Birkhäuser Advanced Texts
Basler Lehrbücher

Gerardo F. Torres del Castillo

An Introduction to Hamiltonian Mechanics

Birkhäuser

Birkhäuser

Birkhäuser Advanced Texts Basler Lehrbücher

More information about this series at http://www.springer.com/series/4842

Gerardo F. Torres del Castillo

An Introduction to Hamiltonian Mechanics

Birkhäuser

Gerardo F. Torres del Castillo
Instituto de Ciencias, BUAP
Puebla, Puebla, Mexico

ISSN 1019-6242 ISSN 2296-4894 (electronic)
Birkhäuser Advanced Texts Basler Lehrbücher
ISBN 978-3-319-95224-6 ISBN 978-3-319-95225-3 (eBook)
https://doi.org/10.1007/978-3-319-95225-3

Library of Congress Control Number: 2018948684

Mathematics Subject Classification (2010): 34A34, 70-01, 70E17, 70F20, 70H03, 70H05, 70H15, 70H20, 70H25, 70H33, 78A05

This book is published under the imprint Birkhäuser, www.birkhauser-science.com, by the registered company Springer Nature Switzerland AG part of Springer Nature.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The aim of this book is to present in an elementary manner the fundamentals of the Hamiltonian formulation of classical mechanics, making use of a basic knowledge of linear algebra (matrices, properties of the determinant and the trace), differential calculus in several variables (the differential of a function of several variables, the chain rule, and the inverse function theorem), analytic geometry, and ordinary differential equations. Even though the main purpose of this book is the exposition of the Hamiltonian formalism of classical mechanics, the first chapters are devoted to the Lagrangian formalism and, for a reader not familiarized with the Lagrange equations, these introductory chapters should suffice to understand the basic elements of analytical mechanics.

This book is intended for advanced undergraduate or graduate students in physics or applied mathematics and for researchers working in related subjects. It is assumed that the reader has some familiarity with some elementary notions about classical mechanics, such as the inertial reference frames and Newton's second law. This book has been written having in mind readers trying to learn the subject by themselves, including detailed examples and exercises with complete solutions, but it can also be used as a class text.

This book does not attempt to be an exhaustive treatment of analytical mechanics or even of the Hamiltonian formulation of classical mechanics. Deliberately, some subjects are not treated in this book; the subjects not included here have been omitted for at least one of the following reasons: they are not essential to understand the basic formalism, or there exist books or articles containing a good discussion of the subject that would be difficult to improve.

Some of the subjects not treated here are dissipative systems, nonholonomic constraints, adiabatic invariants, action-angle variables for systems with more than one degree of freedom, perturbation theory, continuous systems, normal modes of vibration, integral invariants, relativistic mechanics, singular Lagrangians, KAM theory, and chaos.

Throughout the book I have avoided the use of terms like "small parameter," "small variation," "infinitesimal parameter," "infinitesimal transformation," and a diversity of variations denoted by symbols like Δ, δ, $\overline{d}$, $\tilde{\delta}$, $\tilde{\tilde{\delta}}$, and so on. Apart from this, the notation employed throughout the book coincides with that found in the traditional books on analytical mechanics.

While there seems to exist a consensus about the importance of the Hamiltonian formalism, in the standard textbooks on analytical mechanics one finds a variety of opinions about the Hamilton–Jacobi equation, ranging from the conviction that the Hamilton–Jacobi formalism is the most powerful tool of analytical mechanics to the claim that the Hamilton–Jacobi equation has no practical value and that it is only interesting because of its relation with the Schrödinger equation. The point of view adopted here is that, apart from its deep and useful connections with quantum mechanics, the Hamilton–Jacobi formalism is interesting and useful by itself, as I have tried to illustrate in Chapter 6.

Among the differences between this book and the existing textbooks on the subject are the presence of:

- Application of the various formulations to equations not related to classical mechanics.
- Systematic use of equivalent Hamiltonians, which allows us to relate different problems and to find constants of motion without integrating the equations of motion.
- Detailed derivations of the canonical transformations, emphasizing the distinction between what is a generating function and what is not.
- Study of the continuous groups of canonical transformations, avoiding the use of "infinitesimals."
- Precise definition and examples of the symmetries of a Hamiltonian, including the case of transformations that involve the time explicitly.
- Emphasis on the fact that, in the Hamiltonian formalism, there are infinitely many generating functions of translations and rotations (and, therefore, that, e.g., the linear momentum cannot be defined as *the* generating function of translations).
- Study of the canonoid transformations and the associated constants of motion.
- Definition and examples of R-separable solutions of the Hamilton–Jacobi equation.
- General statement, simplified proof, and detailed examples of the Liouville theorem on solutions of the Hamilton–Jacobi equation.
- Discussion and examples of the mapping of solutions of the Hamilton–Jacobi equation under canonical transformations.
- Discussion of the Hamilton–Jacobi equation as an evolution equation for the principal function.
- Presentation of geometrical optics as an application of the Hamiltonian formalism.
- Detailed solution of all the exercises.

Many textbooks on analytical mechanics written in the last few decades make use of the language of modern differential geometry (manifolds, vector fields, differential forms), which is particularly useful and elegant when the Hamiltonian does not depend on the time. In fact, in classical mechanics one has one of the nicest applications of this formalism. One of the aims of this book is to show that it is possible to obtain many interesting results making use of elementary mathematics only.

Throughout the book, there is a collection of examples worked out in detail, which form an essential part of the book, and a set of exercises is also given. It is advisable that the reader attempts to solve them all and to fill in the details of the computations presented in the book. The detailed solutions of all the exercises are collected at the end of the book, but the reader is encouraged to try to find the solutions before seeing the answers. Some sections go beyond the basic level and can be skipped; these sections are Section 2.5 (Variational Symmetries), Section 4.3.1 (The Kepler Problem Revisited), Section 5.5 (Canonoid Transformations), Section 6.3 (Mapping of Solutions of the Hamilton–Jacobi Equation Under Canonical Transformations), Section 6.4 (Transformation of the Hamilton–Jacobi Equation Under Arbitrary Point Transformations), and Section 6.5 (Geometrical Optics).

Throughout the book, references are given to some books or papers when the subject is not commonly treated in the standard textbooks.

Some words about the notation: I have avoided the use of superscripts to label coordinates or components, considering that they are not indispensable at this level and, sometimes, its use can complicate the expressions. The sign $\equiv$ indicates a definition and in all cases it should be clear which side of the sign contains the object being defined.

I would like to thank Dr. Iraís Rubalcava-García for her help with the figures and the reviewers for their helpful comments. I would also like to thank Samuel DiBella at Springer Nature for his valuable support.

Puebla, Puebla, Mexico Gerardo F. Torres del Castillo
May 2018

Contents

Chapter 1
The Lagrangian Formalism

In this chapter we show that the equations of motion of certain mechanical systems, obtained from Newton's second law, can be expressed in a convenient manner in terms of a single real-valued function. In Section 1.1 we analyze some simple mechanical systems, which serve to introduce various concepts and to formulate the d'Alembert principle; then, following a standard procedure, assuming that the applied forces are derivable from a potential, in Section 1.2 we obtain the Lagrange equations. Several examples of the application of the Lagrange equations are presented.

1.1 Introductory Examples. The D'Alembert Principle

In what follows we assume that the reader is familiarized with the basic notions of classical mechanics, such as the concept of inertial reference frame, the Newton laws of motion, and the definitions of kinetic and potential energy.

We begin by analyzing some simple mechanical systems, which will allow us to introduce several concepts that will be useful later. Most of the examples considered in this section will be treated in Section 1.2, making use of the Lagrangian formalism.

Example 1.1. We consider the system formed by a wedge of mass m_1 which can move on a frictionless horizontal surface along a straight line, which we take as the x-axis (see Figure 1.1), and a block of mass m_2 which moves on the wedge, without friction. The position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, shown in Figure 1.1, can be parameterized in the form

$$\mathbf{r}_1 = x\,\mathbf{i}, \qquad \mathbf{r}_2 = (x + y\ \cot\theta_0)\,\mathbf{i} + y\,\mathbf{j}, \tag{1.1}$$

G. F. Torres del Castillo, *An Introduction to Hamiltonian Mechanics*, Birkhäuser Advanced Texts Basler Lehrbücher, https://doi.org/10.1007/978-3-319-95225-3_1

but there are infinitely many alternative parameterizations, e.g.,

$$\mathbf{r}_1 = x\,\mathbf{i}, \qquad \mathbf{r}_2 = x\,\mathbf{i} + s(\cos\theta_0\,\mathbf{i} + \sin\theta_0\,\mathbf{j}) \tag{1.2}$$

(see Figure 1.2). The parameters employed to express the position vectors *need not* be lengths or angles; the only requisites are that these parameters be independent and determine the configuration of the system. For instance, looking at Figure 1.1 one can convince oneself that the value of x does not determine the value of y or vice versa, whereas Equations (1.1) show that x and y completely determine the position vectors $\mathbf{r_1}$ and $\mathbf{r_2}$, which, in turn, define the configuration of the system.

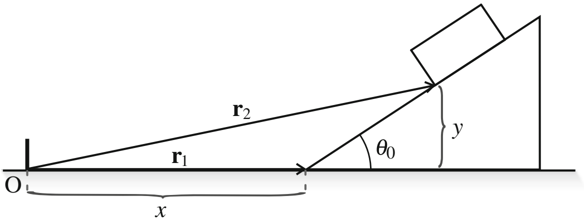
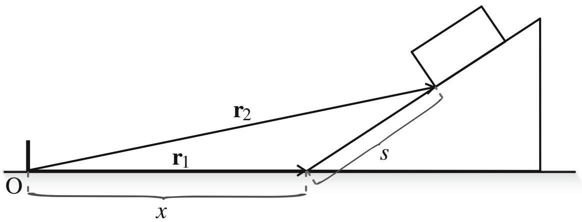


Fig. 1.1 The block can slide, without friction, on the wedge-shaped block, which can move freely on a horizontal surface. The configuration of the system can be specified giving the values of x and y, which determine the position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$

Fig. 1.2 The position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ can be expressed in terms of the lengths x and s shown



The forces acting *on the wedge* are an upward vertical force $N_1\,\mathbf{j}$, produced by the horizontal surface on which the wedge lies, a force $N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j})$, produced by the block, and the weight $-m_1 g\,\mathbf{j}$. Then, assuming that the coordinate axes define an *inertial reference frame*, according to Newton's second law we have

$$m_1\ddot{\mathbf{r}}_1 = N_1\,\mathbf{j} + N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j}) - m_1 g\,\mathbf{j}. \tag{1.3}$$

Similarly, the forces acting *on the block* are $-N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j})$, due to the contact with the wedge (taking into account Newton's third law) and the gravitational force $-m_2 g\,\mathbf{j}$. Hence,

$$m_2\ddot{\mathbf{r}}_2 = -N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j}) - m_2 g\,\mathbf{j}. \tag{1.4}$$

Substituting the expressions (1.1) into the left-hand sides of (1.3) and (1.4) we obtain the system of ordinary differential equations (ODEs)

$$\begin{aligned} m_1\ddot{x}\,\mathbf{i} &= N_1\,\mathbf{j} + N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j}) - m_1 g\,\mathbf{j}, \\ m_2[(\ddot{x} + \ddot{y}\cot\theta_0)\,\mathbf{i} + \ddot{y}\,\mathbf{j}] &= -N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j}) - m_2 g\,\mathbf{j}, \end{aligned} \tag{1.5}$$

which amounts to four scalar equations that determine the values of $N_1, N_2, \ddot{x}$, and $\ddot{y}$.

Note that we only know the direction of the normal forces, $N_1\,\mathbf{j}$ and $\pm N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j})$; their magnitudes, N_1 and N_2, must have the exact values necessary to maintain the two bodies in contact, and the wedge in contact with the horizontal surface, without deformations. These forces are known as *constraint forces* and, as we shall see now, there is a simple way to eliminate them from the equations of motion.

We consider the sums

$$\sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial x} \qquad \text{and} \qquad \sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial y}, \tag{1.6}$$

where $\mathbf{F}_\alpha^{(\text{constr})}$ is the sum of the constraint forces acting on the α-th body, and x, y are the only two parameters appearing in (1.1). In this case, $\mathbf{F}_1^{(\text{constr})} = N_1\,\mathbf{j} + N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j})$, and $\mathbf{F}_2^{(\text{constr})} = -N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j})$. Making use of Equations (1.1) we find that

$$\begin{aligned} \sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial x} &= \big[N_1\,\mathbf{j} + N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j})\big] \cdot \mathbf{i} + \big[-N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j})\big] \cdot \mathbf{i} \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} \sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial y} &= \big[N_1\,\mathbf{j} + N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j})\big] \cdot \mathbf{0} \\ &\quad + \big[-N_2(\sin\theta_0\,\mathbf{i} - \cos\theta_0\,\mathbf{j})\big] \cdot (\cot\theta_0\,\mathbf{i} + \mathbf{j}) = 0. \end{aligned}$$

That is, both sums in (1.6) are equal to zero. In what follows it will be convenient to write Equations (1.6) in the unified form

$$\sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} = 0, \tag{1.7}$$

where, e.g., $q_1 = x$ and $q_2 = y$.

Writing Equations (1.3) and (1.4) in the form

$$m_\alpha \ddot{\mathbf{r}}_\alpha = \mathbf{F}_\alpha^{(\text{appl})} + \mathbf{F}_\alpha^{(\text{constr})},$$

where $\alpha = 1, 2$, and $\mathbf{F}_\alpha^{(\text{appl})}$ is the sum of the forces acting on the α-th body that are not constraint forces, which are often called the "applied forces," from Equations (1.7) we obtain the two scalar equations

$$\sum_{\alpha=1}^{2} \left(m_\alpha \ddot{\mathbf{r}}_\alpha - \mathbf{F}_\alpha^{(\text{appl})} \right) \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} = 0, \tag{1.8}$$

where $i = 1, 2$. These equations do not contain the constraint forces.

Exercise 1.2. Show that Equations (1.8) amount to the explicit expressions

$$m_1 \ddot{x} + m_2(\ddot{x} + \ddot{y} \cot\theta_0) = 0,$$
$$m_2(\ddot{x} + \ddot{y} \cot\theta_0) \cot\theta_0 + m_2 \ddot{y} + m_2 g = 0$$

and find the values of $\ddot{x}$ and $\ddot{y}$.

In the case of the parametrization (1.1), each pair of real numbers (x, y), with $x \in \mathbb{R}$, $y \in (0, b)$, where b is the maximum value of y for which the block still rests on the wedge (its value depends on the height of the wedge, the length of the block, and the location of its center of mass), represents a *configuration* of this system. The set of all configurations is the *configuration space*; the variables x, y form a coordinate system of the configuration space. (Roughly speaking, a configuration is a snapshot of the mechanical system and, in order to distinguish one snapshot from another, we only have to know the values of the parameters x, y.)

The possibility of expressing the position vectors as in Equations (1.1) and (1.2), in terms of two parameters only, is *equivalent* to the existence of some equations linking the position vectors. Such equations are called *constraint equations*. One can readily verify that Equations (1.1) [or (1.2)] represent the general solution of the constraint equations

$$\mathbf{r}_1 \cdot \mathbf{j} = 0, \qquad (\mathbf{r}_2 - \mathbf{r}_1) \cdot (\sin\theta_0 \mathbf{i} - \cos\theta_0 \mathbf{j}) = 0 \tag{1.9}$$

(assuming that $\mathbf{r}_1$ and $\mathbf{r}_2$ lie on the xy-plane). Any expression for the general solution of the constraint equations amounts to a parametrization of the position vectors. When the configuration of the system is restricted by equations that involve the position vectors and the time only [such as Equations (1.9)], we say that the system has *holonomic constraints* or that the system is *holonomic*.

Example 1.3. A block of mass m can slide freely on a wedge that is moving in a *specified* manner on a horizontal surface; that is, the coordinate X shown in Figure 1.3 is assumed to be a *given* function of the time. The position vector $\mathbf{r}$, which determines the position of the block, can be expressed in the form

$$\mathbf{r} = X\mathbf{i} + s(\cos\theta_0\,\mathbf{i} + \sin\theta_0\,\mathbf{j}) \tag{1.10}$$

in terms of a *single* parameter s [cf. (1.2)]. In this case, by contrast with Example 1.1, there is only one body of interest and only one parameter. The constraint force on the block is the normal force $\mathbf{F}^{(\text{constr})} = N(-\sin\theta_0\,\mathbf{i} + \cos\theta_0\,\mathbf{j})$, exerted by the wedge on the block, and the constraint equation is

$$(\mathbf{r} - X\mathbf{i}) \cdot (-\sin\theta_0\,\mathbf{i} + \cos\theta_0\,\mathbf{j}) = 0.$$

Note that we do not have to specify the mass of the wedge because it moves in a given manner. The configuration of this mechanical system is determined by a single parameter [e.g., the coordinate s in (1.10)] since the position of the wedge at any instant is assumed known.

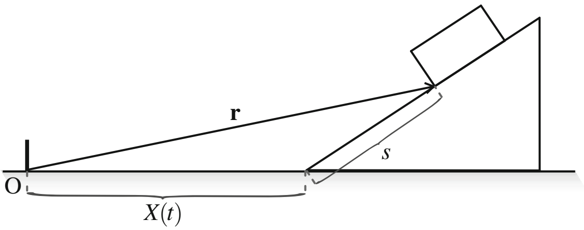


Fig. 1.3 The position of the wedge is determined by a *given* function of time, $X(t)$. The block can slide freely on the wedge under the influence of a uniform gravitational field

The analog of the sum on the left-hand side of (1.7) is

$$\mathbf{F}^{(\text{constr})} \cdot \frac{\partial\mathbf{r}}{\partial s} = N(-\sin\theta_0\,\mathbf{i} + \cos\theta_0\,\mathbf{j}) \cdot (\cos\theta_0\,\mathbf{i} + \sin\theta_0\,\mathbf{j})$$

and is equal to zero. The applied force is the weight of the block, $-mg\mathbf{j}$, and from Newton's second law we have

$$\begin{aligned} 0 &= \left(m\ddot{\mathbf{r}} - \mathbf{F}^{(\text{appl})}\right) \cdot \frac{\partial\mathbf{r}}{\partial s} \\ &= \left\{m[\ddot{X}\mathbf{i} + \ddot{s}(\cos\theta_0\,\mathbf{i} + \sin\theta_0\,\mathbf{j})] + mg\mathbf{j}\right\} \cdot (\cos\theta_0\,\mathbf{i} + \sin\theta_0\,\mathbf{j}) \\ &= m(\ddot{X}\cos\theta_0 + \ddot{s} + g\sin\theta_0). \end{aligned}$$

Thus, in terms of the parameter s, the *equation of motion* is given by the second-order ODE

$$\ddot{s} = -\ddot{X}\cos\theta_0 - g\sin\theta_0.$$

The solution of this equation is $s = -X\cos\theta_0 - \frac{1}{2}g\sin\theta_0\, t^2 + c_1 t + c_2$, where c_1 and c_2 are constants (as usual, determined by the initial conditions).

