilde{S}}{\partial J_r} = \frac{\partial S}{\partial E}\frac{\partial E}{\partial J_r} \\ w_\theta &= \frac{\partial \tilde{S}}{\partial J_\theta} = \frac{\partial S}{\partial E}\frac{\partial E}{\partial J_\theta} + \frac{\partial S}{\partial \beta_2}\,2(J_\theta + J_\phi) \\ w_\phi &= \frac{\partial \tilde{S}}{\partial J_\phi} = \frac{\partial S}{\partial E}\frac{\partial E}{\partial J_\phi} + \frac{\partial S}{\partial \beta_2}\,2(J_\theta + J_\phi) + \frac{\partial S}{\partial \beta_3}\,2J_\phi, \end{aligned} \tag{2.64}$$

where

$$\begin{aligned} \frac{\partial S}{\partial E} &= \int \frac{\partial p_r}{\partial E}\,\mathrm{d}r \\ \frac{\partial S}{\partial \beta_2} &= \int \frac{\partial p_r}{\partial \beta_2}\,\mathrm{d}r + \int \frac{\partial p_\theta}{\partial \beta_2}\,\mathrm{d}\theta \\ \frac{\partial S}{\partial \beta_3} &= \int \frac{\partial p_\theta}{\partial \beta_3}\,\mathrm{d}\theta + \int \frac{\partial p_\phi}{\partial \beta_3}\,\mathrm{d}\phi. \end{aligned} \tag{2.65}$$

As for the actions, some of these integral cannot be calculated, because $U(r)$ has not yet been given. Those that may be calculated are

$$\begin{aligned} \int \frac{\partial p_\theta}{\partial \beta_2}\,\mathrm{d}\theta &= \int \frac{\mathrm{d}\theta}{2p_\theta} \\ \int \frac{\partial p_\theta}{\partial \beta_3}\,\mathrm{d}\theta &= -\int \frac{\mathrm{d}\theta}{2p_\theta \sin^2\theta} \end{aligned} \tag{2.66}$$

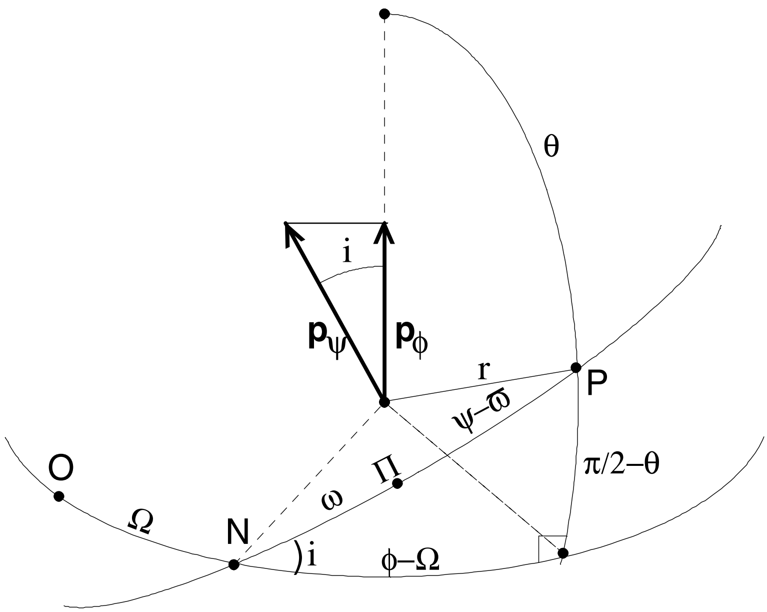

Fig. 2.5. Geometry of central motions ($\varpi = \Omega + \omega$)

$$\int \frac{\partial p_\phi}{\partial \beta_3}\,\mathrm{d}\phi \;=\; \int \frac{\mathrm{d}\phi}{2p_\phi}.$$

The third of these integrals is trivial since $p_\phi = J_\phi = \text{const}$. Omitting the integration constant,

$$\int \frac{\partial p_\phi}{\partial \beta_3}\,\mathrm{d}\phi = \frac{\phi}{2J_\phi}.$$

The two other integrals are also easy to calculate using an immediate relation between the angles and momenta ($\dot{\theta}/p_\theta = \dot{\psi}/p_\psi = \dot{\phi}\sin^2\theta/p_\phi$) and recalling that $p_\phi = J_\phi$ and $p_\psi = J_\psi = J_\phi + J_\theta$ are constants. We thus obtain

$$\int \frac{\partial p_\theta}{\partial \beta_2}\,\mathrm{d}\theta = \int \frac{\mathrm{d}\theta}{2p_\theta} = \int \frac{\mathrm{d}\psi}{2p_\psi} = \frac{\psi}{2(J_\phi + J_\theta)}$$

and

$$\int \frac{\partial p_\theta}{\partial \beta_3}\,\mathrm{d}\theta = -\int \frac{\mathrm{d}\theta}{2p_\theta \sin^2\theta} = -\int \frac{\mathrm{d}\phi}{2p_\phi} = -\frac{\phi - \Omega}{2J_\phi}.$$

In general, we have omitted integration constants, since this arbitrariness is intrinsic to the definition of the angles w_i. However, in the last equation, to shift the x-axis to the ascending node (N) of the orbit (see Fig. 2.5), we have introduced the integration constant $\Omega/2J_\phi$.