Example 1.4. As a third example we shall consider a bead of mass m that can slide along a rod which rotates about a vertical axis with constant angular velocity ω. The rod makes a constant angle θ_0 with the vertical axis (see Figure 1.4). The position vector of the bead can be parameterized in the form

$$\mathbf{r} = r\,(\sin\theta_0\cos\omega t\,\mathbf{i} + \sin\theta_0\sin\omega t\,\mathbf{j} + \cos\theta_0\,\mathbf{k}), \tag{1.11}$$

in terms of a single parameter, r. (Note that, by contrast with Equations (1.1) and (1.2), in (1.10) and (1.11) there is an explicit dependence on the time.) The relevant derivatives of $\mathbf{r}$ with respect to the time are

$$\begin{aligned}\dot{\mathbf{r}} &= \dot{r}\,(\sin\theta_0\cos\omega t\,\mathbf{i} + \sin\theta_0\sin\omega t\,\mathbf{j} + \cos\theta_0\,\mathbf{k})\\ &\quad + \omega r\sin\theta_0(-\sin\omega t\,\mathbf{i} + \cos\omega t\,\mathbf{j}), \qquad (1.12)\\ \ddot{\mathbf{r}} &= \ddot{r}\,(\sin\theta_0\cos\omega t\,\mathbf{i} + \sin\theta_0\sin\omega t\,\mathbf{j} + \cos\theta_0\,\mathbf{k})\\ &\quad + 2\omega\dot{r}\sin\theta_0(-\sin\omega t\,\mathbf{i} + \cos\omega t\,\mathbf{j}) - \omega^2 r\sin\theta_0(\cos\omega t\,\mathbf{i} + \sin\omega t\,\mathbf{j}). \qquad (1.13)\end{aligned}$$

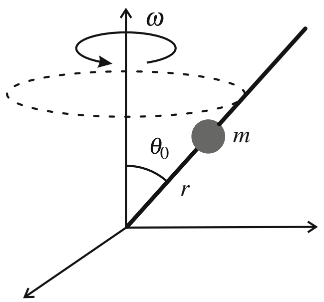


Fig. 1.4 The bead can slide freely along the rod, which rotates with a constant angular velocity ω, forming a constant angle θ_0 with the vertical. The distance r from the bead to the origin can be used as the only coordinate necessary to define the configuration

In the present example, the expression (1.11) is the general solution of the constraint equations

$$\mathbf{r}\cdot(-\sin\omega t\,\mathbf{i} + \cos\omega t\,\mathbf{j}) = 0, \qquad \mathbf{r}\cdot(\cos\theta_0\cos\omega t\,\mathbf{i} + \cos\theta_0\sin\omega t\,\mathbf{j} - \sin\theta_0\,\mathbf{k}) = 0,$$

which involve the time explicitly (the vectors inside the parentheses are two linearly independent vectors orthogonal to the rod).

The forces acting on the bead are its weight (the applied force, $\mathbf{F}^{(\mathrm{appl})} = -mg\mathbf{k}$) and, assuming that there is no friction, a force normal to the rod (the constraint force). Taking into account that, in this case, there is only one position vector, $\mathbf{r}$, which depends on one parameter only, the analog of the sums (1.6) is [making use of (1.11)]

$$\mathbf{F}^{(\text{constr})} \cdot \frac{\partial \mathbf{r}}{\partial r} = \mathbf{F}^{(\text{constr})} \cdot \frac{\mathbf{r}}{r} = 0, \tag{1.14}$$

where the last equality follows from the fact that $\mathbf{F}^{(\text{constr})}$ is normal to the rod and, therefore, perpendicular to the position vector $\mathbf{r}$.

Assuming that $\mathbf{r}$ is the position vector of the bead with respect to an inertial frame, from Newton's second law and (1.14) we have

$$\left(m\ddot{\mathbf{r}} - \mathbf{F}^{(\text{appl})}\right) \cdot \frac{\partial \mathbf{r}}{\partial r} = 0 \tag{1.15}$$

which, with the aid of (1.13) and the relation $\partial \mathbf{r}/\partial r = \mathbf{r}/r$, reduces to

$$m(\ddot{r} - \omega^2 r \sin^2\theta_0) + mg\cos\theta_0 = 0. \tag{1.16}$$

This is the equation of motion for this system, expressed in terms of the parameter r. It is a second-order ODE that can be readily solved because it is linear and the coefficients are constant. However, we can reduce its order as a first step making use of the following standard trick. Multiplying both sides of (1.16) by $\dot{r}$ we find that the result is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\tfrac{1}{2}m\dot{r}^2 - \tfrac{1}{2}m\omega^2 r^2 \sin^2\theta_0 + mgr\cos\theta_0\right) = 0,$$

which means that the expression inside the parenthesis is a *constant of motion* (that is, a function of the parameters, their first time derivatives, and possibly of the time whose value does not change with the time if the equations of motion are satisfied; the constants of motion are also called *first integrals* or *conserved quantities*). Hence,

$$\tfrac{1}{2}m\dot{r}^2 - \tfrac{1}{2}m\omega^2 r^2 \sin^2\theta_0 + mgr\cos\theta_0 = E', \tag{1.17}$$

where E' is a constant (whose value depends on the initial conditions). (E' has dimensions of energy, but it is not the total energy, which is given by [see (1.12)]

$$E = \tfrac{1}{2}m\dot{\mathbf{r}}^2 + mgr\cos\theta_0 = \tfrac{1}{2}m\dot{r}^2 + \tfrac{1}{2}m\omega^2 r^2 \sin^2\theta_0 + mgr\cos\theta_0$$

and, by contrast with (1.17), it is *not* conserved.)

Equation (1.17) is a first-order ODE that can be readily solved separating variables; however, it is convenient to note that, depending on the value of the constant E', Equation (1.17) represents a hyperbola, or a straight line in the $r\dot{r}$-plane. In fact, writing Equation (1.17) in the form

$$\frac{1}{2}m\dot{r}^2 - \frac{1}{2}m\omega^2\sin^2\theta_0\left(r - \frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}\right)^2 = E' - \frac{mg^2\cot^2\theta_0}{2\omega^2}$$

one concludes that, for $E' \neq mg^2 \cot^2\theta_0/2\omega^2$, the so-called *phase curves* in the $r\dot{r}$-plane are hyperbolas with center $(\frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}, 0)$, while for $E' = mg^2\cot^2\theta_0/2\omega^2$, the phase curves are straight lines with slopes $\pm\omega\sin\theta_0$ passing through the point $(\frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}, 0)$ (see Figure 1.5). Actually, each of these straight lines contain three disjoint solution curves corresponding to $r < \frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}$, $r = \frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}$, and $r > \frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}$. The single point $(\frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}, 0)$ is by itself a solution curve, representing an unstable equilibrium point. (The crossing of two phase curves would mean that for the initial conditions represented by the intersection point, the solution of the equations of motion is not unique.) Each point of the $r\dot{r}$-plane (or *phase plane*), with $r > 0$, corresponds to an initial condition for the ODE (1.17) and there is only one solution of this equation passing through that point.

The set of phase curves (called *phase portrait*) is very useful even if we have the explicit solution of the equations of motion; with the aid of the phase portrait one can readily see the qualitative behavior of the motion. For instance, Figure 1.5 shows that for $E' < mg^2\cot^2\theta_0/2\omega^2$, there exist turning points where the sign of the radial velocity is inverted, passing through zero; for $E' > mg^2\cot^2\theta_0/2\omega^2$, the sign of the radial velocity does not change, either the bead is getting away from the origin all the time, or is approaching the origin all the time.

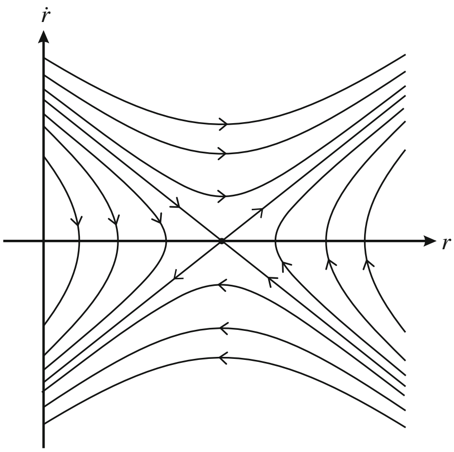


Fig. 1.5 For each value of the constant E', Equation (1.17) defines a curve (a phase curve) in the $r\dot{r}$-plane (the phase plane), which is a hyperbola centered at $(\frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}, 0)$ if $E' \neq mg^2\cot^2\theta_0/2\omega^2$, or part of a straight line with slope $\pm\omega\sin\theta_0$ passing through the point $(\frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}, 0)$ if $E' = mg^2\cot^2\theta_0/2\omega^2$. Each point of the half-plane $r > 0$ corresponds to a possible initial condition and the phase curve passing though this point gives the subsequent states of the bead. The point $(\frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}, 0)$ is a state of unstable equilibrium. The arrows indicate the sense of the time evolution: for points in the upper half-plane $\dot{r}$ is positive, which means that r is an increasing function of time; while for the points in the lower half-plane $\dot{r}$ is negative, which means that r is a decreasing function of time

As pointed out above, the equation of motion (1.16) can be readily solved. Since $g\cos\theta_0/\omega^2\sin^2\theta_0$ is a constant, Equation (1.16) is equivalent to

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(r - \frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}\right) = \omega^2\sin^2\theta_0\left(r - \frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}\right),$$

and by inspection one finds that its general solution can be written as

$$r - \frac{g\cos\theta_0}{\omega^2\sin^2\theta_0} = c_1\mathrm{e}^{\omega t\sin\theta_0} + c_2\mathrm{e}^{-\omega t\sin\theta_0}, \tag{1.18}$$

where c_1, c_2 are constants, which can be related to the initial values of r and $\dot{r}$. In fact, the derivative of Equation (1.18) with respect to the time gives

$$\frac{\dot{r}}{\omega\sin\theta_0} = c_1\mathrm{e}^{\omega t\sin\theta_0} - c_2\mathrm{e}^{-\omega t\sin\theta_0}. \tag{1.19}$$

Setting $t = 0$ in Equations (1.18) and (1.19) we see that

$$c_1 = \frac{1}{2}\left(r_0 - \frac{g\cos\theta_0}{\omega^2\sin^2\theta_0} + \frac{\dot{r}_0}{\omega\sin\theta_0}\right), \qquad c_2 = \frac{1}{2}\left(r_0 - \frac{g\cos\theta_0}{\omega^2\sin^2\theta_0} - \frac{\dot{r}_0}{\omega\sin\theta_0}\right),$$

where r_0 and $\dot{r}_0$ are the values of r and $\dot{r}$ at $t = 0$, respectively.

On the other hand, solving (1.18) and (1.19) for c_1 and c_2 we obtain

$$\begin{aligned} c_1 &= \frac{1}{2}\mathrm{e}^{-\omega t\sin\theta_0}\left(r - \frac{g\cos\theta_0}{\omega^2\sin^2\theta_0} + \frac{\dot{r}}{\omega\sin\theta_0}\right), \\ c_2 &= \frac{1}{2}\mathrm{e}^{\omega t\sin\theta_0}\left(r - \frac{g\cos\theta_0}{\omega^2\sin^2\theta_0} - \frac{\dot{r}}{\omega\sin\theta_0}\right). \end{aligned} \tag{1.20}$$

The expressions on the right-hand sides of (1.20) (which depend on r, $\dot{r}$, and t) are constants of motion, as one can readily verify by calculating their derivatives with respect to the time, making use of (1.16). Furthermore, these two constants of motion are *functionally independent* (that is, one of them cannot be expressed as a function of the other) and *any* constant of motion for this system must be a function of these two constants. For instance, we know that E' is a constant of motion [see (1.17)] and one can verify that

$$E' = -2m\omega^2\sin^2\theta_0 c_1 c_2 + \frac{mg^2\cot^2\theta_0}{2\omega^2}.$$

The solutions with c_1 or c_2 equal to zero correspond to the straight lines of the phase portrait. When $c_1 = c_2 = 0$ we have the equilibrium point $(\frac{g\cos\theta_0}{\omega^2\sin^2\theta_0}, 0)$ of the phase plane. (Equation (1.18) shows that the straight lines of the phase portrait approaching the equilibrium point do not reach that point for finite values of the time and, in fact, the phase curves do not intersect one another.)

This example illustrates an important fact: if the equations of motion are given by n second order ODEs, then there exist $2n$ functionally independent constants of motion. Solving the equations of motion is equivalent to finding $2n$ functionally independent constants of motion.

Example 1.5 (The simple pendulum). A very common example is the simple pendulum which consists of a particle of mass m attached at one end of a massless rigid rod of length l. The rod can only rotate in a fixed vertical plane, with the other end of the rod fixed at a point which is taken as the origin (this point is fixed with respect to some inertial frame). The position vector of the particle can be parameterized as

$$\mathbf{r} = l(\sin\theta\,\mathbf{i} - \cos\theta\,\mathbf{j}), \tag{1.21}$$

where θ is the angle formed by the rod with the downward vertical (see Figure 1.6), then

$$\dot{\mathbf{r}} = l\dot{\theta}(\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}) \tag{1.22}$$

and

$$\ddot{\mathbf{r}} = l\ddot{\theta}(\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}) + l\dot{\theta}^2(-\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j}). \tag{1.23}$$

Equation (1.21) expresses the general solution of the constraint equation $|\mathbf{r}| = l$.

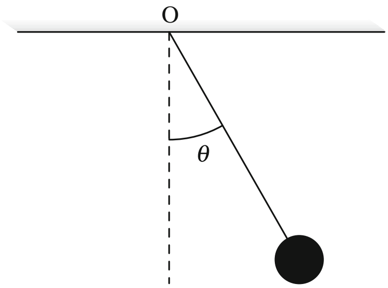


Fig. 1.6 The angle θ between the rod and the downward vertical is used as a coordinate specifying the configuration of the pendulum. It is assumed that the coordinate axes belong to an inertial frame

The forces acting on the pendulum bob are its weight, $-mg\,\mathbf{j}$, and a constraint force, produced by the rod, which maintains the particle on a circle of radius l; this force is directed along the rod and, therefore, we have $\mathbf{F}^{(\text{constr})} = T(-\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j})$, where $|T|$ is the magnitude of the tension of the rod. Here again we have only one particle and only one parameter, or coordinate, θ. Hence, the analog of the sums (1.6) is

$$\mathbf{F}^{(\text{constr})} \cdot \frac{\partial \mathbf{r}}{\partial \theta} = T(-\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j}) \cdot l(\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}) = 0. \tag{1.24}$$

(Alternatively we can make use of the fact that the constraint force is collinear with the position vector, $\mathbf{F}^{(\text{constr})} = -(T/l)\,\mathbf{r}$ and, therefore

$$\mathbf{F}^{(\text{constr})} \cdot \frac{\partial \mathbf{r}}{\partial \theta} = -\frac{T}{l}\mathbf{r} \cdot \frac{\partial \mathbf{r}}{\partial \theta} = -\frac{T}{2l}\frac{\partial}{\partial \theta}\mathbf{r} \cdot \mathbf{r} = -\frac{T}{2l}\frac{\partial}{\partial \theta}l^2 = 0.)$$

Taking into account Equation (1.24), from Newton's second law we get

$$\begin{aligned} 0 &= \left(m\ddot{\mathbf{r}} - \mathbf{F}^{(\text{appl})}\right) \cdot \frac{\partial \mathbf{r}}{\partial \theta} = \left(m\ddot{\mathbf{r}} + mg\mathbf{j}\right) \cdot l(\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}) \\ &= ml(l\ddot{\theta} + g\sin\theta). \end{aligned} \tag{1.25}$$

Thus, in this case, the equation of motion is

$$\ddot{\theta} = -\frac{g}{l}\sin\theta \tag{1.26}$$

but, by contrast with (1.16), this ODE is nonlinear (because of the presence of the function $\sin\theta$). However, as in the case of Equation (1.16), the order of the equation of motion (1.26) can be reduced: multiplying Equation (1.25) by $\dot{\theta}$, the result is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\tfrac{1}{2}ml^2\dot{\theta}^2 - mgl\cos\theta\right) = 0,$$

which means that

$$E \equiv \tfrac{1}{2}ml^2\dot{\theta}^2 - mgl\cos\theta \tag{1.27}$$

is a constant of motion (in this case, the total energy). Some of the phase curves, defined by $\frac{1}{2}ml^2\dot{\theta}^2 - mgl\cos\theta = E$, are shown in Figure 1.7.

Equation (1.27) is a first-order ODE that can be reduced to quadrature. In fact, we obtain

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \pm\sqrt{\frac{2E}{ml^2} + \frac{2g}{l}\cos\theta},$$

which leads to the elliptic integral of the first kind

$$\pm\sqrt{\frac{g}{l}}\int \mathrm{d}t = \int \frac{\mathrm{d}\theta}{\sqrt{\frac{2E}{mgl} + 2\cos\theta}} = \int \frac{\mathrm{d}\theta}{2\sqrt{k^2 - \sin^2(\theta/2)}},$$

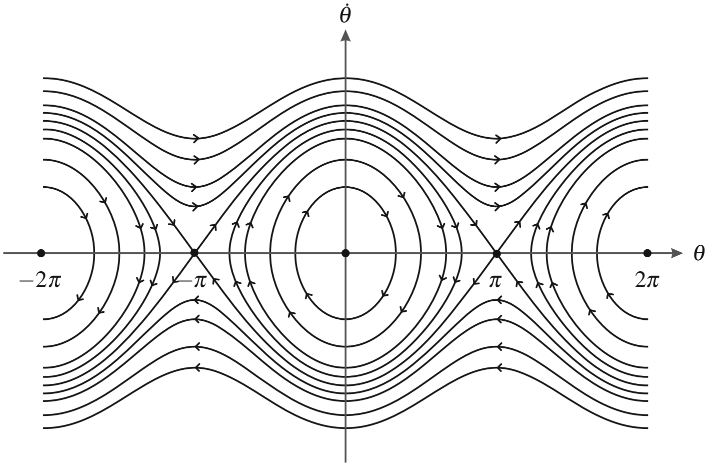


Fig. 1.7 For each value of the constant E, Equation (1.27) defines a phase curve in the $\theta\dot{\theta}$-plane. The curves with $E = mgl$, called *separatrices*, are the boundaries between the closed curves encircling the origin and the curves for which the sign of the velocity does not change. The arrows indicate the sense of the time evolution. The point $(0, 0)$ is a point of stable equilibrium and is the phase curve with $E = -mgl$, while the points $(\pm\pi, 0)$ correspond to a single point of unstable equilibrium. These points also belong to the set $E = mgl$, but are disconnected from the separatrices

where $k^2 \equiv (E + mgl)/2mgl$. (The separatrices in Figure 1.7 correspond to $k = 1$. The time required to go from $\theta = -\pi$ to $\theta = \pi$, or from $\theta = \pi$ to $\theta = -\pi$, along the separatrix is infinite and the phase curves do not intersect at the unstable equilibrium points.) With the change of variable $u = (1/k)\sin(\theta/2)$ one gets the standard form

$$\pm\sqrt{\frac{g}{l}}\int \mathrm{d}t = \int \frac{\mathrm{d}u}{\sqrt{(1-k^2u^2)(1-u^2)}}.$$

The nonlinear equation of motion (1.26) can be replaced by the *linearized* equation

$$\ddot{\theta} = -\frac{g}{l}\theta, \tag{1.28}$$

which is obtained from (1.26) substituting the right-hand side of the equation by its Taylor expansion about the equilibrium point $\theta = 0$, $\dot{\theta} = 0$, keeping only the linear terms in θ and $\dot{\theta}$. By contrast with (1.26), the solution of the linear equation (1.28) can be expressed in terms of elementary functions: $\theta = c_1 \cos\omega t + c_2 \sin\omega t$, with $\omega \equiv \sqrt{g/l}$. This is a periodic function of the time with period $2\pi\sqrt{l/g}$.

Exercise 1.6. A simple pendulum of mass m_2 hangs from a block of mass m_1 which is attached to a spring of stiffness k. The block lies on a horizontal surface without friction. Express the position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ shown in Figure 1.8 in terms of the parameters x and θ, and give the constraint equations. Show that the constraint forces satisfy Equation (1.7), with $q_1 = x$, $q_2 = \theta$, and use this result to show that the equations of motion are

$$\begin{aligned} m_1\ddot{x} + kx + m_2(\ddot{x} + l\ddot{\theta}\cos\theta - l\dot{\theta}^2\sin\theta) &= 0, \\ \ddot{x}\cos\theta + l\ddot{\theta} + g\sin\theta &= 0, \end{aligned} \tag{1.29}$$

assuming that the spring obeys Hooke's law.

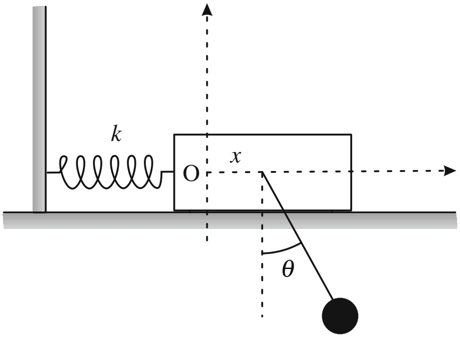


Fig. 1.8 The origin coincides with the suspension point of the pendulum when the spring is relaxed. x is the coordinate of the suspension point of the rod and θ is the angle between the rod and the downward vertical

The equations of motion (1.29) constitute a system of two coupled second-order nonlinear ODEs. For some purposes, it is enough to find its approximate solutions in the vicinity of the stable equilibrium point $x = 0$, $\theta = 0$. The system of equations (1.29) linearized about $x = 0$, $\theta = 0$, $\dot{x} = 0$, $\dot{\theta} = 0$, is

$$\begin{aligned} m_1\ddot{x} + kx + m_2(\ddot{x} + l\ddot{\theta}) &= 0, \\ \ddot{x} + l\ddot{\theta} + g\theta &= 0. \end{aligned} \tag{1.30}$$

By construction, this is a linear system of ODEs with constant coefficients, which can be solved exactly in various ways. An elementary method of solution consists in looking for solutions of the form

$$x = A\cos(\omega t + \delta), \qquad \theta = B\cos(\omega t + \delta), \tag{1.31}$$

where A, B, ω, and δ are constants, that is, x and θ have the *same* time-dependence given by a sinusoidal function with an unknown frequency ω. Substituting (1.31) into (1.30) one obtains the homogeneous system of equations for A and B

$$\begin{aligned} \big[-(m_1+m_2)\omega^2+k\big]A - m_2\omega^2 l B &= 0, \\ -\omega^2 A + (-\omega^2 l + g)B &= 0. \end{aligned} \tag{1.32}$$

This system has a nontrivial solution if and only if the determinant

$$\begin{vmatrix} -(m_1+m_2)\omega^2+k & -m_2\omega^2 l \\ -\omega^2 & -\omega^2 l + g \end{vmatrix}$$

is equal to zero. This condition leads to the quadratic equation

$$m_1 l\omega^4 - \big[kl + (m_1+m_2)g\big]\omega^2 + gk = 0$$

for ω^2. Thus,

$$\omega^2 = \frac{kl + (m_1+m_2)g \pm \sqrt{k^2l^2 + (m_1+m_2)^2g^2 + 2(m_2-m_1)gkl}}{2m_1 l}. \tag{1.33}$$

The radicand can be written as $\big[kl + (m_2-m_1)g\big]^2 + 4m_1m_2g^2$, which cannot be negative. In fact, one can readily see that (1.33) gives two distinct positive values for ω^2. Substituting any of these values of ω^2 into (1.32), one of these equations is a multiple of the other and, therefore, from (1.32) one only obtains a relation between the amplitudes A and B of the oscillatory motions of the block and the pendulum.

The solutions of the form (1.31), where all the bodies of the system move with the same frequency, are called *normal modes of oscillation*, and the corresponding frequencies are called *normal frequencies*. The general solution of (1.30) is a superposition of normal modes of oscillation (see below).

In order to illustrate these concepts, it will be convenient to consider the special case with $m_1 = m_2$ and $kl = mg$, where m is the common value of the masses m_1 and m_2. Then Equation (1.33) reduces to

$$\omega^2 = \frac{3\pm\sqrt{5}}{2}\frac{g}{l}.$$

Substituting this expression into the second equation of (1.32), one gets

$$-(3\pm\sqrt{5})A = (1\pm\sqrt{5})lB.$$

Thus, one of the normal modes of oscillation, is

$$x = A_1\cos(\omega_1 t + \delta_1), \qquad \theta = -\frac{3+\sqrt{5}}{1+\sqrt{5}}\frac{A_1}{l}\cos(\omega_1 t + \delta_1), \tag{1.34}$$

where A_1 and δ_1 are arbitrary constants, with

$$\omega_1{}^2 \equiv \frac{3+\sqrt{5}}{2}\frac{g}{l}.$$

The second normal mode is given by

$$x = A_2 \cos(\omega_2 t + \delta_2), \qquad \theta = \frac{3-\sqrt{5}}{\sqrt{5}-1}\frac{A_2}{l}\cos(\omega_2 t + \delta_2), \tag{1.35}$$

where A_2 and δ_2 are arbitrary constants, and

$$\omega_2{}^2 \equiv \frac{3-\sqrt{5}}{2}\frac{g}{l}.$$

Owing to the linearity of the system (1.30), its general solution is the sum (or superposition) of the solutions (1.34) and (1.35), e.g., for the coordinate x we have $x = A_1 \cos(\omega_1 t + \delta_1) + A_2 \cos(\omega_2 t + \delta_2)$. Note that this general solution contains four arbitrary constants, as would be expected for a system of two second-order ODEs.

Example 1.7. We now consider two particles of masses m_1 and m_2 joined by a rod of length l. The position vector of the first particle can be parameterized by its Cartesian coordinates,

$$\mathbf{r}_1 = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}, \tag{1.36}$$

and with the aid of two additional parameters, θ, ϕ, we can express the position vector of the second particle in the form (see Figure 1.9)

$$\mathbf{r}_2 = \mathbf{r}_1 + l(\sin\theta\cos\phi\,\mathbf{i} + \sin\theta\sin\phi\,\mathbf{j} + \cos\theta\,\mathbf{k}). \tag{1.37}$$

Hence, we need a total of five independent parameters to express the position vectors of the two particles and, therefore, we shall say that this system possesses five *degrees of freedom*. (Again, there are an infinite number of alternative parameterizations, but the number of independent parameters is fixed by the mechanical system.)

The rod can only exert a force (a tension) on the particles, of some unknown magnitude, T, which must be directed along the rod itself. This tension will have the exact magnitude and sense required to maintain the two particles separated by the fixed distance l. If the vector $\mathbf{T}$ denotes the tension acting on m_1, then $-\mathbf{T}$ is the tension on m_2 (that is, $\mathbf{F}_1^{(\text{constr})} = \mathbf{T}$ and $\mathbf{F}_2^{(\text{constr})} = -\mathbf{T}$). Thus, denoting by $q_1, \ldots, q_5$ the five parameters appearing in Equations (1.36)–(1.37) (or any other set of variables employed to parameterize $\mathbf{r}_1$ and $\mathbf{r}_2$),

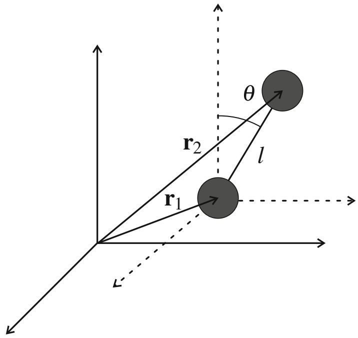


Fig. 1.9 The two particles are held together by a rod of constant length l

$$\sum_{\alpha=1}^{2} \mathbf{F}_{\alpha}^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_{\alpha}}{\partial q_i} = \mathbf{T} \cdot \frac{\partial \mathbf{r}_1}{\partial q_i} + (-\mathbf{T}) \cdot \frac{\partial \mathbf{r}_2}{\partial q_i} = \mathbf{T} \cdot \frac{\partial (\mathbf{r}_1 - \mathbf{r}_2)}{\partial q_i}.$$

Since $\mathbf{T}$ is parallel to $\mathbf{r}_1 - \mathbf{r}_2$ and l is the magnitude of $\mathbf{r}_1 - \mathbf{r}_2$, we can write $\mathbf{T} = (T/l)(\mathbf{r}_1 - \mathbf{r}_2)$, where $|T|$ is the magnitude of the tension, hence,

$$\mathbf{T} \cdot \frac{\partial (\mathbf{r}_1 - \mathbf{r}_2)}{\partial q_i} = \frac{T}{l}(\mathbf{r}_1 - \mathbf{r}_2) \cdot \frac{\partial (\mathbf{r}_1 - \mathbf{r}_2)}{\partial q_i} = \frac{T}{2l}\frac{\partial}{\partial q_i}(\mathbf{r}_1 - \mathbf{r}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2) = \frac{T}{2l}\frac{\partial}{\partial q_i} l^2 = 0.$$

Thus,

$$\sum_{\alpha=1}^{2} \mathbf{F}_{\alpha}^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_{\alpha}}{\partial q_i} = 0$$

for all the values of the index i. (Note that the explicit parametrization of the position vectors, (1.36), (1.37), was not required.)

Exercise 1.8. A bead of mass m can slide freely along a hoop of radius a. The hoop rotates with constant angular velocity ω about a fixed point of the hoop on a horizontal plane (see Figure 1.10). Show that the constraint forces satisfy (1.7) and use this fact to obtain the equations of motion using as generalized coordinate the angle, ϕ, between the diameter of the hoop passing through the fixed point and the radius passing through the bead.

Exercise 1.9. A block of mass m_2 is tied to a wall by means of a rope that passes over a pulley attached to a wedge of mass m_1. The block lies on the wedge, which is free to move on a horizontal surface (see Figure 1.11). Neglect the masses of the rope and the pulley, and the friction forces, show that Equation (1.38) holds, find the equation of motion, and the phase portrait.

Summarizing, in all the examples presented above, we have considered mechanical systems formed by N bodies or particles, whose positions with respect to an inertial frame are determined by N position vectors $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N$, which can be

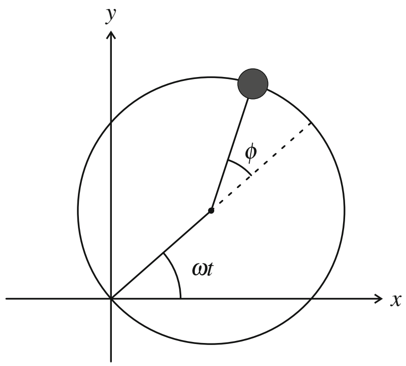


Fig. 1.10 The hoop rotates about the origin with constant angular velocity ω. The hoop is maintained in contact with the horizontal plane $z = 0$

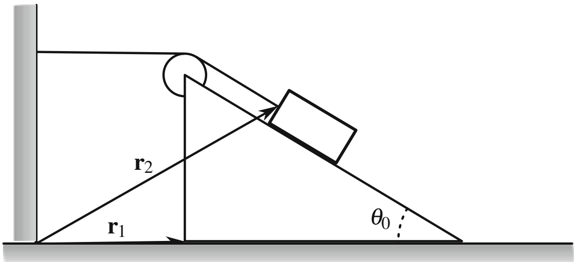


Fig. 1.11 The configuration of this mechanical system can be specified by a single parameter, such as the distance between the wall and the wedge

expressed in terms of a set of independent parameters (or *generalized coordinates*), $q_1, q_2, \ldots, q_n$, and possibly of the time, and in all cases we found that the constraint forces are such that, for $i = 1, 2, \ldots, n$,

$$\sum_{\alpha=1}^{N} \mathbf{F}_\alpha^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} = 0, \tag{1.38}$$

and, therefore, from Newton's second law it follows that

$$\sum_{\alpha=1}^{N} \left(m_\alpha \ddot{\mathbf{r}}_\alpha - \mathbf{F}_\alpha^{(\text{appl})} \right) \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} = 0, \tag{1.39}$$

for $i = 1, 2, \ldots, n$. Usually (as in all the foregoing examples), Equations (1.39) amount to a system of n second-order ODEs for the coordinates q_i as functions of the time. The solution of this system determines all the possible motions of the mechanical system. (Once we have the solution of the equations of motion, the constraint forces can be calculated using again Newton's second law: $\mathbf{F}_\alpha^{(\text{constr})} = m_\alpha \ddot{\mathbf{r}}_\alpha - \mathbf{F}_\alpha^{(\text{appl})}$.)

The number of independent parameters necessary to express the position vectors $\mathbf{r}_\alpha$ is the number of *degrees of freedom* of the system. In what follows we shall assume that the constraint forces satisfy (1.38). Equations (1.38) constitute the *d'Alembert principle*. In the standard approach, the d'Alembert principle is formulated making use of the concepts of virtual work, virtual displacement and, especially, infinitesimal virtual displacement. By formulating the d'Alembert principle in the form (1.38), one avoids the use of all such concepts.

As pointed out at the beginning of this section, there are an infinite number of alternative parameterizations of the position vectors but we can readily see that Equations (1.38) do not depend on the coordinates being employed. In fact, if q'_i is a second set of generalized coordinates, it must be possible to express the q'_i in terms of the q_j (and possibly of the time) and vice versa. Then, by virtue of the chain rule, if Equations (1.38) hold, we have

$$\sum_{\alpha=1}^{N} \mathbf{F}_\alpha^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q'_i} = \left(\sum_{\alpha=1}^{N} \mathbf{F}_\alpha^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_j} \right) \frac{\partial q_j}{\partial q'_i} = 0,$$

where, as in what follows, a repeated lowercase Latin index, $i, j, k, \ldots,$ implies sum over all values of the index.

1.2 The Lagrange Equations

Equations (1.39) follow from Newton's second law, d'Alembert's principle, and the assumption that the position vectors of the bodies forming the system can be written in terms of a set of n independent parameters $q_1, q_2, \ldots, q_n$ and, possibly, of the time

$$\mathbf{r}_\alpha = \mathbf{r}_\alpha(q_1, q_2, \ldots, q_n, t), \qquad \alpha = 1, 2, \ldots, N. \tag{1.40}$$

As pointed out above, this last condition is equivalent to say that the position vectors $\mathbf{r}_\alpha$ satisfy certain constraint equations or that the constraints are holonomic. (Observe that the expressions (1.1), (1.2), (1.10), (1.11), (1.21), (1.36), and (1.37), have the form (1.40).)

In order to simplify the writing, in what follows, a function of the coordinates and the time, $f(q_1, q_2, \ldots, q_n, t)$, will be expressed in the abbreviated form $f(q_i, t)$, where the subscript i indicates the existence of several variables. In this sense, it is equivalent to write $f(q_i, t)$, $f(q_j, t)$, or $f(q_k, t)$. (Other convenient equivalent notations are $f(\mathbf{q}, t)$ and $f(q, t)$.) Thus, the right-hand side of Equation (1.40) will be written as $\mathbf{r}_\alpha(q_i, t)$.

As we have seen in the preceding examples, Equations (1.39) are useful as they stand because they do not contain the constraint forces. However, there is a convenient way of expressing the left-hand side of these equations (which avoids the direct calculation of the second derivatives $\ddot{\mathbf{r}}_\alpha$).