We may summarize the results by writing

$$\begin{aligned} w_r &= \frac{\partial E}{\partial J_r}\int \frac{\partial p_r}{\partial E}\,\mathrm{d}r \\ w_\theta &= \frac{\partial E}{\partial J_\theta}\int \frac{\partial p_r}{\partial E}\,\mathrm{d}r + 2(J_\theta + J_\phi)\int \frac{\partial p_r}{\partial \beta_2}\,\mathrm{d}r + \psi \\ w_\phi &= \frac{\partial E}{\partial J_\phi}\int \frac{\partial p_r}{\partial E}\,\mathrm{d}r + 2(J_\theta + J_\phi)\int \frac{\partial p_r}{\partial \beta_2}\,\mathrm{d}r + \psi + \Omega. \end{aligned} \tag{2.67}$$

2.6 Kepler Motion

In the case of the heliocentric motion of a planet, we have

$$U(r) = -\frac{\mu}{r},$$

where $\mu = G(M+m)$; G is the universal gravitation constant and M is the mass of the Sun. Now, we can consider the several integrals left uncalculated in the last section. The first one is the radial action J_r (see 2.61). We have

$$J_r = \frac{1}{2\pi}\oint \frac{1}{r}\sqrt{2mEr^2 + 2\mu m^2 r - \beta_2}\;\mathrm{d}r. \tag{2.68}$$

The radicand has the real roots

$$r_{1,2} = \frac{-\mu m}{2E}\left(1 \pm \sqrt{1 + \frac{2E\beta_2}{\mu^2 m^3}}\right). \tag{2.69}$$

One may note that, if $E > 0$, the two roots are real, but one is negative. In this case, the motion is only possible for r larger than the positive root and has no upper bound. For $-\mu^2 m^3/2\beta_2 < E < 0$, the two roots are real and positive, say $r_1 < r_2$; the motion is periodic and is a libration between the two roots. In this case, we may calculate the action J_r. The integral of (2.68) may be done along a path in a two-sheet Riemann surface enclosing the two branch points r_1, r_2. It has been thoroughly studied by Sommerfeld (see [93], Note 6). The sophisticated procedure idealized by Sommerfeld has, since then, been reproduced in many treatises on Mechanics. However, there is a simpler way of doing it. We introduce the *mean distance* to the force center

$$a \stackrel{\text{def}}{=} \frac{r_1 + r_2}{2} = -\frac{\mu m}{2E}, \tag{2.70}$$

the *eccentricity*

$$e \stackrel{\text{def}}{=} \frac{r_2 - r_1}{2a} = \sqrt{1 + \frac{2E\beta_2}{\mu^2 m^3}} \tag{2.71}$$

and the angle u (*eccentric anomaly*) defined through

$$r = a(1 - e\cos u). \tag{2.72}$$

A lengthy but elementary calculation gives

$$p_r = \sqrt{-2mE}\,\frac{e\sin u}{1 - e\cos u}$$

and the given integral becomes

$$J_r = \frac{1}{2\pi}\, a e^2 \sqrt{-2mE} \int_0^{2\pi} \frac{\sin^2 u \,\mathrm{d}u}{1 - e\cos u}. \tag{2.73}$$

The integral to be solved is trivial. We may just introduce $z = \mathrm{e}^{\mathrm{i}u}$ and perform the integration along the circle $|z| = 1$ in the complex plane, with recourse to the theory of residues. We obtain[7]

$$\int_0^{2\pi} \frac{\sin^2 u \,\mathrm{d}u}{1 - e\cos u} = \frac{2\pi}{e^2}\left(1 - \sqrt{1 - e^2}\right). \tag{2.74}$$

After some elementary calculations, we obtain

$$J_r = \sqrt{-2mE}\, a\left(1 - \sqrt{1 - e^2}\right) = \mu m \sqrt{\frac{m}{-2E}} - \sqrt{\beta_2} \tag{2.75}$$

[7] This integral is also found in tables, e.g. [25], [41].

and the inversion of this equation gives

$$E = -\frac{\mu^2 m^3}{2(J_r + J_\theta + J_\phi)^2} \tag{2.76}$$

(since $\sqrt{\beta_2} = J_\psi = J_\theta + J_\phi$).

We may now proceed with the remaining integrals. They are

$$\begin{aligned}\int \frac{\partial p_r}{\partial E}\,\mathrm{d}r &= \int \frac{m\,\mathrm{d}r}{p_r},\\ \int \frac{\partial p_r}{\partial \beta_2}\,\mathrm{d}r &= -\int \frac{\mathrm{d}r}{2p_r r^2}.\end{aligned}$$

Introducing the eccentric anomaly u, these integrals are changed into elementary ones. The first one is

$$\int \frac{m\,\mathrm{d}r}{p_r} = \frac{a^{3/2}}{\sqrt{\mu}}\int (1 - e\cos u)\,\mathrm{d}u = \frac{a^{3/2}}{\sqrt{\mu}}(u - e\sin u) \tag{2.77}$$

which introduces the *mean anomaly*

$$\ell \stackrel{\text{def}}{=} u - e\sin u.$$

The second one is

$$\begin{aligned}-\int \frac{\mathrm{d}r}{2p_r r^2} &= -\frac{1}{2m\sqrt{\mu a}}\int \frac{\mathrm{d}u}{1 - e\cos u}\\ &= -\frac{1}{m\sqrt{\mu a(1-e^2)}}\arctan\sqrt{\frac{1+e}{1-e}}\tan\frac{u}{2}\end{aligned} \tag{2.78}$$

which introduces the *true anomaly*

$$v \stackrel{\text{def}}{=} 2\arctan\sqrt{\frac{1+e}{1-e}}\tan\frac{u}{2}. \tag{2.79}$$