First, by virtue of the chain rule, from (1.40) we have

$$\dot{\mathbf{r}}_\alpha = \frac{\partial \mathbf{r}_\alpha}{\partial q_i} \dot{q}_i + \frac{\partial \mathbf{r}_\alpha}{\partial t}. \tag{1.41}$$

An essential point of the Lagrangian formalism, to be developed in this section, is that, in spite of the way in which we obtained the expression on the right-hand side of (1.41), the $2n + 1$ variables $q_1, q_2, \ldots, q_n, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n, t$ appearing in it are going to be treated as independent. For instance, in Equation (1.12) it makes sense to give values to r, $\dot{r}$, and t, in an arbitrary manner; that is, any combination of values of r, $\dot{r}$, and t (with $r > 0$) yields a possible velocity vector $\dot{\mathbf{r}}$. Following this rule, from (1.41) we obtain

$$\frac{\partial \dot{\mathbf{r}}_\alpha}{\partial \dot{q}_j} = \frac{\partial \mathbf{r}_\alpha}{\partial q_j}, \tag{1.42}$$

for $\alpha = 1, 2, \ldots, N$, and $j = 1, 2, \ldots, n$. In a similar way we have, again by the chain rule, assuming, as in everything that follows, that the partial derivatives commute,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathbf{r}_\alpha}{\partial q_i} = \frac{\partial^2 \mathbf{r}_\alpha}{\partial q_j q_i}\dot{q}_j + \frac{\partial^2 \mathbf{r}_\alpha}{\partial t \partial q_i} = \frac{\partial}{\partial q_i}\left(\frac{\partial \mathbf{r}_\alpha}{\partial q_j}\dot{q}_j + \frac{\partial \mathbf{r}_\alpha}{\partial t}\right) = \frac{\partial \dot{\mathbf{r}}_\alpha}{\partial q_i}, \tag{1.43}$$

thus, considering the first term in (1.39), making use of the Leibniz rule, (1.42) and (1.43),

$$\begin{aligned} m_\alpha \ddot{\mathbf{r}}_\alpha \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(m_\alpha \dot{\mathbf{r}}_\alpha \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i}\right) - m_\alpha \dot{\mathbf{r}}_\alpha \cdot \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathbf{r}_\alpha}{\partial q_i} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left(m_\alpha \dot{\mathbf{r}}_\alpha \cdot \frac{\partial \dot{\mathbf{r}}_\alpha}{\partial \dot{q}_i}\right) - m_\alpha \dot{\mathbf{r}}_\alpha \cdot \frac{\partial \dot{\mathbf{r}}_\alpha}{\partial q_i} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{q}_i}\left(\frac{1}{2}m_\alpha \dot{\mathbf{r}}_\alpha \cdot \dot{\mathbf{r}}_\alpha\right) - \frac{\partial}{\partial q_i}\left(\frac{1}{2}m_\alpha \dot{\mathbf{r}}_\alpha \cdot \dot{\mathbf{r}}_\alpha\right). \end{aligned}$$

Hence, Equations (1.39) can be written in the equivalent form

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = Q_i, \tag{1.44}$$

where

$$T \equiv \sum_{\alpha=1}^{N} \frac{1}{2} m_\alpha \dot{\mathbf{r}}_\alpha \cdot \dot{\mathbf{r}}_\alpha \tag{1.45}$$

is the kinetic energy of the system and

$$Q_i \equiv \sum_{\alpha=1}^{N} \mathbf{F}_\alpha^{(\mathrm{appl})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} \tag{1.46}$$

($i = 1, 2, \ldots, n$) are usually called *generalized forces*. (Since the coordinates q_i may not have dimensions of length, the generalized forces may not have dimensions of force.) (In the foregoing section, T stood for a tension but, in the rest of this chapter, T will denote kinetic energy; there is no risk of confusion because we will not deal with constraint forces henceforth.)

Note that the value of the kinetic energy, T, does not depend on the coordinates being employed [in fact, the coordinates q_i do not appear in the right-hand side of (1.45)], but it does depend on the inertial frame with respect to which the velocities $\dot{\mathbf{r}}_\alpha$ are defined. On the other hand, the Q_i do depend on the coordinates chosen. With the aid of the chain rule, from (1.46), one finds that if the coordinates q_i are replaced by another set of coordinates q'_i, then the new generalized forces are given by

$$Q'_i = \frac{\partial q_j}{\partial q'_i} Q_j. \tag{1.47}$$

Note also that the generalized forces can depend on the time even if the applied forces do not, since the position vectors $\mathbf{r}_\alpha$ appearing in (1.46) may depend explicitly on the time.

In the examples usually considered, the generalized forces are derivable from a potential, that is, there exists a function $V(q_i, t)$ such that

$$Q_i = -\frac{\partial V}{\partial q_i}. \tag{1.48}$$

The validity of (1.48) is highly convenient because instead of the n functions Q_i of the coordinates and the time, it is enough to know the potential V, from which all the generalized forces can be calculated. The existence of the potential V does not depend on the coordinates being employed, and it is therefore a property of the mechanical system under consideration. If the coordinates q_i are replaced by other coordinates $q'_i = q'_i(q_j, t)$, then from (1.47), (1.48), and the chain rule we have

$$Q'_i = -\frac{\partial q_j}{\partial q'_i}\frac{\partial V}{\partial q_j} = -\frac{\partial V}{\partial q'_i}.$$

Note that if a potential exists, it is defined up to an additive function of t only.

Substituting (1.48) into (1.44), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial (T - V)}{\partial q_i} = 0.$$

Taking into account that V is a function of q_i and t only, the last equation can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial (T - V)}{\partial \dot{q}_i} - \frac{\partial (T - V)}{\partial q_i} = 0$$

or, equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \tag{1.49}$$

where

$$L \equiv T - V \tag{1.50}$$

is the *Lagrangian* of the system, and Equations (1.49) are known as the *Lagrange equations*. As we shall see later, there exists an infinite number of functions, L, such that, when substituted into the Lagrange equations lead to the equations of motion of the system being considered (see Proposition 1.25 and Sections 2.3 and 2.4). The expression (1.50) will be called the *natural Lagrangian* or the *standard Lagrangian*.

In order to emphasize the fact that the $q_i, \dot{q}_i$, and t appearing in a Lagrangian are independent variables, the Lagrange equations (1.49) can be written in the more explicit form

$$\frac{\partial^2 L}{\partial t \partial \dot{q}_i} + \frac{\partial^2 L}{\partial q_j \partial \dot{q}_i}\dot{q}_j + \frac{\partial^2 L}{\partial \dot{q}_j \partial \dot{q}_i}\ddot{q}_j - \frac{\partial L}{\partial q_i} = 0 \tag{1.51}$$

which shows that if

$$\det\left(\frac{\partial^2 L}{\partial \dot{q}_j \partial \dot{q}_i}\right) \neq 0 \tag{1.52}$$

then Equations (1.51) constitute a set of $2n$ second-order ODEs for the q_i. (If the condition (1.52) holds, then Equations (1.51) can be inverted to express $\ddot{q}_i$ as functions of $q_j, \dot{q}_j$ and t.) When condition (1.52) is satisfied, which happens in most examples of elementary mechanics, we say that the Lagrangian is regular or non-singular. Otherwise, L is a *singular* Lagrangian.

If the applied forces are conservative, which means that there exists a function $E_P(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ (the potential energy of the system) such that $\mathbf{F}_\alpha^{(\mathrm{appl})} = -\nabla_\alpha E_P$, where ∇_α denotes the gradient with respect to $\mathbf{r}_\alpha$, then from (1.46) and the chain rule we find that

$$Q_i = -\sum_{\alpha=1}^{N} \nabla_\alpha E_P \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} = -\frac{\partial}{\partial q_i} E_P\big(\mathbf{r}_1(q_i, t), \mathbf{r}_2(q_i, t), \ldots, \mathbf{r}_N(q_i, t)\big),$$

which is of the form (1.48), with

$$V(q_i, t) = E_P\big(\mathbf{r}_1(q_i, t), \mathbf{r}_2(q_i, t), \ldots, \mathbf{r}_N(q_i, t)\big). \tag{1.53}$$

Thus, *when the applied forces are conservative, the generalized forces are derivable from a potential, which can be chosen according to (1.53).*

Example 1.10. We shall consider again the system studied in Example 1.4, with the parametrization defined by Equation (1.11). From (1.12), using the fact that the vectors $\sin\theta_0 \cos\omega t\,\mathbf{i} + \sin\theta_0 \sin\omega t\,\mathbf{j} + \cos\theta_0\,\mathbf{k}$, and $-\sin\omega t\,\mathbf{i} + \cos\omega t\,\mathbf{j}$ are orthogonal to each other and have norm equal to 1, we find that the kinetic energy of the particle is given by

$$T = \tfrac{1}{2}m\dot{\mathbf{r}}^2 = \tfrac{1}{2}m(\dot{r}^2 + \omega^2 r^2 \sin^2\theta_0).$$

On the other hand, from (1.46) and (1.11) we find that the only generalized force is

$$Q = \mathbf{F}^{(\text{appl})} \cdot \frac{\partial \mathbf{r}}{\partial r} = -mg\mathbf{k} \cdot (\sin\theta_0 \cos\omega t\,\mathbf{i} + \sin\theta_0 \sin\omega t\,\mathbf{j} + \cos\theta_0\,\mathbf{k}) = -mg\cos\theta_0$$

and is derivable from the potential $V = mgr\cos\theta_0$ [see (1.48)]. (In this case the applied force is the weight of the particle, $-mg\mathbf{k}$, which, as is well known, is conservative; in fact, $-mg\mathbf{k} = -\nabla(mg\mathbf{k}\cdot\mathbf{r})$, that is, the potential energy can be chosen as $E_P(\mathbf{r}) = mg\mathbf{k}\cdot\mathbf{r}$, so that, according to (1.11), $E_P(\mathbf{r}(r,t)) = mgr\cos\theta_0$, in agreement with (1.53).)

Thus, the standard Lagrangian is [see (1.50)]

$$L(r,\dot{r},t) = \tfrac{1}{2}m(\dot{r}^2 + \omega^2 r^2 \sin^2\theta_0) - mgr\cos\theta_0 \tag{1.54}$$

and substituting this expression into (1.49) [or (1.51)] we obtain the single equation

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = \frac{\mathrm{d}}{\mathrm{d}t}m\dot{r} - (m\omega^2 r\sin^2\theta_0 - mg\cos\theta_0) \\ &= m\ddot{r} - m\omega^2 r\sin^2\theta_0 + mg\cos\theta_0, \end{aligned}$$

which coincides with Equation (1.16).

The generalized forces are derivable from a potential also in some cases where the applied forces *are not* conservative. For example, a force field of the form $\mathbf{F}(\mathbf{r}) = \mathbf{k}\times\mathbf{r} = -y\,\mathbf{i} + x\,\mathbf{j}$ is not conservative (if E_P is a function such that $\mathbf{k}\times\mathbf{r} = -\nabla E_P$, then, using Cartesian coordinates, we would have $\partial E_P/\partial x = y$, $\partial E_P/\partial y = -x$, $\partial E_P/\partial z = 0$, but then $\partial^2 E_P/\partial y\partial x = 1$ and $\partial^2 E_P/\partial x\partial y = -1$). However, for a particle restricted to move on the plane $x = a$, where a is some constant, the position vector can be parameterized in the form

$$\mathbf{r} = a\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k},$$

and (identifying $q_1 = y$, $q_2 = z$) we have [see (1.46)]

$$Q_1 = \mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial y} = (-y\,\mathbf{i} + a\,\mathbf{j}) \cdot \mathbf{j} = a, \qquad Q_2 = \mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial z} = (-y\,\mathbf{i} + a\,\mathbf{j}) \cdot \mathbf{k} = 0,$$

which can be written in the form (1.48) with, e.g., $V = -ay$.

If the generalized forces $Q_i(q_j, t)$ are derivable from a potential, then

$$\frac{\partial Q_i}{\partial q_j} = \frac{\partial Q_j}{\partial q_i}, \tag{1.55}$$

for $i, j = 1, 2, \ldots, n$. Indeed, if $Q_i = -\partial V/\partial q_i$, then

$$\frac{\partial Q_i}{\partial q_j} = -\frac{\partial^2 V}{\partial q_j \partial q_i} = -\frac{\partial^2 V}{\partial q_i \partial q_j} = \frac{\partial Q_j}{\partial q_i}.$$

The converse is *locally* true, that is, if the conditions (1.55) are satisfied, then, for each point, P, in the domain of the functions Q_i, there exists a function V defined in some neighborhood of P, such that $Q_i = -\partial V/\partial q_i$. The proof of this assertion is given in any standard textbook on advanced calculus in several variables. Condition (1.55) is the well-known criterion for the exactness of the differential form $Q_1 \mathrm{d}q_1 + Q_2 \mathrm{d}q_2 + \cdots + Q_n \mathrm{d}q_n$ encountered, e.g., in the theory of differential equations.

An example usually employed to show that condition (1.55) only guarantees the local existence of a potential is given by the functions

$$Q_1 = -\frac{q_2}{q_1{}^2 + q_2{}^2}, \qquad Q_2 = \frac{q_1}{q_1{}^2 + q_2{}^2},$$

which are defined (and differentiable) in the open set $D \equiv \mathbb{R}^2 \setminus \{(0,0)\}$. One readily verifies that

$$\frac{\partial Q_1}{\partial q_2} = \frac{q_2{}^2 - q_1{}^2}{(q_1{}^2 + q_2{}^2)^2} = \frac{\partial Q_2}{\partial q_1},$$

however, there does not exist a function V, defined on all of D, such that $Q_i = -\partial V/\partial q_i$, as can be seen by evaluating the line integral

$$I \equiv \oint_C (Q_1 \mathrm{d}q_1 + Q_2 \mathrm{d}q_2)$$

on a closed curve, C, enclosing the point $(0, 0)$. For instance, if C is the curve given by $(q_1(t), q_2(t)) = (\cos t, \sin t)$ with $t \in [0, 2\pi]$ (that is, C is a circle of radius 1 centered at the origin in the $q_1 q_2$-plane), one readily finds that

$$I = \oint \frac{-q_2 \mathrm{d}q_1 + q_1 \mathrm{d}q_2}{q_1{}^2 + q_2{}^2} = \int_0^{2\pi} [-\sin t(-\sin t\,\mathrm{d}t) + \cos t(\cos t\,\mathrm{d}t)] = 2\pi.$$

On the other hand, the existence of a function V such that $Q_i = -\partial V/\partial q_i$ would imply that

$$I = -\oint_C \mathrm{d}V = -\int_0^{2\pi} \frac{\mathrm{d}[V(\cos t, \sin t)]}{\mathrm{d}t}\,\mathrm{d}t = -[V(1,0) - V(1,0)] = 0,$$

which contradicts the previous result. It may be noticed that the function $V = \arctan(q_2/q_1)$, which is well defined for $q_1 > 0$, satisfies the conditions

$$\frac{\partial V}{\partial q_1} = -\frac{q_2}{{q_1}^2 + {q_2}^2}, \qquad \frac{\partial V}{\partial q_2} = \frac{q_1}{{q_1}^2 + {q_2}^2},$$

but $\arctan(q_2/q_1)$ is not a single-valued, continuous function defined in all the region D.

When $n = 1$ the conditions (1.55) are trivially satisfied, regardless of the form of the applied forces. That is, for any system with one degree of freedom, the generalized force is always derivable from a potential (provided that it is a function of q and t only).

Example 1.11. We shall consider the problem treated in Example 1.1, this time employing the Lagrangian formalism. In order to facilitate the comparison with the results of Example 1.1, we shall make use of the parametrization (1.1). We find that the kinetic energy of the bodies of the system is given by

$$\begin{aligned} T &= \tfrac{1}{2}m_1\dot{\mathbf{r}}_1{}^2 + \tfrac{1}{2}m_2\dot{\mathbf{r}}_2{}^2 \\ &= \tfrac{1}{2}m_1(\dot{x}\,\mathbf{i})^2 + \tfrac{1}{2}m_2[(\dot{x} + \dot{y}\,\cot\theta_0)\,\mathbf{i} + \dot{y}\,\mathbf{j}]^2 \\ &= \tfrac{1}{2}m_1\dot{x}^2 + \tfrac{1}{2}m_2[(\dot{x} + \dot{y}\,\cot\theta_0)^2 + \dot{y}^2] \\ &= \tfrac{1}{2}[(m_1 + m_2)\dot{x}^2 + 2m_2\cot\theta_0\dot{x}\dot{y} + m_2\csc^2\theta_0\dot{y}^2] \end{aligned}$$

(a homogeneous function of the $\dot{q}_i$ of degree two).

On the other hand, according to the definition of the generalized forces (1.46), with $q_1 = x$, $q_2 = y$, we have

$$Q_1 \equiv \sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\mathrm{appl})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial x} = (-m_1 g\mathbf{j}) \cdot \mathbf{i} + (-m_2 g\mathbf{j}) \cdot \mathbf{i} = 0,$$

and

$$Q_2 \equiv \sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\mathrm{appl})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial y} = (-m_1 g\mathbf{j}) \cdot \mathbf{0} + (-m_2 g\mathbf{j}) \cdot (\cot\theta_0\mathbf{i} + \mathbf{j}) = -m_2 g,$$

showing that these forces are derivable from the potential $V = m_2 g y$. Hence, the standard Lagrangian for this problem, in these coordinates, is

$$L = \tfrac{1}{2}\left[(m_1 + m_2)\dot{x}^2 + 2m_2 \cot\theta_0 \dot{x}\dot{y} + m_2 \csc^2\theta_0 \dot{y}^2\right] - m_2 g y. \tag{1.56}$$

(Alternatively, taking into account that the applied forces are gravitational and that the potential energy is given by $E_P(\mathbf{r}_1, \mathbf{r}_2) = m_1 g\mathbf{j} \cdot \mathbf{r}_1 + m_2 g\mathbf{j} \cdot \mathbf{r}_2$, from Equations (1.53) and (1.1) we obtain $V(x, y) = m_1 g\mathbf{j} \cdot x\,\mathbf{i} + m_2 g\mathbf{j} \cdot [(x + y\cot\theta_0)\,\mathbf{i} + y\,\mathbf{j}] = m_2 g y$.)