The two remaining integrals are, then,

$$\begin{aligned}\int \frac{\partial p_r}{\partial E}\,\mathrm{d}r &= \frac{a^{3/2}\ell}{\sqrt{\mu}}\\ \int \frac{\partial p_r}{\partial \beta_2}\,\mathrm{d}r &= -\frac{v}{2m\sqrt{\mu a(1-e^2)}}.\end{aligned}$$

Substituting these integrals into (2.67), and noting that

$$\frac{\partial E}{\partial J_r} = \frac{\partial E}{\partial J_\theta} = \frac{\partial E}{\partial J_\phi} = \frac{\mu^2 m^3}{(J_r + J_\theta + J_\phi)^3}$$

$$J_r + J_\theta + J_\phi = m\sqrt{\mu a} \tag{2.80}$$

and

$$J_\theta + J_\phi = \sqrt{\beta_2} = m\sqrt{\mu a(1 - e^2)}$$

it follows that

$$\begin{aligned} w_r &= \ell \\ w_\theta &= \ell - v + \psi \qquad = \ell + \omega \\ w_\phi &= \ell - v + \psi + \Omega = \ell + \omega + \Omega, \end{aligned} \tag{2.81}$$

where we have introduced the so-called argument of pericenter $\omega = \psi - v$, giving the distance of the pericenter (Π) to the ascending node (N) (see Fig. 2.5).

To complete the definition of the angle–action variables of the Kepler motion, we write[8]

$$J_\phi = m\sqrt{\mu a(1 - e^2)}\cos i. \tag{2.82}$$

2.7 Degeneracy

In the example studied in the previous section, the three frequencies

$$\nu_k = \frac{\partial E}{\partial J_k} \tag{2.83}$$

of the system are equal. We follow Schwarzschild and call this case *degenerate*. In general, degeneracy is said to occur when there exists a commensurability relation

$$(h \mid \nu) = \sum_{k=1}^{N} h_k \nu_k = 0 \qquad\qquad h \in \mathbf{Z}^N \tag{2.84}$$

amongst the frequencies of the system. Degeneracy may be essential or accidental. A degeneracy is said to be *essential* when it does not depend on the initial conditions. We shall stress that this does not mean that the frequencies themselves are independent of the initial conditions. The Keplerian motion is a good example: the frequencies $\nu_r, \nu_\theta, \nu_\phi$ (defined by the derivatives of E with respect to J_r, J_θ, J_ϕ) depend on the initial conditions but they are always equal, regardless of the initial conditions.

Otherwise, a degeneracy is called *accidental* when it only occurs for some particular values of the initial conditions. One example is the motion of an asteroid in an orbit whose period is commensurable with Jupiter's. In this case, the commensurability relation ceases to exist if the asteroid orbit is moved inward (or outward). The main consequence of an accidental degeneracy is

[8] The inclination is introduced by the fact that p_ψ is the angular momentum of the motion and p_ϕ is the angular momentum of the motion projected on the reference plane: $p_\phi = p_\psi \cos i$.

the appearance of small divisors, which impair the performance of perturbation theories. Motions affected by accidental commensurabilities are called *resonant* and are the subject of several of the next chapters.

A separable multiperiodic system may be such that multiple commensurability relations exist. Degeneracy affects the degree of periodicity of the solutions: the solutions of a degenerate separable multiperiodic system with N degrees of freedom and D independent commensurability relations are $(N-D)$-periodic. When $D = N - 1$, the system is said to be *completely degenerate.* For instance, the degeneracy of the Keplerian motion is complete, since we may write two independent commensurability relations, viz. $\nu_\theta - \nu_r = 0$ and $\nu_\phi - \nu_\theta = 0$. As a consequence, the Keplerian motion is periodic. The central motions of Sect. 2.5 are always degenerate, since $\nu_\phi - \nu_\theta = 0$. However, they are not completely degenerate, except in some particular cases such as Keplerian motion and the harmonic oscillator (Bertrand's theorem). In these cases, besides $\nu_\phi - \nu_\theta = 0$, we also have $\nu_\theta - \nu_r = 0$. For other laws of force, a second commensurability relation may only occur for given initial conditions (accidental degeneracy or resonance).

In Kolmogorv's theorem, the non-degeneracy of an integrable Hamiltonian $H(J)$ is defined as

$$\det\left(\frac{\partial^2 H}{\partial J_i \partial J_j}\right) \neq 0, \tag{2.85}$$

which guarantees the reversibility of the transformation from actions to frequencies. This definition is more restrictive than Schwarzschild's. Indeed, all Hamiltonians linear in one of the actions are degenerate in Kolmogorov's sense[9]. For these Hamiltonians, one whole row of the Hessian determinant consists of zeros. It happens that a common operation in the applications of Hamiltonian Mechanics to Astronomy is the extension of the phase space, because of time-dependent applied forces. In such an extension, a new generalized momentum (or action) is added to the given Hamiltonian. The extended Hamiltonian will always be such that the Hessian determinant is zero. If the condition given by (2.85) were a universal restriction, almost all dynamical systems of Astronomy would be excluded from the possibility of application of the theories discussed in this book. However, when frequency relocation is not done, the most general non-degeneracy condition is Schwarzschild's, that is, $(h \mid \nu) \neq 0$ for all $h \in \mathbf{D_k} \subset \mathbf{Z}^N \backslash 0$.

2.7.1 Schwarzschild Transformation

In the study of degenerate systems, it is often convenient to redefine angles and actions to introduce angles whose frequencies are equal to zero. Let a separable multiperiodic system of N degrees of freedom have L essential commensurability relations

[9] For a more accurate discussion, see Sect. 3.11.4.

$$\sum_{k=1}^{N} h_k^{(\varrho)} \nu_k = 0 \qquad\qquad \varrho = N - L + 1, \cdots, N \tag{2.86}$$

and let us introduce the point transformation of the angles,

$$\begin{array}{lcl} w_1 & = & \ell_1 \\ w_2 & = & \ell_2 \\ \cdots & & \\ w_M & = & \ell_M \\ \sum_k h_k^{(M+1)} w_k & = & \ell_{M+1} \\ \cdots & & \\ \sum_k h_k^{(N)} w_k & = & \ell_N, \end{array} \tag{2.87}$$

where, for simplicity, we have introduced $M = N - L$. Extending this transformation to the momenta, we obtain

$$\begin{array}{lcll} J_1 & = & x_1 & + \sum_\varrho h_1^{(\varrho)} x_\varrho \\ J_2 & = & x_2 & + \sum_\varrho h_2^{(\varrho)} x_\varrho \\ \cdots & & & \\ J_M & = & x_M & + \sum_\varrho h_M^{(\varrho)} x_\varrho \\ J_{M+1} & = & & \sum_\varrho h_{M+1}^{(\varrho)} x_\varrho \\ \cdots & & & \\ J_N & = & & \sum_\varrho h_N^{(\varrho)} x_\varrho, \end{array} \tag{2.88}$$

where the x_k are the momenta conjugate to the new angles ℓ_k.

The angles $\ell_\mu \ \ (\mu = 1, \cdots, M)$ are called *non-degenerate*[10] while the remaining ones, $\ell_\varrho \ \ (\varrho = M + 1, \cdots, N)$, are called *degenerate*.

With the new variables, the Hamiltonian depends only on the actions conjugate to non-degenerate angles. Thus, the frequencies of the degenerate angles are

$$\tilde{\nu}_\varrho = \frac{\mathrm{d}\ell_\varrho}{\mathrm{d}t} = \frac{\partial \tilde{H}(x)}{\partial x_\varrho} = 0 \qquad\qquad (\varrho = M + 1, \cdots, N). \tag{2.89}$$

The equations $\tilde{\nu}_\varrho = 0$ are the new commensurability relations.

2.7.2 Delaunay Variables

The usual angle–action variables of the Keplerian motion, the Delaunay variables, are the result of the application of the Schwarzschild transformation

[10] The actions conjugate to non-degenerate angles are sometimes called *proper*. However, the word *proper* is used, in this book, to indicate the almost constant actions resulting from an averaging process. Thus, to avoid ambiguities, the word proper will not be used to mean *non-degenerate*.

to the angle–action variables obtained in Sect. 2.6. Indeed, in this case, the commensurabilities are

$$\begin{aligned} \nu_\theta - \nu_r &= 0 \\ \nu_\phi - \nu_\theta &= 0. \end{aligned} \tag{2.90}$$

Then

$$\begin{aligned} \ell_1 &= w_r &&= \ell \\ \ell_2 &= w_\theta - w_r &&= \omega \\ \ell_3 &= w_\phi - w_\theta &&= \Omega \end{aligned} \tag{2.91}$$

and

$$\begin{aligned} J_r &= x_1 - x_2 \\ J_\theta &= x_2 - x_3 \\ J_\phi &= x_3 \end{aligned} \tag{2.92}$$

or

$$\begin{aligned} x_1 &= J_r + J_\theta + J_\phi &&= m\sqrt{\mu a} \\ x_2 &= J_\theta + J_\phi &&= m\sqrt{\mu a(1-e^2)} \\ x_3 &= J_\phi &&= m\sqrt{\mu a(1-e^2)}\cos i. \end{aligned} \tag{2.93}$$

For $m = 1$, these variables are exactly the variables ℓ, g, h, L, G, H of Delaunay. Indeed, point dynamics problems often are such that the mass of the moving particle cancels in the equations and does not affect the results. In this case, energies, momenta and actions are considered per unit mass and we write

$$\begin{aligned} x_1 &= \sqrt{\mu a} \\ x_2 &= x_1\sqrt{1-e^2} \\ x_3 &= x_2 \cos i \end{aligned} \tag{2.94}$$

and

$$E = -\frac{\mu^2}{2x_1^2}. \tag{2.95}$$

2.8 The Separable Cases of Liouville and Stäckel

Autonomous systems whose energy consists of a kinetic energy quadratic in the velocities and a potential energy independent of the velocities have been thoroughly studied in the past. Sufficient conditions for their separability were established by Liouville and Stäckel. These cases are generally presented as sets of conditions for the potential and kinetic energies, separately. In what follows, kinetic and potential energies are considered together to give a set of conditions for the Hamiltonian; this choice is more appropriate for the scope of this book.