The Lagrange equations give the equations of motion

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = \frac{\mathrm{d}}{\mathrm{d}t}[(m_1 + m_2)\dot{x} + m_2 \cot\theta_0 \dot{y}] = (m_1 + m_2)\ddot{x} + m_2 \cot\theta_0 \ddot{y}$$

and

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = \frac{\mathrm{d}}{\mathrm{d}t}[m_2 \cot\theta_0 \dot{x} + m_2 \csc^2\theta_0 \dot{y}] + m_2 g \\ &= m_2 \cot\theta_0 \ddot{x} + m_2 \csc^2\theta_0 \ddot{y} + m_2 g, \end{aligned}$$

which are equivalent to the results given in Exercise 1.2.

Exercise 1.12. Find the equation of motion for the system considered in Exercise 1.6, making use of the Lagrange formalism.

Exercise 1.13. Solve Exercise 1.8 making use of the Lagrange equations.

Exercise 1.14. Find the standard Lagrangian for the system considered in Exercise 1.9 and use it to obtain the equation of motion.

Exercise 1.15. A right circular cylinder of uniform density, with mass m and radius a, rolls without slipping inside a hollow cylinder of radius b, under the influence of a uniform gravitational field (see Figure 1.12). Find the period of the small oscillations. (From elementary mechanics, it is known that the kinetic energy of a rigid body rotating about a fixed axis is equal to $\frac{1}{2}I\omega^2$, where I is the moment of inertia of the body about that axis and ω is the angular velocity. Keep in mind that "fixed axis" means fixed with respect to an inertial frame and that the angular velocity must be measured with respect to an inertial frame.)

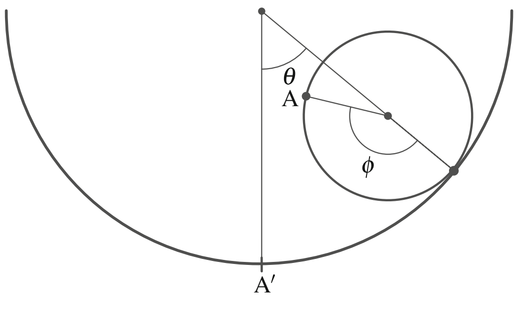


Fig. 1.12 The cylinder, of radius a, rolls without slipping inside the cylindrical surface of radius b. The points A and A′ coincide when the cylinder is at the equilibrium position

Example 1.16 (One-dimensional harmonic oscillator). A very simple but important example, which we will encounter many times in the following chapters, is that of the one-dimensional harmonic oscillator. This system consists of a body of mass m that can move along a straight line, subject to a force directed to a point, O, fixed with respect to an inertial frame. The magnitude of the force is proportional to the distance between the body and O. Hence, if q is a Cartesian coordinate along the line of the motion, with the origin at O, the force has the form $-kq$, where k is a positive constant, and this force is derivable from the potential $V = \frac{1}{2}kq^2$ [see (1.48)]. It is customary to express the constant k as $m\omega^2$, so that the standard Lagrangian is given by

$$L = \tfrac{1}{2}m\dot{q}^2 - \tfrac{1}{2}m\omega^2 q^2. \tag{1.57}$$

Then from (1.49) we obtain the only equation

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = \frac{\mathrm{d}}{\mathrm{d}t}(m\dot{q}) + m\omega^2 q = m\ddot{q} + m\omega^2 q,$$

whose solution is $q = c_1 \cos\omega t + c_2 \sin\omega t$, where c_1 and c_2 are arbitrary constants [cf. Equation (1.28)]. This shows that the constant ω introduced above represents the angular frequency of the motion.

Generalized Potentials. The Electromagnetic Force

Going back to the equations of motion of a holonomic system that obeys the d'Alembert principle expressed in the form (1.44), we see that these equations can be written in the form (1.49), with $L = T - V$, if and only if the generalized forces can be expressed as

$$Q_i = -\frac{\partial V}{\partial q_i} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial V}{\partial \dot{q}_i}, \tag{1.58}$$

for $i = 1, 2, \ldots, n$, in terms of some function V, which we will *assume* that is a function of $q_i, \dot{q}_i$ and t only [cf. Equation (1.48)]. When V depends on the $\dot{q}_i$, it is said that V is a *generalized potential* or a *velocity-dependent potential* (the last name seems more convenient in view of the proliferation of the adjective "generalized").

The total derivative $\mathrm{d}/\mathrm{d}t$ appearing in (1.58) is just an abbreviation. We have to keep in mind that here the variables $q_i, \dot{q}_i$ and t are independent. Writing (1.58) explicitly we have

$$Q_i = -\frac{\partial V}{\partial q_i} + \frac{\partial^2 V}{\partial t \partial \dot{q}_i} + \frac{\partial^2 V}{\partial q_j \partial \dot{q}_i}\dot{q}_j + \frac{\partial^2 V}{\partial \dot{q}_j \partial \dot{q}_i}\ddot{q}_j. \tag{1.59}$$

If, additionally, we assume that the components Q_i are functions of $(q_j, \dot{q}_j, t)$ only, then the second partial derivatives of V with respect to the $\dot{q}_i$ must be equal to zero, and this means that V must be of the form

$$V = \alpha_k \dot{q}_k + \beta, \tag{1.60}$$

where the α_k and β are $n+1$ functions of the coordinates q_i and t only. Substituting (1.60) into (1.59) we obtain

$$\begin{aligned} Q_i &= -\frac{\partial \alpha_j}{\partial q_i}\dot{q}_j - \frac{\partial \beta}{\partial q_i} + \frac{\partial \alpha_i}{\partial t} + \frac{\partial \alpha_i}{\partial q_j}\dot{q}_j \\ &= \left(\frac{\partial \alpha_i}{\partial q_j} - \frac{\partial \alpha_j}{\partial q_i}\right)\dot{q}_j - \frac{\partial \beta}{\partial q_i} + \frac{\partial \alpha_i}{\partial t}, \end{aligned} \tag{1.61}$$

that is, the generalized forces must be functions of the first degree in the velocities.

Exercise 1.17. Show that, given a set of functions $Q_i(q_j, \dot{q}_j, t)$, there exist (locally) functions $\alpha_i(q_j, t)$ and $\beta(q_j, t)$ such that Equations (1.61) hold if and only if the functions Q_i satisfy the conditions

$$\begin{gathered} \frac{\partial^2 Q_i}{\partial \dot{q}_j \partial \dot{q}_k} = 0, \qquad \frac{\partial Q_i}{\partial \dot{q}_j} = -\frac{\partial Q_j}{\partial \dot{q}_i}, \\ \frac{\partial^2 Q_i}{\partial q_k \partial \dot{q}_j} + \frac{\partial^2 Q_j}{\partial q_i \partial \dot{q}_k} + \frac{\partial^2 Q_k}{\partial q_j \partial \dot{q}_i} = 0, \\ \frac{\partial Q_i}{\partial q_j} - \frac{\partial Q_j}{\partial q_i} = \frac{\partial^2 Q_i}{\partial t \partial \dot{q}_j} + \frac{\partial^2 Q_i}{\partial q_k \partial \dot{q}_j}\dot{q}_k. \end{gathered} \tag{1.62}$$

Surprisingly, there exists a force satisfying conditions (1.62), with $n = 3$. The force on a charged particle, with electric charge e, in an electromagnetic field (using Gaussian cgs units) is given by

$$\mathbf{F} = e(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}), \tag{1.63}$$

where $\mathbf{v}$ is the velocity of the particle (as usual, with respect to some inertial frame) and c is the speed of light in vacuum. One can verify using, e.g., Cartesian coordinates that the components of the force (1.63) satisfy conditions (1.62) if and only if

$$\nabla \cdot \mathbf{B} = 0, \qquad \nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}. \tag{1.64}$$

These equations are two of the Maxwell equations and, in classical electrodynamics, it is assumed that they hold everywhere (in the absence of magnetic monopoles). Hence, the force (1.63), known as the *Lorentz force*, can be expressed in terms of a generalized potential. In fact, Equations (1.64) are locally equivalent to the existence of the electromagnetic potentials, φ and $\mathbf{A}$, such that

$$\mathbf{E} = -\nabla\varphi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \qquad \mathbf{B} = \nabla \times \mathbf{A}. \tag{1.65}$$

(It should be kept in mind that, for given electromagnetic fields **E** and **B**, the potentials φ and **A** are not unique. Given **E** and **B**, the potentials are defined up to a gauge transformation:

$$\mathbf{A} \mapsto \mathbf{A} + \nabla\xi, \qquad \varphi \mapsto \varphi - \frac{1}{c}\frac{\partial\xi}{\partial t}, \tag{1.66}$$

where ξ is an arbitrary function.) Substituting Equations (1.65) into (1.63) and comparing the Cartesian components of **F** with (1.61), one finds that one can take

$$\alpha_i = -\frac{e}{c}A_i, \qquad \beta = e\varphi, \tag{1.67}$$

$i = 1, 2, 3$, where the A_i are the Cartesian components of **A**, and, therefore,

$$V = -\frac{e}{c}A_i\dot{q}_i + e\varphi = -\frac{e}{c}\mathbf{A}\cdot\mathbf{v} + e\varphi. \tag{1.68}$$

Thus, the equations of motion for a charged particle of mass m and electric charge e, in an electromagnetic field described by the electromagnetic potentials φ, **A**, can be obtained from the Lagrangian

$$L = \frac{1}{2}m\mathbf{v}^2 + \frac{e}{c}\mathbf{A}\cdot\mathbf{v} - e\varphi. \tag{1.69}$$

As we shall see in the next paragraphs, even though the expression (1.68) was derived making use of the Cartesian coordinates, Equations (1.68) and (1.69) are valid in any coordinate system.

It may be pointed out that, in the preceding pages, the Lagrange equations were introduced as a convenient way for obtaining the equations of motion, but in the case of a charged particle in a given electromagnetic field we already know the equations of motion and, therefore, it might seem pointless to look for the Lagrangian in this case; however, as we shall see in the next few pages, the Lagrangian itself helps to find the solution of the equations of motion, after all, in most cases one may be more interested in the solution of the equations of motion than in the equations of motion themselves.

Coordinate Transformations. Covariance of the Lagrange Equations

Sometimes, after having obtained the expression of the Lagrangian using a set of coordinates q_i, we may want to use another coordinate system, $q'_i = q'_i(q_j, t)$. For instance, in many cases, the Lagrangian is known in Cartesian coordinates, and it may seem more convenient to employ another set of generalized coordinates (see, e.g., Examples 1.18 and 1.21, below). As we shall see, a Lagrangian adequate for the new coordinates is given by the original Lagrangian with the coordinates and velocities, q_i and $\dot{q}_i$, replaced by their equivalent expressions in terms of q'_i and $\dot{q}'_i$.

More precisely, we shall demonstrate that if we have a system of ODEs for the variables q_i as functions of t, given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \tag{1.70}$$

where L is a given function of $(q_i, \dot{q}_i, t)$, and we consider a new set of coordinates q'_i which are functions of the q_i and t, then the ODEs (1.70) expressed in terms of the q'_i are equivalent to the Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}'_i} - \frac{\partial L}{\partial q'_i} = 0, \tag{1.71}$$

where L is the Lagrangian appearing in (1.70) written in terms of $(q'_i, \dot{q}'_i, t)$.

Indeed, assuming that the new coordinates are given by some expressions of the form $q'_i = q'_i(q_j, t)$, the chain rule gives

$$\dot{q}'_i = \frac{\partial q'_i}{\partial q_j}\dot{q}_j + \frac{\partial q'_i}{\partial t} \tag{1.72}$$

and from this last equation (taking into account that the partial derivatives $\partial q'_i/\partial q_j$ are functions of the q_i and t only) we obtain

$$\frac{\partial \dot{q}'_i}{\partial \dot{q}_k} = \frac{\partial q'_i}{\partial q_k}. \tag{1.73}$$

Similarly, the chain rule and (1.72) yield

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial q'_j}{\partial q_i} = \frac{\partial^2 q'_j}{\partial q_k \partial q_i}\dot{q}_k + \frac{\partial^2 q'_j}{\partial t \partial q_i} = \frac{\partial}{\partial q_i}\left(\frac{\partial q'_j}{\partial q_k}\dot{q}_k + \frac{\partial q'_j}{\partial t}\right) = \frac{\partial \dot{q}'_j}{\partial q_i} \tag{1.74}$$

where we treat $q_i, \dot{q}_i$ and t as independent variables [cf. Equations (1.41)–(1.43)]. Then, making use repeatedly of the chain rule, (1.73), (1.74), and the Leibniz rule we obtain the identity (valid for any function L, regardless of its meaning)

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}'_j}\frac{\partial \dot{q}'_j}{\partial \dot{q}_i}\right) - \left(\frac{\partial L}{\partial \dot{q}'_j}\frac{\partial \dot{q}'_j}{\partial q_i} + \frac{\partial L}{\partial q'_j}\frac{\partial q'_j}{\partial q_i}\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}'_j}\frac{\partial q'_j}{\partial q_i}\right) - \frac{\partial L}{\partial \dot{q}'_j}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial q'_j}{\partial q_i} - \frac{\partial L}{\partial q'_j}\frac{\partial q'_j}{\partial q_i} \\
&= \frac{\partial q'_j}{\partial q_i}\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}'_j} - \frac{\partial L}{\partial q'_j}\right).
\end{aligned} \tag{1.75}$$

Since the matrix $(\partial q'_j/\partial q_i)$ must be invertible, it follows that Equations (1.70) hold if and only if Equations (1.71) do. Another proof of the equivalence of Equations (1.70) and (1.71) will be given in Section 2.4.

Example 1.18 (Charged particle in a uniform magnetic field). Making use of the preceding results, we can study the motion of a charged particle of mass m and electric charge e in a uniform magnetic field. Choosing a set of Cartesian axes in such a way that the z-axis coincides with the direction of the magnetic field, the magnetic field is given by $\mathbf{B} = B_0\mathbf{k}$, where B_0 is a constant. One can verify that a vector potential for this magnetic field is $\mathbf{A} = \frac{1}{2}B_0(-y\,\mathbf{i} + x\,\mathbf{j})$. Hence, according to (1.69), the standard Lagrangian for this particle, written in Cartesian coordinates, is

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \frac{eB_0}{2c}(x\dot{y} - y\dot{x}). \tag{1.76}$$

Besides the Cartesian coordinates, the circular cylindrical coordinates, (ρ, ϕ, z), are also convenient to solve this problem. Making use of the standard formulas relating the cylindrical coordinates with the Cartesian ones,

$$x = \rho\cos\phi, \qquad y = \rho\sin\phi, \qquad z = z,$$

one finds that

$$\dot{x} = \cos\phi\,\dot{\rho} - \rho\sin\phi\,\dot{\phi}, \qquad \dot{y} = \sin\phi\,\dot{\rho} + \rho\cos\phi\,\dot{\phi}, \qquad \dot{z} = \dot{z},$$

and substituting these expressions into (1.76) we obtain

$$L = \frac{m}{2}(\dot{\rho}^2 + \rho^2\dot{\phi}^2 + \dot{z}^2) + \frac{eB_0}{2c}\rho^2\dot{\phi}. \tag{1.77}$$

Substituting this Lagrangian into the Lagrange equations we can obtain the equations of motion expressed in terms of the cylindrical coordinates of the particle. However, as will be shown in Exercise 1.23, below, it is neither necessary nor convenient to do such a substitution.

As we shall see in Section 2.4, the Lagrange equations also maintain their form if the coordinates and the time are simultaneously transformed [that is, the coordinates q_i and t are replaced by new variables $q'_i = q'_i(q_j, t)$, $t' = t'(q_j, t)$], provided that the Lagrangian is suitably transformed too.

Ignorable Coordinates. Invariance of the Lagrangian

One of the advantages of expressing a set of ODEs in the form of the Lagrange equations (1.49) is that with the aid of the Lagrangian, in some cases, one can solve these equations partially or totally more easily than with a direct approach. Some of such simplifications occur when the Lagrangian of interest does not contain one of the coordinates q_i, or the time.

If $L(q_i, \dot{q}_i, t)$ does not depend on one of the generalized coordinates, q_k, say, then from the Lagrange equations (1.49) we see that the *generalized momentum* conjugate to q_k,

$$p_k \equiv \frac{\partial L}{\partial \dot{q}_k}, \tag{1.78}$$

is a constant of motion:

$$\frac{\mathrm{d}p_k}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_k} = \frac{\partial L}{\partial q_k} = 0.$$

When $L(q_i, \dot{q}_i, t)$ does not contain the coordinate q_k it is said that q_k is an *ignorable*, or *cyclic*, coordinate.

For instance, the Lagrangian (1.56) does not contain the coordinate x and, therefore, the generalized momentum conjugate to x,

$$p_x \equiv \frac{\partial L}{\partial \dot{x}} = (m_1 + m_2)\dot{x} + m_2 \cot\theta_0 \dot{y},$$

is a constant of motion. In this case, p_x is the x-component of the usual linear momentum of the entire system.

It may be noticed that a Lagrangian can have more than one ignorable coordinate; the ultimate example is that of a free particle in Cartesian coordinates, for which the standard Lagrangian is

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2). \tag{1.79}$$

All the coordinates are ignorable and their conjugate momenta ($m\dot{x}$, $m\dot{y}$, and $m\dot{z}$) are constants of motion. It should be stressed that the existence of ignorable coordinates depends not only on the mechanical system under consideration, but also on the coordinates being employed. For instance, the Lagrangian (1.79), expressed in spherical coordinates, (r, θ, ϕ), has the form

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta\dot{\theta}^2), \tag{1.80}$$

and only the coordinate ϕ is ignorable.

The existence of an ignorable coordinate, q_k, is equivalent to the *invariance* of the Lagrangian under the coordinate transformation

$$q'_k = q_k + s, \qquad \text{and} \qquad q'_i = q_i, \quad \text{for } i \neq k, \tag{1.81}$$

where s is a parameter that *does not* depend on the coordinates or the time, so that all the generalized velocities, $\dot{q}_i$, are left unchanged by the transformation (1.81).