Theorem 2.8.1 (Liouville). *The dynamical systems whose Hamiltonian may be written as*

$$H = \frac{f_1(q_1, p_1) + \cdots + f_N(q_N, p_N)}{g_1(q_1, p_1) + \cdots + g_N(q_N, p_N)} \tag{2.96}$$

are separable.

The Hamilton–Jacobi equation in this case is

$$\sum_{i=1}^{N} f_i \left(q_i, \frac{\partial S}{\partial q_i} \right) = E \sum_{i=1}^{N} g_i \left(q_i, \frac{\partial S}{\partial q_i} \right) \tag{2.97}$$

which, after the introduction of $S = \sum_i S_i(q_i)$, may be separated into N equations

$$f_i \left(q_i, \frac{\mathrm{d}S_i}{\mathrm{d}q_i} \right) - E g_i \left(q_i, \frac{\mathrm{d}S_i}{\mathrm{d}q_i} \right) = \beta_i, \tag{2.98}$$

the integration constants β_i being such that $\sum_i \beta_i = 0$. These equations may be solved with respect to $\mathrm{d}S_i/\mathrm{d}q_i$ when

$$\left(\frac{\partial f_i}{\partial p_i} \right) - E \left(\frac{\partial g_i}{\partial p_i} \right) \neq 0 \qquad \text{for all } i.$$

□

Theorem 2.8.2 (Stäckel). *The dynamical systems whose Hamiltonian may be written as*

$$H = \frac{1}{\Delta} \sum_{i=1}^{N} A_i f_i(q_i, p_i), \tag{2.99}$$

where Δ is the determinant of a square matrix of rank N in which each column depends only on the coordinate of the same subscript as the column:

$$\Delta = \det \left(a_{ji}(q_i) \right), \tag{2.100}$$

and the A_i are the cofactors of the elements of any of the rows of the matrix, are separable.

The Hamilton–Jacobi equation in this case is

$$\sum_{i=1}^{N} A_i f_i \left(q_i, \frac{\partial S}{\partial q_1} \right) = E\Delta. \tag{2.101}$$

This partial differential equation has a complete integral of the form $S = \sum_i S_i(q_i)$. If we assume, for instance, that the A_i are the cofactors of the elements of the first row, the theorems of Laplace allow us to write

$$\begin{aligned}\sum_{i=1}^{N} a_{1i} A_i &= \Delta \qquad (2.102)\\ \sum_{i=1}^{N} a_{\varrho i} A_i &= 0 \qquad (\varrho = 2, \cdots, N).\end{aligned}$$

Because of these relations, the Hamilton–Jacobi equation is not affected when we introduce the sum

$$-\sum_{i=1}^{N} A_i \sum_{\varrho=2}^{N} \beta_\varrho a_{\varrho i}(q_i),$$

where the β_ϱ are $N-1$ arbitrary constants. Using also the Laplacian expression for Δ, the Hamilton–Jacobi equation becomes

$$\sum_{i=1}^{N} A_i \left(f_i \left(q_i, \frac{\partial S}{\partial q_i} \right) - \sum_{\varrho=2}^{N} \beta_\varrho a_{\varrho i}(q_i) \right) = E \sum_{i=1}^{N} a_{1i}(q_i) A_i, \qquad (2.103)$$

which may be separated into N equations

$$f_i \left(q_i, \frac{\mathrm{d}S_i}{\mathrm{d}q_i} \right) = \sum_{\varrho=2}^{N} \beta_\varrho a_{\varrho i}(q_i) + E a_{1i}(q_i). \qquad (2.104)$$

These equations may be solved with respect to $\mathrm{d}S_i/\mathrm{d}q_i$ if

$$\frac{\partial f_i}{\partial p_i} \neq 0 \qquad \text{for all } i.$$

□

2.8.1 Example: Liouville Systems

The original form of Liouville's separability conditions says that the kinetic and potential energies may be written, respectively, as

$$T = \frac{1}{2}(A_1 + A_2 + \cdots + A_N)(B_1 \dot{q}_1^2 + B_2 \dot{q}_2^2 + \cdots + B_N \dot{q}_N^2) \qquad (2.105)$$

and

$$V = \frac{V_1 + V_2 + \cdots + V_N}{A_1 + A_2 + \cdots + A_N}, \qquad (2.106)$$

where $A_i = A_i(q_i), B_i = B_i(q_i)$ and $V_i = V_i(q_i)$. (The function with subscript i depends only on the generalized coordinate q_i.) A simple calculation shows that the energy $H = T + V$ has the form given in the above theorem and that the Hamilton–Jacobi equation is separated into the N equations:

$$\frac{1}{2B_k} \left(\frac{\mathrm{d}S_k}{\mathrm{d}q_k} \right)^2 = E A_k + \beta_k - V_k. \qquad (2.107)$$