More specifically, this invariance of the Lagrangian means that its value does not change when we perform the coordinate transformation (1.81), that is,

$$L(q_1, \ldots, q_{k-1}, q_k, q_{k+1}, \ldots, q_n, \dot{q}_i, t) = L(q_1, \ldots, q_{k-1}, q_k + s, q_{k+1}, \ldots, q_n, \dot{q}_i, t), \tag{1.82}$$

for all values of s in some neighborhood of 0. (We have already found examples where the coordinates cannot take all real values, e.g., in (1.54), the coordinate r has to be greater than zero.) In fact, recalling the definition of the partial derivative of a function, (1.82) implies that $\partial L/\partial q_k = 0$ and, conversely, if $\partial L/\partial q_k = 0$, then $L(q_i, \dot{q}_i, t)$ does not depend on q_k and, therefore, Equation (1.82) holds.

For instance, the Lagrangian (1.56) is invariant under the translation defined by $x' = x + s$, $y' = y$, which geometrically amounts to the displacement of the origin by a distance s to the left. This invariance corresponds to the fact that we can choose the origin at any point of the x-axis, and this choice does not affect the expression (1.56). Here we are considering the translation $x' = x + s$ as a *passive* transformation, that is, we are translating the coordinate system by displacing the origin, leaving the mechanical system untouched. An *active* translation would mean that the coordinate system is left invariant, while the mechanical system is translated as a whole. (The concept of invariance is treated in more detail in Section 2.5.)

Actually, we can obtain a constant of motion under a weaker condition. If there exists a function $G(q_i, t)$ such that

$$\frac{\partial L}{\partial q_k} = \frac{\partial G}{\partial q_i}\dot{q}_i + \frac{\partial G}{\partial t} \tag{1.83}$$

(with implicit sum over the repeated index i), then $p_k - G$ is a constant of motion. Indeed, making use of the Lagrange equations (1.49), the chain rule, and (1.83), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(p_k - G) = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_k} - \frac{\mathrm{d}G}{\mathrm{d}t} = \frac{\partial L}{\partial q_k} - \frac{\partial G}{\partial q_i}\dot{q}_i - \frac{\partial G}{\partial t} = 0.$$

A first simple example is given again by the Lagrangian (1.56); the coordinate y is not ignorable but $\partial L/\partial y = -m_2 g$, which has the form (1.83) with $G = -m_2 g t$, hence $p_y + m_2 g t$ is a constant of motion.

Example 1.19 (Charged particle in a uniform magnetic field). As shown in Example 1.18, the standard Lagrangian for a charged particle of mass m and electric charge e in a uniform magnetic field $\mathbf{B} = B_0\mathbf{k}$, in Cartesian coordinates, can be taken as

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \frac{eB_0}{2c}(x\dot{y} - y\dot{x}). \tag{1.84}$$

Clearly, z is an ignorable coordinate and therefore its conjugate momentum is a constant of motion

$$p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z} \tag{1.85}$$

(the z-component of the usual linear momentum of the particle). On the other hand, $\partial L/\partial x \neq 0$ (that is, x is not ignorable), but

$$\frac{\partial L}{\partial x} = \frac{eB_0}{2c}\dot{y},$$

which is of the form (1.83) with $G = eB_0 y/2c$. Hence, $p_x - eB_0 y/2c$ is conserved, that is,

$$m\dot{x} - \frac{eB_0}{c}y \tag{1.86}$$

is also a constant of motion. Similarly, y is not ignorable, but

$$\frac{\partial L}{\partial y} = -\frac{eB_0}{2c}\dot{x},$$

which is also of the form (1.83) with $G = -eB_0 x/2c$ and, therefore, $p_y + eB_0 x/2c$ is conserved, that is,

$$m\dot{y} + \frac{eB_0}{c}x \tag{1.87}$$

is a constant of motion.

As pointed out above, for a system with n degrees of freedom and a regular Lagrangian [such as (1.84)], the solution of the equations of motion is equivalent to finding $2n$ functionally independent constants of motion. In the present example, we have already identified three (functionally independent) constants of motion, which are useful to obtain the general solution of the equations of motion (which we have not written down, and we will not require). The conservation of the functions (1.86) and (1.87) can be expressed in the form

$$m\dot{x} = \frac{eB_0}{c}(y - y_0), \qquad m\dot{y} = -\frac{eB_0}{c}(x - x_0),$$

where x_0 and y_0 are two constants, respectively, or, equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}t}(x - x_0) = \omega_c(y - y_0), \qquad \frac{\mathrm{d}}{\mathrm{d}t}(y - y_0) = -\omega_c(x - x_0),$$

with $\omega_c \equiv eB_0/mc$. The solution of this linear system of first-order ODEs is readily found to be

$$x - x_0 = c_1 \cos\omega_c t + c_2 \sin\omega_c t, \qquad y - y_0 = -c_1 \sin\omega_c t + c_2 \cos\omega_c t,$$

where c_1 and c_2 are two additional constants. These last two equations show that in the xy-plane, the particle describes a circle centered at (x_0, y_0) with radius $\sqrt{c_1{}^2 + c_2{}^2}$, and angular velocity ω_c. (The angular frequency ω_c is known as the *cyclotron frequency*.)

The motion along the z-axis (the direction of the magnetic field) is rectilinear and uniform: from (1.85), using the fact that p_z is constant, we find that $z = p_z t/m + z_0$, where z_0 is another constant (making a total of six independent constants, namely p_z, x_0, y_0, c_1, c_2, and z_0).

It may be remarked that the magnetic field $\mathbf{B} = B_0\mathbf{k}$ can also be obtained from the vector potentials $\mathbf{A} = -B_0 y\,\mathbf{i}$, or $\mathbf{A} = B_0 x\,\mathbf{j}$ (among an infinite number of other possibilities); with the first of these, x is ignorable and with the second one, y is ignorable, but there is no choice of the vector potential for which all the coordinates are ignorable.

Note also that the original purpose for finding the Lagrangian was to use it to obtain the equations of motion, by substituting the Lagrangian (1.84) into the Lagrange equations (which would lead to a system of three second-order ODEs). Fortunately, in this example, it was not necessary to write down the Lagrange equations because from the Lagrangian we were able to find three constants of motion, which constitute a system of three first-order ODEs, and by solving this system, we found the general solution of the equations of motion.

It should be pointed out that in few cases we are able to find n, or more, constants of motion, with n being the number of degrees of freedom of the system.

Exercise 1.20. Find the natural Lagrangian of a particle of mass m in a uniform gravitational field in Cartesian coordinates. Show that even though not all of the coordinates are ignorable, there exists a constant of motion associated with each coordinate.

Example 1.21 (Angular momentum of a charged particle in an axially symmetric magnetic field). The Lagrangian (1.69) for a charged particle in a static magnetic field (not necessarily uniform), expressed in circular cylindrical coordinates, is

$$L = \frac{m}{2}(\dot{\rho}^2 + \rho^2\dot{\phi}^2 + \dot{z}^2) + \frac{e}{c}(A_1\dot{\rho} + A_2\dot{\phi} + A_3\dot{z}), \tag{1.88}$$

where the functions A_i are defined by $A_i \equiv \mathbf{A}\cdot(\partial\mathbf{r}/\partial q_i)$ (among other things, this implies that the dimensions of the A_i need not coincide with those of $\mathbf{A}$). Assuming that the magnetic field is invariant under rotations about the z-axis, we want to find a constant of motion associated with this symmetry.

As shown above, we will have a constant of motion associated with the rotations about the z-axis if the coordinate ϕ is ignorable or if there exists a function $G(q_i, t)$ such that

$$\frac{\partial L}{\partial \phi} = \frac{\partial G}{\partial q_i}\dot{q}_i + \frac{\partial G}{\partial t}$$

[see (1.83)]. Making use of (1.88) we see that this last condition takes the explicit form

$$\frac{e}{c}\left(\frac{\partial A_1}{\partial\phi}\dot{\rho}+\frac{\partial A_2}{\partial\phi}\dot{\phi}+\frac{\partial A_3}{\partial\phi}\dot{z}\right)=\frac{\partial G}{\partial\rho}\dot{\rho}+\frac{\partial G}{\partial\phi}\dot{\phi}+\frac{\partial G}{\partial z}\dot{z}+\frac{\partial G}{\partial t},$$

where $(\rho, \phi, z, \dot{\rho}, \dot{\phi}, \dot{z}, t)$ are independent variables (that is, we are not imposing yet the equations of motion). Hence, owing to this independence (functional, not linear, independence), the total differential of G must be given by

$$\mathrm{d}G=\frac{e}{c}\left(\frac{\partial A_1}{\partial\phi}\mathrm{d}\rho+\frac{\partial A_2}{\partial\phi}\mathrm{d}\phi+\frac{\partial A_3}{\partial\phi}\mathrm{d}z\right). \tag{1.89}$$

On the other hand, the existence of a function $G(\rho, \phi, z)$ satisfying (1.89) implies the conservation of [see (1.88)]

$$\frac{\partial L}{\partial\dot{\phi}}-G=m\rho^2\dot{\phi}+\frac{e}{c}A_2-G. \tag{1.90}$$

Thus, we consider the differential of the combination $G - eA_2/c$, which according to (1.89) must be

$$\begin{aligned}\mathrm{d}\left(G-\frac{e}{c}A_2\right)&=\frac{e}{c}\left[\left(\frac{\partial A_1}{\partial\phi}-\frac{\partial A_2}{\partial\rho}\right)\mathrm{d}\rho+\left(\frac{\partial A_3}{\partial\phi}-\frac{\partial A_2}{\partial z}\right)\mathrm{d}z\right]\\&=\frac{e}{c}(\rho B_\rho\mathrm{d}z-\rho B_z\mathrm{d}\rho),\end{aligned} \tag{1.91}$$

where B_ρ, B_ϕ, B_z are the components of the magnetic field corresponding to the vector potential $\mathbf{A}$, with respect to the *orthonormal* basis $(\hat{\rho}, \hat{\phi}, \hat{z})$ defined by the circular cylindrical coordinates. Thus, the existence of G is equivalent to the exactness of the differential form $\rho B_\rho \mathrm{d}z - \rho B_z \mathrm{d}\rho$.

Making use of the standard criterion for the exactness of a differential form, one finds the conditions

$$\frac{\partial(\rho B_\rho)}{\partial\rho}=-\frac{\partial(\rho B_z)}{\partial z},\qquad\frac{\partial(\rho B_\rho)}{\partial\phi}=0,\qquad\frac{\partial(\rho B_z)}{\partial\phi}=0,$$

and with the aid of the expression for the divergence of $\mathbf{B}$,

$$\nabla\cdot\mathbf{B}=\frac{1}{\rho}\frac{\partial(\rho B_\rho)}{\partial\rho}+\frac{1}{\rho}\frac{\partial B_\phi}{\partial\phi}+\frac{\partial B_z}{\partial z},$$

and the fact that $\nabla\cdot\mathbf{B}=0$, it follows that $\rho B_\rho \mathrm{d}z - \rho B_z \mathrm{d}\rho$ is an exact differential if and only if all the components B_ρ, B_ϕ, B_z are functions of ρ and z only.

In conclusion, for a charged particle in an axially symmetric magnetic field (which means that if the z-axis is the axis of symmetry, the components B_ρ, B_ϕ, B_z do not depend on ϕ), $m\rho^2\dot{\phi} + eA_2/c - G$ is a constant of motion, with the function $G - eA_2/c$ determined by the (gauge-independent) expression (1.91).

For instance, in the case of the uniform field considered in Example 1.19, the only cylindrical component of $\mathbf{B}$ different from zero is $B_z = B_0$ and the constant of motion (1.90) is $m\rho^2\dot{\phi} + eB_0\rho^2/2c$. (Note that this is the conserved momentum conjugate to the ignorable coordinate ϕ in the Lagrangian (1.77).) Other examples of axially symmetric magnetic fields are those corresponding to linear magnetic multipoles.

Exercise 1.22 (Linear momentum of a charged particle in a magnetic field invariant under translations). Show that the Lagrangian (1.88) possesses a constant of motion associated with the invariance under translations along the z-axis, given by $m\dot{z} + eA_3/c - G$, where $\mathrm{d}(G - eA_3/c) = (e/c)(B_\phi \mathrm{d}\rho - \rho B_\rho \mathrm{d}\phi)$, if and only if the cylindrical components of the magnetic field do not depend on z.

When a Lagrangian does not contain the variable t there is also an associated constant of motion. Making use of the Leibniz rule and the chain rule we obtain the *identity*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L\right) &= \frac{\mathrm{d}\dot{q}_i}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} + \dot{q}_i \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i}\dot{q}_i - \frac{\partial L}{\partial \dot{q}_i}\frac{\mathrm{d}\dot{q}_i}{\mathrm{d}t} - \frac{\partial L}{\partial t} \\ &= \dot{q}_i\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial t}.\end{aligned} \tag{1.92}$$

Hence, if L does not depend explicitly on t (i.e., $\partial L/\partial t = 0$), then

$$J \equiv \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L \tag{1.93}$$

(with implicit sum over the repeated index i) is a constant of motion, that is, its time derivative is equal to zero as a consequence of the equations of motion (1.49) (this constant is known as the Jacobi integral). When L has the standard form $L = T - V$, the Jacobi integral corresponds frequently, but not always, to the total energy $T + V$ (see below).

For instance, the Lagrangian (1.54) does not depend explicitly on the time and substituting it into Equation (1.93) one finds that in this case the Jacobi integral is

$$\begin{aligned}\dot{r}\frac{\partial L}{\partial \dot{r}} - L &= \dot{r}(m\dot{r}) - \tfrac{1}{2}m(\dot{r}^2 + \omega^2 r^2 \sin^2\theta_0) + mgr\cos\theta_0 \\ &= \tfrac{1}{2}m\dot{r}^2 - \tfrac{1}{2}m\omega^2 r^2 \sin^2\theta_0 + mgr\cos\theta_0,\end{aligned}$$

which coincides with (1.17) (and is not the total energy).

Similarly, the Lagrangian (1.84) does not depend explicitly on the time and, correspondingly, Equation (1.93) is a constant of motion. A direct substitution of (1.84) into (1.93) gives

$$\dot{x}\left(m\dot{x}-\frac{eB_0}{2c}y\right)+\dot{y}\left(m\dot{y}+\frac{eB_0}{2c}x\right)+\dot{z}\,(m\dot{z})-\frac{m}{2}(\dot{x}^2+\dot{y}^2+\dot{z}^2)-\frac{eB_0}{2c}(x\dot{y}-y\dot{x})$$
$$=\frac{m}{2}(\dot{x}^2+\dot{y}^2+\dot{z}^2),$$

which is the kinetic energy of the particle (not $T+V$).

Exercise 1.23. Find three constants of motion from the Lagrangian (1.77) and use them to obtain the solution of the equations of motion for a charged particle in a uniform magnetic field.

We also get a constant of motion if there exists a function $G(q_i, t)$ such that

$$\frac{\partial L}{\partial t}=\frac{\partial G}{\partial q_i}\dot{q}_i+\frac{\partial G}{\partial t} \tag{1.94}$$

[cf. (1.83)]. In fact, with the aid of (1.92), (1.94), and the Lagrange equations we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{q}_i\frac{\partial L}{\partial \dot{q}_i}-L+G\right)=0.$$

A (somewhat artificial) example is given by a charged particle in a static, uniform, electric field $\mathbf{E}=E_0\mathbf{k}$, where E_0 is a constant. Such a field is usually expressed making use of an electrostatic potential $\varphi=-E_0z$ [see Equations (1.65)], but we can also represent this field making use of the potentials $\varphi=0$ and $\mathbf{A}=-E_0ct\mathbf{k}$ (note that $\mathbf{B}=\nabla\times\mathbf{A}=\mathbf{0}$). Then, from (1.69) we find that, in Cartesian coordinates, the standard Lagrangian is

$$L=\frac{m}{2}(\dot{x}^2+\dot{y}^2+\dot{z}^2)-eE_0t\dot{z},$$

and, in this manner, L depends on the time (in spite of the fact that the field is static), however,

$$\frac{\partial L}{\partial t}=-eE_0\dot{z},$$

which has the form (1.94) with $G=-eE_0z$, therefore, we conclude that

$$\dot{q}_i\frac{\partial L}{\partial \dot{q}_i}-L+G=\frac{m}{2}(\dot{x}^2+\dot{y}^2+\dot{z}^2)-eE_0z$$

is a constant of motion.

In the case of an ordinary potential, $V(q_i, t)$, the ambiguity in its definition is somewhat trivial: if $V(q_i, t)$ and $V'(q_i, t)$ are two potentials for the generalized forces Q_i, then from the definition (1.48) we have

$$-\frac{\partial V}{\partial q_i} = -\frac{\partial V'}{\partial q_i},$$

which is equivalent to the existence of a function, $f(t)$, such that $V' = V + f(t)$. By contrast, the ambiguity in the definition of a velocity-dependent potential involves a function of the $n+1$ variables q_i and t.

Proposition 1.24. *The velocity-dependent potentials* $V(q_i, \dot{q}_i, t)$ *and* $V'(q_i, \dot{q}_i, t)$ *lead to the same generalized forces, that is*

$$-\frac{\partial V}{\partial q_i} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial V}{\partial \dot{q}_i} = -\frac{\partial V'}{\partial q_i} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial V'}{\partial \dot{q}_i}, \tag{1.95}$$

for $i = 1, 2, \ldots, n$ *[see Equations (1.58)], if and only if there exists a function* $F(q_i, t)$ *such that*

$$V'(q_i, \dot{q}_i, t) = V(q_i, \dot{q}_i, t) + \frac{\partial F}{\partial q_i}\dot{q}_i + \frac{\partial F}{\partial t}. \tag{1.96}$$

(Usually, the last two terms of the previous equation are written as $\mathrm{d}F/\mathrm{d}t$, which is not appropriate because in Equation (1.96) the variables $q_i, \dot{q}_i, t$ are independent and this point is crucial in the proof of this proposition.)