2.8.2 Example: Stäckel Systems

The original form of Stäckel's separability conditions says that the kinetic and potential energies must be, respectively,

$$T = \frac{1}{2}\Delta\left(\frac{\dot{q}_1^2}{A_1} + \frac{\dot{q}_2^2}{A_2} + \cdots + \frac{\dot{q}_N^2}{A_N}\right) \tag{2.108}$$

and

$$V = \frac{1}{\Delta}\sum_{i=i}^{N} g_i(q_i)\, A_i, \tag{2.109}$$

where Δ and A_i are the same as in the given theorem. The resulting energy $H = T + V$ has the form as given in the theorem and the Hamilton–Jacobi equation is separated into the N equations:

$$\frac{1}{2}\left(\frac{\mathrm{d}S_i}{\mathrm{d}q_i}\right)^2 = \sum_{\varrho=2}^{N} \beta_\varrho a_{\varrho i}(q_i) + E a_{1i}(q_i) - g_i(q_i). \tag{2.110}$$

2.8.3 Example: Central Motions

The example of the motion of a particle in a central force field, considered in the previous section, is also an example of a separable Stäckel system. The Hamiltonian of this system is (see 2.52):

$$H = \frac{1}{2m}\left[p_r^2 + \frac{1}{r^2}\left(p_\theta^2 + \frac{p_\phi^2}{\sin^2\theta}\right)\right] + V(r). \tag{2.111}$$

In order to see that this system satisfies the conditions of the Stäckel theorem, we introduce the matrix

$$(a_{ij}) = \begin{pmatrix} -r^{-2} & 1 & 0 \\ 0 & -\sin^{-2}\theta & 1 \\ 1 & 0 & 0 \end{pmatrix} \tag{2.112}$$

whose determinant is $\Delta = 1$ and the cofactors of the elements of the third row are:

$$\begin{aligned} A_1 &= 1 \\ A_2 &= r^{-2} \\ A_3 &= (r\sin\theta)^{-2}. \end{aligned} \tag{2.113}$$

Comparison to (2.97) shows that the functions f_i are

$$\begin{aligned} f_1 &= \frac{1}{2m}p_r^2 + V(r) \\ f_2 &= \frac{1}{2m}p_\theta^2 \\ f_3 &= \frac{1}{2m}p_\phi^2 \end{aligned} \tag{2.114}$$

and the proof is completed.

2.9 Angle–Action Variables of a Quadratic Hamiltonian

Let us consider the case of a Hamiltonian given by a quadratic form in q, p, with purely imaginary eigenvalues. Let it be

$$H_2(z) = \sum_{i,j=1}^{2N} \frac{1}{2}\, a_{ij} z_i z_j, \tag{2.115}$$

where $z \equiv (q, p) \in \mathbf{R}^{2N}$. In this case, the techniques discussed in the previous sections to obtain angle–action variables cannot be used because the Hamiltonian does not have the form of the considered separable systems. However, the resulting differential equations are homogeneous and linear with constant coefficients and a few steps are enough to solve them. These equations are

$$\frac{\mathrm{d}z}{\mathrm{d}t} = -\mathsf{J}\frac{\partial H_2}{\partial z} = -\mathsf{JS}z, \tag{2.116}$$

where J is the symplectic unit matrix of order $2N$ and $\mathsf{S} = \left(\dfrac{\partial^2 H_2}{\partial z_i \partial z_j}\right) = (a_{ij})$ is the Hessian matrix of H_2. Let λ_i and $\tilde{A}_i$ be, respectively, the eigenvalues and eigenvectors of $-\mathsf{JS}$. If we assume that all eigenvalues are distinct, the general solution of (2.116) is

$$z = \sum_{i=1}^{2N} c_i \tilde{A}_i \exp \lambda_i t, \tag{2.117}$$

where c_i are arbitrary constants. The characteristic polynomial $P(\lambda) = \det(-\mathsf{JS} - \lambda\mathsf{I})$ is even and, if λ is an eigenvalue of $-\mathsf{JS}$, then so is $-\lambda$. The eigenvalues of $-\mathsf{JS}$, which were assumed to be imaginary, may thus be written as

$$\lambda_k = -\mathrm{i}\omega_k, \qquad \lambda_{N+k} = \mathrm{i}\omega_k \qquad (k = 1, 2, \cdots, N). \tag{2.118}$$

Let us now consider the matrix formed by the $2N$ eigenvectors, $\mathsf{A} \equiv (\tilde{A}_i)$, its transpose A' and let us form the matrix $\mathsf{R} = \mathsf{A}'\mathsf{JA}$. A simple calculation shows that the elements of R are

$$\varrho_{ij} \stackrel{\text{def}}{=} \tilde{A}_i' \mathsf{J} \tilde{A}_j.$$

We have to prove the following lemma (see [71] Sect. II.C):

Lemma 2.9.1. *If $\tilde{A}_i$ and $\tilde{A}_j$ are eigenvectors of $-\mathsf{JS}$ corresponding to two eigenvalues λ_i, λ_j such that $\lambda_i + \lambda_j \neq 0$, then $\tilde{A}_i'\mathsf{J}\tilde{A}_j = 0$.*

The proof of this statement is very simple. We just have to recall that the eigenvalue λ_i and the eigenvector $\tilde{A}_i$ corresponding to it are related by $\mathsf{JS}\tilde{A}_i = -\lambda_i\tilde{A}_i$. It then follows that:

$$\begin{aligned} \lambda_i\tilde{A}_i'\mathsf{J}\tilde{A}_j &= -\tilde{A}_i'\mathsf{S}\tilde{A}_j \qquad \text{and} \\ \lambda_j\tilde{A}_i'\mathsf{J}\tilde{A}_j &= \tilde{A}_i'\mathsf{S}\tilde{A}_j; \end{aligned}$$

and so, $(\lambda_i + \lambda_j)\tilde{A}_i'\mathsf{J}\tilde{A}_j = 0$, that is, $\tilde{A}_i'\mathsf{J}\tilde{A}_j = 0$.