Proof. Letting $\Delta \equiv V' - V$, Equations (1.95) amount to

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \Delta}{\partial \dot{q}_i} - \frac{\partial \Delta}{\partial q_i} = 0$$

or, equivalently,

$$\frac{\partial^2 \Delta}{\partial t \partial \dot{q}_i} + \frac{\partial^2 \Delta}{\partial q_j \partial \dot{q}_i}\dot{q}_j + \frac{\partial^2 \Delta}{\partial \dot{q}_j \partial \dot{q}_i}\ddot{q}_j - \frac{\partial \Delta}{\partial q_i} = 0. \tag{1.97}$$

Since Δ is a function of $(q_i, \dot{q}_i, t)$ only, Equations (1.97) are *identically* satisfied only if the coefficients of the second derivatives $\ddot{q}_j$ vanish, that is, $\partial^2 \Delta/\partial \dot{q}_i \partial \dot{q}_j = 0$, for $i, j = 1, 2, \ldots, n$, which imply that Δ must be of the form

$$\Delta = \mu_j \dot{q}_j + \nu, \tag{1.98}$$

where the μ_i and ν are functions of (q_j, t) only. Substituting this expression into (1.97) we have

$$\frac{\partial \mu_i}{\partial t} + \frac{\partial \mu_i}{\partial q_j}\dot{q}_j - \frac{\partial \mu_j}{\partial q_i}\dot{q}_j - \frac{\partial \nu}{\partial q_i} = 0.$$

Now, since μ_i and ν are functions of (q_j, t) only, the last equations are identically satisfied only if the coefficients of the $\dot{q}_j$ vanish. This leads to the conditions

$$\frac{\partial \mu_i}{\partial q_j} = \frac{\partial \mu_j}{\partial q_i} \tag{1.99}$$

and

$$\frac{\partial \mu_i}{\partial t} - \frac{\partial \nu}{\partial q_i} = 0. \tag{1.100}$$

Equations (1.99) imply the local existence of a function $\tilde{F}(q_i, t)$ such that [cf. (1.55)]

$$\mu_i = \frac{\partial \tilde{F}}{\partial q_i} \tag{1.101}$$

and substituting this expression into (1.100) we obtain

$$\frac{\partial}{\partial q_i}\left(\frac{\partial \tilde{F}}{\partial t} - \nu\right) = 0,$$

which means that $\nu = \partial\tilde{F}/\partial t + f(t)$, where $f(t)$ is a function of t only. Hence, letting $F \equiv \tilde{F} + \chi(t)$, where $\chi(t)$ is an anti-derivative of $f(t)$, we have

$$\nu = \frac{\partial F}{\partial t}$$

and, from (1.101),

$$\mu_i = \frac{\partial F}{\partial q_i}.$$

Substituting these expressions into Equation (1.98) we obtain

$$\varDelta = \frac{\partial F}{\partial q_i}\dot{q}_i + \frac{\partial F}{\partial t}.$$

Conversely, if $V' = V + (\partial F/\partial q_j)\dot{q}_j + \partial F/\partial t$, for some $F(q_i, t)$, then

$$\begin{aligned}
-\frac{\partial V'}{\partial q_i} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial V'}{\partial \dot{q}_i} &= -\left(\frac{\partial V}{\partial q_i} + \frac{\partial^2 F}{\partial q_i \partial q_j}\dot{q}_j + \frac{\partial^2 F}{\partial q_i \partial t}\right) + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V}{\partial \dot{q}_i} + \frac{\partial F}{\partial q_i}\right) \\
&= -\frac{\partial V}{\partial q_i} - \frac{\partial^2 F}{\partial q_i \partial q_j}\dot{q}_j - \frac{\partial^2 F}{\partial q_i \partial t} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial V}{\partial \dot{q}_i} + \frac{\partial^2 F}{\partial q_j \partial q_i}\dot{q}_j + \frac{\partial^2 F}{\partial t \partial q_i} \\
&= -\frac{\partial V}{\partial q_i} + \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial V}{\partial \dot{q}_i}.
\end{aligned}$$

□

As shown above, if the generalized forces Q_i are functions of $(q_i, \dot{q}_i, t)$ only, then $V = \alpha_i \dot{q}_i + \beta$ [see (1.60)]. Hence, from Proposition 1.24 we conclude that the functions α_i and β are not uniquely defined. The sets of functions (α_i, β) and (α'_i, β') correspond to the same velocity-dependent potential if and only if there exists a function $F(q_i, t)$ such that

$$\alpha'_i = \alpha_i + \frac{\partial F}{\partial \dot{q}_i}, \qquad \beta' = \beta + \frac{\partial F}{\partial t}$$

[cf. Equations (1.66) and (1.67)].

The proof given of Proposition 1.24 shows that the following Proposition also holds.

Proposition 1.25. *Two functions, $L(q_i, \dot{q}_i, t)$ and $L'(q_i, \dot{q}_i, t)$, lead to the same Lagrange equations, in the sense that*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L'}{\partial \dot{q}_i} - \frac{\partial L'}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i}, \tag{1.102}$$

for $i = 1, 2, \ldots, n$, if and only if there exists a function $F(q_i, t)$ such that

$$L'(q_i, \dot{q}_i, t) = L(q_i, \dot{q}_i, t) + \frac{\partial F}{\partial q_i}\dot{q}_i + \frac{\partial F}{\partial t}. \tag{1.103}$$

This does not mean that two Lagrangians must be related in the form (1.103) in order to lead to the same set of equations of motion. A simple example is provided by the Lagrangians

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - mgy, \qquad L' = m\dot{x}\dot{y} - mgx.$$

Substituting these functions into the Lagrange equations one readily finds that both Lagrangians lead to $m\ddot{x} = 0$ and $m\ddot{y} + mg = 0$, but the difference $L' - L$ is not a linear function of $\dot{x}$ and $\dot{y}$ as required by (1.103). (See also Section 2.3.)

Summarizing, we have seen in this chapter that if we consider a holonomic mechanical system and we assume that the constraint forces satisfy d'Alembert's principle, then Newton's second law gives

$$\sum_{\alpha=1}^{N} \left(m_\alpha \ddot{\mathbf{r}}_\alpha - \mathbf{F}_\alpha^{(\mathrm{appl})}\right) \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} = 0 \tag{1.104}$$

$(i = 1, 2, \ldots, n)$, where the constraint forces do not appear. The q_i are generalized coordinates employed to parameterize the position vectors, $\mathbf{r}_\alpha$, of the particles of the system (with respect to some inertial frame). The number of equations in (1.104) is equal to the number of degrees of freedom of the system. When, additionally, the generalized forces (1.46) are derivable from a potential [Equation (1.58)], the left-hand side of Equations (1.104) can be written in the abbreviated form

$$\sum_{\alpha=1}^{N} \left(m_\alpha \ddot{\mathbf{r}}_\alpha - \mathbf{F}_\alpha^{(\text{appl})} \right) \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i},$$

with $L = T - V$, T is the kinetic energy of the system and V is the potential.

The position vectors $\mathbf{r}_\alpha$ have to be defined with respect to an inertial frame, but the generalized coordinates may be defined in relation to a non-inertial frame (see, e.g., Examples 1.3 and 2.4)

Chapter 2
Some Applications of the Lagrangian Formalism

As we have seen in the preceding chapter, the equations of motion of a mechanical system subject to holonomic constraints, with forces derivable from a potential, can be expressed in terms of a single function. The fact that a single function determines the whole set of equations of motion is very useful, as we have seen in Section 1.2. In this chapter we shall consider several additional examples of mechanical systems whose equations of motion are derived with the aid of the Lagrangian formalism. We also show that in the case of an arbitrary second-order ODE, there exist an infinite number of Lagrangians that reproduce the given equation, regardless of the meaning of the equation and of the variables appearing in it.

In Section 2.4 we briefly discuss the relationship between the Lagrange equations and some problems of the calculus of variations, and in Section 2.5 we show that there is a general procedure to find constants of motion associated with certain symmetries of the Lagrangian.

2.1 Central Forces

An important and relatively simple example of a mechanical system in classical mechanics corresponds to a particle subject to a central force. This problem can be defined in the following manner: there is fixed point (with respect to an inertial frame), called the center of force, such that, placing the origin at the center of force, the force on a particle in this field is of the form

$$f(r)\frac{\mathbf{r}}{r},$$

where r is the magnitude of the position vector $\mathbf{r}$ and $f(r)$ is a real-valued function of r only.

G. F. Torres del Castillo, *An Introduction to Hamiltonian Mechanics*, Birkhäuser Advanced Texts Basler Lehrbücher, https://doi.org/10.1007/978-3-319-95225-3_2

A central force is always derivable from a potential, which is a function of r only. In fact, from the definition of the generalized forces, Equation (1.46), we have

$$Q_i = f(r)\frac{\mathbf{r}}{r} \cdot \frac{\partial \mathbf{r}}{\partial q_i} = \frac{f(r)}{r}\frac{1}{2}\frac{\partial(\mathbf{r}\cdot\mathbf{r})}{\partial q_i} = \frac{f(r)}{r}\frac{1}{2}\frac{\partial r^2}{\partial q_i} = f(r)\frac{\partial r}{\partial q_i}.$$

If the function $V(r)$ is defined by the condition

$$\frac{\mathrm{d}V(r)}{\mathrm{d}r} = -f(r) \tag{2.1}$$

(that is, V is an indefinite integral of $-f$), then, by the chain rule,

$$Q_i = -\frac{\mathrm{d}V}{\mathrm{d}r}\frac{\partial r}{\partial q_i} = -\frac{\partial V}{\partial q_i},$$

which shows that, in effect, any central force is derivable from a potential, which is defined, up to a constant term, by Equation (2.1). (Note that it was not necessary to specify which set of coordinates q_i is used.)

The standard Lagrangian for a particle in a central field, expressed in spherical coordinates (r, θ, ϕ), is, therefore [see Equation (1.80)],

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2) - V(r). \tag{2.2}$$

Two of the corresponding Lagrange equations can be written down and partially solved, even if the potential $V(r)$ has not been specified. These are

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = \frac{\mathrm{d}}{\mathrm{d}t}(mr^2\dot{\theta}) - mr^2\sin\theta\cos\theta\dot{\phi}^2, \\ 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = \frac{\mathrm{d}}{\mathrm{d}t}(mr^2\sin^2\theta\dot{\phi}). \end{aligned} \tag{2.3}$$

The last of these equations expresses the conservation of $mr^2\sin^2\theta\dot{\phi}$, which is the z-component of the angular momentum of the particle about the center of force, $\mathbf{L} \equiv \mathbf{r} \times \mathbf{p}$, where $\mathbf{p}$ is the linear momentum $m\dot{\mathbf{r}}$,

$$L_3 = m(x\dot{y} - y\dot{x}) = mr^2\sin^2\theta\dot{\phi}. \tag{2.4}$$

Restricting ourselves to the case where $L_3 \neq 0$, by virtue of (2.4) we can replace the derivatives with respect to the time by derivatives with respect to ϕ. For instance, the first equation in (2.3) takes the form

$$0 = \frac{L_3}{mr^2\sin^2\theta}\frac{\mathrm{d}}{\mathrm{d}\phi}\left(mr^2\frac{L_3}{mr^2\sin^2\theta}\frac{\mathrm{d}\theta}{\mathrm{d}\phi}\right) - mr^2\sin\theta\cos\theta\left(\frac{L_3}{mr^2\sin^2\theta}\right)^2,$$

which reduces to

$$\frac{d^2}{d\phi^2}\cot\theta = -\cot\theta.$$

The general solution of this equation is

$$\cot\theta = A\cos\phi + B\sin\phi, \tag{2.5}$$

where A and B are constants. Recalling the relation between the spherical coordinates and the Cartesian ones, the last equation amounts to $z = Ax + By$, which corresponds to a plane passing through the origin (with normal $-A\mathbf{i} - B\mathbf{j} + \mathbf{k}$). In the case $L_3 = 0$, ϕ is a constant, which corresponds to some plane passing through the origin and containing the z-axis. Thus, we conclude that the motion of a particle subject to a central force lies on a plane passing through the center of force.

Since the Lagrangian (2.2) does not depend on t, there exists another constant of motion, which is given by [see (1.93)]

$$\dot{r}\frac{\partial L}{\partial \dot{r}} + \dot{\theta}\frac{\partial L}{\partial \dot{\theta}} + \dot{\phi}\frac{\partial L}{\partial \dot{\phi}} - L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2) + V(r). \tag{2.6}$$

This constant of motion corresponds to the total energy and will be denoted as E in what follows.

As is well known, the three Cartesian components of $\mathbf{L}$ are conserved, which can be readily demonstrated with the aid of the vector formalism and the direct use of the Newton's second law. For our purposes, it will be convenient to prove that the norm of $\mathbf{L}$ is conserved making use of the Lagrange equations (2.3). As a first step we note that $\mathbf{L}^2 = (\mathbf{r}\times\mathbf{p})^2 = \mathbf{r}^2\mathbf{p}^2 - (\mathbf{r}\cdot\mathbf{p})^2$ (which follows from the Lagrange identity, $(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) = (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) - (\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c})$, or from the fact that $|\mathbf{r}\times\mathbf{p}| = |\mathbf{r}||\mathbf{p}|\sin\alpha$ and $\mathbf{r}\cdot\mathbf{p} = |\mathbf{r}||\mathbf{p}|\cos\alpha$, where α is the angle between $\mathbf{r}$ and $\mathbf{p}$). Noting that

$$\mathbf{r}\cdot\mathbf{p} = \mathbf{r}\cdot m\frac{d\mathbf{r}}{dt} = \frac{m}{2}\frac{d(\mathbf{r}\cdot\mathbf{r})}{dt} = \frac{m}{2}\frac{dr^2}{dt} = mr\dot{r},$$

we have

$$\mathbf{L}^2 = r^2(m\dot{\mathbf{r}})^2 - (mr\dot{r})^2 = m^2r^2(\dot{\mathbf{r}}^2 - \dot{r}^2) = m^2r^4(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2). \tag{2.7}$$

On the other hand, differentiating Equation (2.5) with respect to the time and making use of Equation (2.4) we get

$$-\csc^2\theta\,\dot{\theta} = -(A\sin\phi - B\cos\phi)\dot{\phi} = -(A\sin\phi - B\cos\phi)\frac{L_3\csc^2\theta}{mr^2},$$

therefore, using again (2.4) and (2.5),

$$\begin{aligned} m^2r^4(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) &= {L_3}^2(A\sin\phi - B\cos\phi)^2 + {L_3}^2\csc^2\theta \\ &= {L_3}^2\big[(A\sin\phi - B\cos\phi)^2 + 1 + \cot^2\theta\big] \\ &= {L_3}^2\big[(A\sin\phi - B\cos\phi)^2 + 1 + (A\cos\phi + B\sin\phi)^2\big] \\ &= {L_3}^2(A^2 + B^2 + 1). \end{aligned}$$

Thus, $\mathbf{L}^2$ is a constant of motion [and has the value ${L_3}^2(A^2 + B^2 + 1)$] and, from (2.6), we obtain the ODE

$$\frac{m\dot{r}^2}{2} + \frac{\mathbf{L}^2}{2mr^2} + V(r) = E, \tag{2.8}$$

which determines r as a function of t.

In order to take advantage of the fact that the motion of a particle in a central field of force lies on a plane passing through the center of force, it is customary to select the Cartesian axes in such a way that the xy-plane coincides with the plane of the orbit. In this manner, the problem can be treated as the movement of a particle on the Euclidean plane. Then, making use of the polar coordinates, (r, θ), the standard Lagrangian is given by

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) - V(r), \tag{2.9}$$

Since θ and t do not appear in the Lagrangian, we immediately have the constants of motion

$$l \equiv \frac{\partial L}{\partial\dot{\theta}} = mr^2\dot{\theta}, \tag{2.10}$$

that represents the angular momentum of the particle about the origin [cf. Equation (2.4)], and the Jacobi integral

$$E \equiv \dot{q}_i\frac{\partial L}{\partial\dot{q}_i} - L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) + V(r), \tag{2.11}$$

corresponding to the total energy. (As a consequence of the fact that the vectors $\mathbf{r}$ and $\dot{\mathbf{r}}$ lie on the xy-plane, the x- and y-components of the angular momentum are identically equal to zero. The only nonvanishing Cartesian component of the angular momentum is denoted here as l.) From Equation (2.10) we have

$$\dot{\theta} = \frac{l}{mr^2}, \tag{2.12}$$

which, substituted into Equation (2.11), yields the first-order ODE

$$\frac{m\dot{r}^2}{2} + \frac{l^2}{2mr^2} + V(r) = E \tag{2.13}$$

[which is equivalent to (2.8)]. The solution of (2.13) is given by

$$\pm \int \mathrm{d}t = \int \frac{\mathrm{d}r}{\sqrt{\frac{2}{m}\left[E - V(r) - \frac{l^2}{2mr^2}\right]}}. \tag{2.14}$$

Note that in order to obtain Equations (2.12) and (2.13) it was not necessary to write down the Lagrange equations corresponding to the Lagrangian (2.9) explicitly, and that the problem has been reduced to finding two indefinite integrals [the constants of motion l and E are one half of the constants necessary to have the general solution of the equations of motion, the other two constants of motion correspond to the integration constants arising from (2.14) and (2.12)]; if one is able to find r as a function of t by means of (2.14), then that expression can be substituted into (2.12) to find θ as a function of t. Unfortunately, in most cases, this is not possible in terms of elementary functions.

Making use of the chain rule, the combination of Equations (2.12) and (2.14) gives a first-order ODE for r as a function of θ, which leads to the equation of the orbit, provided that $l \neq 0$. One finds

$$\frac{\mathrm{d}r}{\mathrm{d}\theta} = \pm\frac{r^2}{l}\sqrt{2m\left[E - V(r) - \frac{l^2}{2mr^2}\right]},$$

which leads to

$$\pm \int \mathrm{d}\theta = \int \frac{l\,\mathrm{d}r}{r^2\sqrt{2m\left[E - V(r) - \frac{l^2}{2mr^2}\right]}}. \tag{2.15}$$

This equation shows that, for a given potential, the properties of the orbit (e.g., shape, size) depend on the values of E and l (see the examples below). The solution of (2.15) adds an integration constant, which defines the orientation of the orbit in the plane. The arbitrariness of this integration constant is related to the rotational invariance of the problem.

Equation (2.12) shows that when $l = 0$, the orbit is part of a straight line passing through the center of force, $\theta = \text{const}$. (Note that (2.15) also makes sense if $l = 0$ and gives $\theta = \text{const}$.)

The product $\frac{1}{2}r^2\dot{\theta}$ is the area swept by the line joining the origin with the particle per unit of time, therefore, the conservation of the angular momentum (2.10) means that the line joining the center of force with the particle sweeps area at a constant rate. This behavior was discovered by Kepler in the case of the planetary motion and is known as Kepler's second law. As we have seen, this result holds for any central force field.