Corollary 2.9.1. $\tilde{A}_i'\mathsf{J}\tilde{A}_i = 0$ *for all* $i = 1, 2, \cdots, 2N$.

The following lemma is trivial.

Lemma 2.9.2. *For all $i, j = 1, 2, \cdots, 2N$, we have $\tilde{A}_i'\mathsf{J}\tilde{A}_j = -\tilde{A}_j'\mathsf{J}\tilde{A}_i$.*

A consequence of these lemmas is that the only terms of R that may be different from zero are those arising from eigenvectors corresponding to pairs of eigenvalues $\pm i\omega_k$. We assume $\varrho_{ij} \neq 0$ for the pairs i, j such that $|j-i| = N$. Otherwise, $\varrho_{ij} = 0$:

$$\mathsf{R} = \begin{pmatrix} 0 & 0 & \cdots & -\varrho_{N+1,1} & 0 & \cdots \\ 0 & 0 & \cdots & 0 & -\varrho_{N+2,2} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \varrho_{N+1,1} & 0 & \cdots & 0 & 0 & \cdots \\ 0 & \varrho_{N+2,2} & \cdots & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \tag{2.119}$$

Therefore, it is enough to rescale the eigenvectors (dividing the $\tilde{A}_k$ and $\tilde{A}_{N+k}$ by $\sqrt{\varrho_{N+k,k}}$ for all k) to obtain J instead of R. If D is the diagonal matrix

$$\mathsf{D} \stackrel{\text{def}}{=} \operatorname{diag}\left(\frac{1}{\sqrt{\varrho_{N+1,1}}}, \cdots, \frac{1}{\sqrt{\varrho_{2N,N}}}, \frac{1}{\sqrt{\varrho_{N+1,1}}}, \cdots, \frac{1}{\sqrt{\varrho_{2N,N}}}\right),$$

the matrix $\mathsf{M} = \mathsf{AD}$ is such that $\mathsf{M}'\mathsf{JM} = \mathsf{J}$ and therefore, the linear transformation $\zeta \to z = \mathsf{AD}\zeta$ is canonical (see 1.36).

If we compare the equation of this transformation to (2.117), we obtain for the new canonical variables,

$$\zeta_k = c_k\sqrt{\varrho_{N+k,k}}\ \mathrm{e}^{\lambda_k t} \qquad \zeta_{N+k} = c_{N+k}\sqrt{\varrho_{N+k,k}}\ \mathrm{e}^{\lambda_{N+k} t}$$

$(k = 1, 2, \cdots, N)$. To complete the construction of the angle–action variables (w, J) of H_2, it is enough to introduce them through the Poincaré-like complex canonical variables $\sqrt{\mathrm{i}J_k}\ \mathrm{e}^{-\mathrm{i}w_k}$ and $\sqrt{\mathrm{i}J_k}\ \mathrm{e}^{\mathrm{i}w_k}$ and compare them to ζ. We get

$$\begin{aligned} w_k &= \omega_k t - \alpha_k \\ J_k &= -\mathrm{i}|c_k|^2 \varrho_{N+k,k} \qquad (k = 1, 2, \cdots, N), \end{aligned} \tag{2.120}$$

where α_k is the argument of c_k. Because of the rules of conjugation, it is enough to work with the equations giving the first N variables ζ_k. The other N equations repeat the same results. It is worth stressing some points: (i) the J_k are real since the $\varrho_{N+k,k}$ are imaginary; (ii) the J_k may be either positive or negative, according to the sign of $-\mathrm{i}\varrho_{N+k,k}$; (iii) the N complex integration constants c_k are changed into α_k, J_k; (iv) c_k and c_{N+k} are complex conjugates.

The direct comparison of equations (2.117) and (2.120) gives

$$z = \sum_{k=1}^{N} \sqrt{\frac{\mathrm{i}J_k}{\varrho_{N+k,k}}} \, (\tilde{A}_k \, \mathrm{e}^{-\mathrm{i}w_k} + \tilde{A}_{N+k} \, \mathrm{e}^{\mathrm{i}w_k}). \tag{2.121}$$

This equation is consistent with the fact that z is a real vector.

In terms of the angle–action variables, the new Hamiltonian follows straightforwardly from the equations $\partial H/\partial J_k = \dot{w}_k = \omega_k$, whose integration gives $H = \sum_k \omega_k J_k$, or, as a function of ζ, $H = -\sum_k \mathrm{i}\omega_k \zeta_k \zeta_{N+k}$. If we compare this result to

$$H = \frac{1}{2} \boldsymbol{z}' \mathsf{S} \boldsymbol{z} = \frac{1}{2} \boldsymbol{\zeta}' \mathsf{DA'SAD} \boldsymbol{\zeta},$$

we see that

$$\mathsf{DA'SAD} = \begin{pmatrix} 0 & 0 & \cdots & -\mathrm{i}\omega_1 & 0 & \cdots \\ 0 & 0 & \cdots & 0 & -\mathrm{i}\omega_2 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -\mathrm{i}\omega_1 & 0 & \cdots & 0 & 0 & \cdots \\ 0 & -\mathrm{i}\omega_2 & \cdots & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

This matrix is the Hessian of H calculated with respect to the new canonical variables ζ. It could be easily obtained from the properties of the matrices D, A and S, using the lemma given in Exercise 2.9.6, below.

Exercise 2.9.1. Show that the characteristic polynomial $P(\lambda) = \det(-\mathsf{JS} - \lambda \mathsf{I})$ is even.

Exercise 2.9.2. Show that the eigenvectors $\tilde{A}_k$ and $\tilde{A}_{N+k}$ of $-\mathsf{JS}$ corresponding to two complex conjugate eigenvalues are complex conjugate themselves.

Exercise 2.9.3. Show that, for $|i - j| = N$, the ϱ_{ij} are imaginary.

Exercise 2.9.4. Show that the transformation $(w, J) \to (\sqrt{\mathrm{i}J} \, \mathrm{e}^{-\mathrm{i}w}, \sqrt{\mathrm{i}J} \, \mathrm{e}^{\mathrm{i}w})$ is canonical.

Exercise 2.9.5. Show that ζ_k and ζ_{N+k} are _not_ complex conjugates.

Exercise 2.9.6 (Lemma). Prove that for all $i, j = 1, 2, \cdots, 2N$, we have

$$\tilde{A}'_i \mathsf{S} \tilde{A}_j = -\lambda_j \varrho_{ji}.$$

Hint: Use the characteristic equation $\mathsf{JS}\tilde{A}_j = -\lambda_j \tilde{A}_j$.

2.9.1 Gyroscopic Systems

Let us consider the important particular case of the two-degrees-of-freedom gyroscopic system whose Hamiltonian is

$$H = \frac{\boldsymbol{p}^2}{2} - [\boldsymbol{k}, \boldsymbol{r}, \boldsymbol{p}] + W(\boldsymbol{r}), \tag{2.122}$$

where $\boldsymbol{r} \equiv (x, y), \boldsymbol{p} \equiv (p_x, p_y)$, $\boldsymbol{k}$ is a unit vector perpendicular to the plane of motion and the potential energy is $W = \frac{1}{2}(ax^2 + by^2) + dxy$ $(a, b, d$ are constants). (See Sect. 1.7; for the sake of simplicity, we have chosen units such that $m = 1$ and $|\boldsymbol{\Omega}| = 1$.) The Hessian matrix is

$$\mathsf{S} = \begin{pmatrix} a & d & 0 & -1 \\ d & b & 1 & 0 \\ 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}. \tag{2.123}$$

Then,

$$-\mathsf{JS} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ -a & -d & 0 & 1 \\ -d & -b & -1 & 0 \end{pmatrix} \tag{2.124}$$

and the eigenvalues of $-\mathsf{JS}$ are

$$\lambda_j = \pm\frac{1}{2}\sqrt{-2(a+b+2) \pm 2\sqrt{(a-b)^2 + 8(a+b) + 4d^2}}. \tag{2.125}$$

We assume that these eigenvalues are imaginary and write them as $\pm\mathrm{i}\omega_1$ and $\pm\mathrm{i}\omega_2$. This means that the parameters a, b, d of the given function W are such that $\phi \stackrel{\mathrm{def}}{=} (a-b)^2 + 8(a+b) + 4d^2 \geq 0$ and $-(a+b+2) + \sqrt{\phi} < 0$.

The eigenvectors of $-\mathsf{JS}$ are

$$\tilde{A}_j = \begin{pmatrix} -\lambda_j^3 - (b+1)\lambda_j + d \\ \lambda_j^2 + d\lambda_j - a + 1 \\ a\lambda_j^2 - b + ab - d^2 \\ d(\lambda_j^2 + 1) - (a+b)\lambda_j \end{pmatrix}. \tag{2.126}$$

The quantities $\varrho_{k+2,k}$ are immediate. We just point out the fact that, of the five parameters $a, b, d, \omega_1, \omega_2$, only three are independent. We use (2.125) to eliminate b, d and obtain

$$\begin{aligned} \varrho_{31} &= 2\mathrm{i}\,\omega_1(\omega_1^2 - \omega_2^2)(1 - a + a\omega_1^2 - \omega_1^2\omega_2^2) \\ \varrho_{42} &= 2\mathrm{i}\,\omega_2(\omega_2^2 - \omega_1^2)(1 - a + a\omega_2^2 - \omega_1^2\omega_2^2). \end{aligned} \tag{2.127}$$

The new angle–action variables are

$$\begin{aligned} w_k &= \omega_k t - \alpha_k \\ J_k &= -\mathrm{i}|c_k|^2 \varrho_{N+k,k} \qquad (k = 1, 2, \cdots, N), \end{aligned} \tag{2.128}$$

where $c_k = |c_k|\mathrm{e}^{\mathrm{i}\alpha_k}$ are the integration constants.