The Kepler Problem

It is not an exaggeration to say that the most important example of the central forces is defined by the potential $V(r) = -k/r$, where k is a constant, which defines the so-called Kepler problem. According to Newton's law of gravitation, the gravitational force produced by a static mass located at the origin corresponds to a potential of this form. The orbit of a particle in this field of force is obtained from Equation (2.15), assuming that $l \neq 0$, with the aid of the changes of variable,

$$u = \frac{1}{r} \qquad \text{and} \qquad u - \frac{mk}{l^2} = \frac{1}{l}\sqrt{2mE + \frac{m^2k^2}{l^2}}\,\cos\beta, \tag{2.16}$$

we get

$$\begin{aligned} \pm(\theta - \theta_0) &= -\int \frac{l\,\mathrm{d}u}{\sqrt{2mE + 2mku - l^2u^2}} \\ &= -\int \frac{l\,\mathrm{d}u}{\sqrt{2mE - l^2\left(u - \frac{mk}{l^2}\right)^2 + \frac{m^2k^2}{l^2}}} \\ &= \beta, \end{aligned} \tag{2.17}$$

where θ_0 is an integration constant. Hence, substituting (2.17) into Equations (2.16), we find that the equation of the orbit is

$$\begin{aligned} \frac{1}{r} &= \frac{mk}{l^2}\left[1 + \sqrt{1 + \frac{2El^2}{mk^2}}\,\cos(\theta - \theta_0)\right] \\ &= \frac{mk}{l^2}\big[1 + e\cos(\theta - \theta_0)\big], \end{aligned} \tag{2.18}$$

where

$$e \equiv \sqrt{1 + \frac{2El^2}{mk^2}}. \tag{2.19}$$

Equation (2.18) corresponds to a conic with one of its foci at the origin and eccentricity e. If $E = 0$, the eccentricity is equal to 1, and the orbit (2.18) is a parabola; if $E > 0$, the eccentricity is greater than 1, and the orbit is a hyperbola; finally, if $E < 0$, the eccentricity is less than 1, and the orbit is an ellipse (in particular, when $E = -mk^2/2l^2$, we have $e = 0$, which means that the orbit is a circle).

The phase portrait representing the behavior of r as a function of t is given by the graph of [see Equation (2.13)]

$$\frac{m\dot{r}^2}{2} + \frac{l^2}{2mr^2} - \frac{k}{r} = E,$$

or, equivalently,

$$\frac{m\dot{r}^2}{2} = E + \frac{mk^2}{2l^2} - \frac{l^2}{2m}\left(\frac{1}{r} - \frac{mk}{l^2}\right)^2. \tag{2.20}$$

By contrast with the phase curves considered in the previous chapter, in the present case each phase curve is characterized by two parameters, E and l, determined by the initial conditions. Each phase curve in Figure 2.1 corresponds to a different value of E, but all of them have a common nonzero value of l.

Restricting the attention to the case where the orbit is an ellipse (that is, $e < 1$ or, equivalently, $E < 0$), the minimum value of r is equal to $(1 - e)a$, where a is the major semiaxis of the ellipse (see Figure 2.2), which is attained at $\theta = \theta_0$ [see Equation (2.18)] hence (2.18) gives

$$\frac{1}{(1-e)a} = \frac{mk}{l^2}(1+e),$$

that is, $1 - e^2 = l^2/(mka)$ or, making use of (2.19),

$$a = -\frac{k}{2E}. \tag{2.21}$$

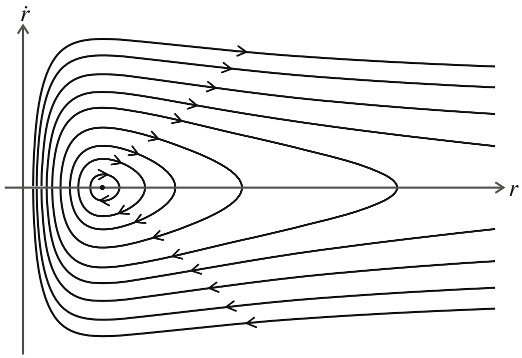


Fig. 2.1 Phase portrait for the radial coordinate in the Kepler problem with a fixed nonzero value of the angular momentum. There is a stable equilibrium point at $(l^2/mk, 0)$ with $E = -mk^2/2l^2$ [see Equation (2.20)], which corresponds to a circular orbit. The closed phase curves correspond to negative values of E and to motions where r is a periodic function of t. In fact, for $E < 0$, r and θ are periodic functions of t with the same period (except in the case of circular orbits) and for that reason the orbits are closed

Thus, while the eccentricity is determined by the values of E and l [see Equation (2.19)], the major semiaxis of the ellipse is a function of E only.

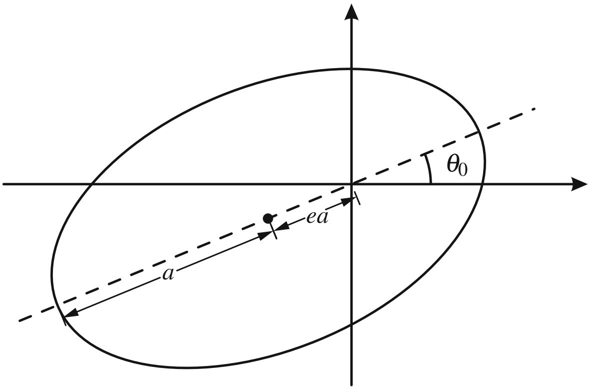


Fig. 2.2 In the case of the central potential $V = -k/r$, when $E < 0$ the orbit is an ellipse with the center of force at one of its foci. The distance from the center of the ellipse to a focus is ea, where e is the eccentricity and a is the major semiaxis of the ellipse

Equation (2.14), with $V(r) = -k/r$, can be rewritten as

$$\begin{aligned}\pm(t - t_0) &= \int \frac{m\,\mathrm{d}r}{\sqrt{2mE + \dfrac{2mk}{r} - \dfrac{l^2}{r^2}}}\\ &= \int \frac{mr\,\mathrm{d}r}{\sqrt{2mEr^2 + 2mkr - l^2}}\\ &= \int \frac{mr\,\mathrm{d}r}{\sqrt{2mE\left(r + \dfrac{k}{2E}\right)^2 - \dfrac{mk^2}{2E}\left(1 + \dfrac{2El^2}{mk^2}\right)}},\end{aligned}$$

where t_0 is an integration constant. Then, with the aid of (2.19) and (2.21), the "mechanical" constants E and l can be eliminated from the last integral in favor of the "geometrical" constants a and e,

$$\pm(t - t_0) = \sqrt{\frac{ma}{k}} \int \frac{r\,\mathrm{d}r}{\sqrt{a^2e^2 - (r-a)^2}}$$

and, with the change of variable $r - a = -ae\cos\psi$, we obtain

$$t - t_0 = \sqrt{\frac{ma^3}{k}} \int (1 - e\cos\psi)\mathrm{d}\psi.$$

(Whereas the sign of $\mathrm{d}r/\mathrm{d}t$ changes from negative to positive when the particle passes through the pericenter—the point of the orbit closest to the center of force—the sign of $\mathrm{d}\psi/\mathrm{d}t$ is always positive.)

Measuring the time in such a way that $t = 0$ when r passes through one minimum value, we obtain the relation

$$t = \sqrt{\frac{ma^3}{k}}(\psi - e\sin\psi). \tag{2.22}$$

Since $r = a(1 - e\cos\psi)$, the particle completes a revolution when ψ is increased by 2π and, therefore, from (2.22) it follows that the period, T, is given by

$$T = 2\pi\sqrt{\frac{ma^3}{k}} \tag{2.23}$$

and Equation (2.22) can be rewritten as

$$\frac{2\pi}{T}t = \psi - e\sin\psi. \tag{2.24}$$

Equation (2.23), relating the period with the semiaxis of the ellipse, is Kepler's third law while (2.24), relating the time with the so-called *eccentric anomaly*, ψ, is Kepler's equation.

The geometrical meaning of the eccentric anomaly will be presented in Section 4.3.1, where the Kepler problem is studied again, making use of another approach.

It would be desirable to invert (2.24), in order to find ψ as a function of t, which would allow us to express r as a function of t; however, if the eccentricity is different from zero, finding ψ in terms of t is a highly complicated problem.

Exercise 2.1. When $E = 0$, the orbit is a parabola with its focus at the center of force, but the angular momentum, l, can take any value. Show that the value of l is related to the minimum distance from the orbit to the center of force by

$$r_{\text{min}} = \frac{l^2}{2mk}$$

and that the dependence of r on the time is given implicitly by

$$\pm(t - t_0) = \frac{2}{3}\sqrt{\frac{m}{2k}}\,(r + 2r_{\text{min}})\sqrt{r - r_{\text{min}}},$$

so that $r = r_{\text{min}}$ when $t = t_0$.

Exercise 2.2. Analyze the case of hyperbolic orbits in the Kepler problem (that is $e > 1$ or, equivalently, $E > 0$) and obtain the analog of the Kepler equation. (In this case, the minimum value of r is $(e - 1)a$, where a is the major semiaxis of the hyperbola.)

The Isotropic Harmonic Oscillator

Another important and well-known example of the central forces is that of the isotropic harmonic oscillator, for which $V(r) = \frac{1}{2}m\omega^2 r^2$, where ω is a constant.

(It turns out that the isotropic harmonic oscillator is deeply related to the Kepler problem, in classical mechanics and in quantum mechanics; see, e.g., Section 4.3.1.) As in the previous example, making use of Equation (2.13) we find the phase portrait representing the behavior of r as a function of t, for a fixed nonzero value of l (see Figure 2.3). The phase curves are the graphs in the $r\dot{r}$-plane of

$$\frac{m\dot{r}^2}{2} + \frac{l^2}{2mr^2} + \frac{1}{2}m\omega^2 r^2 = E, \tag{2.25}$$

or, equivalently, of

$$\frac{m\dot{r}^2}{2} = E - |l|\omega - \frac{l^2}{2m}\left(\frac{1}{r} - \frac{m\omega r}{|l|}\right)^2. \tag{2.26}$$

Assuming that $l \neq 0$, the introduction of the variables v and χ given by

$$v = \frac{1}{r^2} \qquad \text{and} \qquad lv - \frac{mE}{l} = \sqrt{\frac{m^2E^2}{l^2} - m^2\omega^2}\, \cos\chi,$$

allows us to rewrite the integral (2.15) as

$$\pm(\theta - \theta_0) = \int \frac{l\,\mathrm{d}r}{r^3\sqrt{\dfrac{2mE}{r^2} - m^2\omega^2 - \dfrac{l^2}{r^4}}}$$

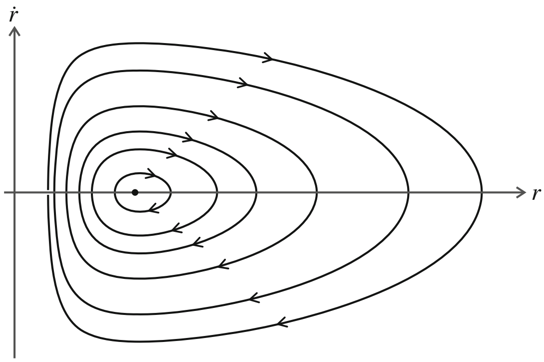


Fig. 2.3 Phase portrait of the radial coordinate of the isotropic harmonic oscillator, for a fixed nonzero value of the angular momentum. There is a stable equilibrium point at $(\sqrt{|l|/m\omega}, 0)$ with $E = |l|\omega$ [see Equation (2.26)], which corresponds to a circular orbit. All the phase curves are closed and r is a periodic function of t. All the orbits are ellipses centered at the origin and, for that reason, if the orbit is not a circle, in each period r completes two cycles

$$= -\int \frac{l\,\mathrm{d}v}{2\sqrt{2mEv - m^2\omega^2 - l^2v^2}}$$

$$= -\frac{l}{2}\int \frac{\mathrm{d}v}{\sqrt{\frac{m^2E^2}{l^2} - m^2\omega^2 - \left(lv - \frac{mE}{l}\right)^2}}$$

$$= \frac{\chi}{2},$$

hence

$$\frac{1}{r^2} = \frac{mE}{l^2}\left[1 + \sqrt{1 - \left(\frac{l\omega}{E}\right)^2}\cos 2(\theta - \theta_0)\right], \tag{2.27}$$

which is the equation of an ellipse centered at the origin. This equation shows that the orbit is a circle if and only if $E = |l|\omega$.

In order to find the dependence of r on t, with the aid of the changes of variable

$$u = r^2 \qquad \text{and} \qquad u - \frac{E}{m\omega^2} = \sqrt{\frac{E^2}{m^2\omega^4} - \frac{l^2}{m^2\omega^2}}\cos\beta,$$

from Equation (2.14) we obtain

$$\pm(t - t_0) = \int \frac{\mathrm{d}r}{\sqrt{\frac{2E}{m} - \omega^2r^2 - \frac{l^2}{m^2r^2}}}$$

$$= \int \frac{r\,\mathrm{d}r}{\omega\sqrt{\frac{2Er^2}{m\omega^2} - r^4 - \frac{l^2}{m^2\omega^2}}}$$

$$= \frac{1}{2\omega}\int \frac{\mathrm{d}u}{\sqrt{\frac{E^2}{m^2\omega^4} - \frac{l^2}{m^2\omega^2} - \left(u - \frac{E}{m\omega^2}\right)^2}}$$

$$= -\frac{\beta}{2\omega},$$

where t_0 is an integration constant, hence

$$r^2 = \frac{E}{m\omega^2}\left[1 + \sqrt{1 - \left(\frac{l\omega}{E}\right)^2}\cos 2\omega(t - t_0)\right], \tag{2.28}$$

which shows that r is a periodic function of the time with period π/ω. Taking into account that in each period the value of r passes through its maximum and minimum values twice, the period of the motion is $2\pi/\omega$. By contrast with the Kepler problem, where the period depends on the length of the major semiaxis of the ellipse (when $e < 1$), in this case the period is constant.

The equations of motion of an isotropic harmonic oscillator can be readily solved in Cartesian coordinates. Indeed, in these coordinates the standard Lagrangian is given by

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) - \tfrac{1}{2}m\omega^2(x^2 + y^2) \tag{2.29}$$

and the Lagrange equations yield $\ddot{x} = -\omega^2 x$ and $\ddot{y} = -\omega^2 y$. The general solution of these equations is

$$x = c_1 \cos\omega t + c_2 \sin\omega t, \qquad y = c_3 \cos\omega t + c_4 \sin\omega t,$$

where c_1, c_2, c_3, c_4 are arbitrary constants, which is the composition of two simple harmonic motions with period $2\pi/\omega$. One can verify directly, by eliminating t from these two equations, that the orbit is an ellipse centered at the origin.

The Lagrangian

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) - \tfrac{1}{2}m({\omega_1}^2 x^2 + {\omega_2}^2 y^2),$$

with $\omega_1 \neq \omega_2$ corresponds to an anisotropic harmonic oscillator; the solution of the equations of motion is

$$x = c_1 \cos\omega_1 t + c_2 \sin\omega_1 t, \qquad y = c_3 \cos\omega_2 t + c_4 \sin\omega_2 t,$$

where c_1, c_2, c_3, c_4 are arbitrary constants. Only when ω_1/ω_2 is a rational number, the orbit is closed and is a Lissajous figure.

The Repulsive Isotropic Harmonic Oscillator

Another simple example, closely related to the isotropic harmonic oscillator, is the so-called "repulsive isotropic harmonic oscillator," which corresponds to a particle subject to a radial repulsive force whose magnitude is proportional to the distance to the center of force. Hence, this example is defined by a potential of the form $V(r) = -\frac{1}{2}m\omega^2 r^2$, where ω is a real constant. Assuming that $l \neq 0$, making use of Equation (2.15) and the changes of variable

$$v = \frac{1}{r^2} \qquad \text{and} \qquad lv - \frac{mE}{l} = \sqrt{\frac{m^2E^2}{l^2} + m^2\omega^2}\,\cos\chi,$$

we find that the orbits are given

$$\begin{aligned}
\pm(\theta - \theta_0) &= \int \frac{l\,\mathrm{d}r}{r^3\sqrt{\dfrac{2mE}{r^2} + m^2\omega^2 - \dfrac{l^2}{r^4}}} \\
&= -\int \frac{l\,\mathrm{d}v}{2\sqrt{2mEv + m^2\omega^2 - l^2v^2}} \\
&= -\frac{l}{2}\int \frac{\mathrm{d}v}{\sqrt{\dfrac{m^2E^2}{l^2} + m^2\omega^2 - \left(lv - \dfrac{mE}{l}\right)^2}} \\
&= \frac{\chi}{2},
\end{aligned}$$

hence

$$\frac{1}{r^2} = \frac{mE}{l^2}\left[1 + \sqrt{1 + \left(\frac{l\omega}{E}\right)^2}\cos 2(\theta - \theta_0)\right], \tag{2.30}$$

which is the equation of a hyperbola centered at the origin [cf. Equation (2.27)].

2.2 Further Examples

In this section we shall consider some few additional examples of the application of the Lagrangian formulation to classical mechanics.

Example 2.3 (The spherical pendulum). Another common example is that of a spherical pendulum. This mechanical system consists of a point mass, m, attached to a weightless rod of constant length, l, which hangs from a fixed point (with respect to an inertial frame) in a uniform gravitational field (see Figure 2.4). It is assumed that the rod can rotate in all directions and therefore the system has two degrees of freedom. The configuration can be specified by the spherical coordinates, θ, ϕ (see Figure 2.4) and from (1.80) we find that the standard Lagrangian is given by

$$L = \frac{m}{2}l^2(\dot{\theta}^2 + \sin^2\theta\,\dot{\phi}^2) - mgl\cos\theta. \tag{2.31}$$

Since L does not depend on ϕ and t, we have the constants of motion

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = ml^2\sin^2\theta\,\dot{\phi}$$

(the z-component of the angular momentum) and the Jacobi integral

$$\frac{m}{2} l^2 (\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) + mgl \cos \theta$$

(the total energy). Denoting by L_3 and E, respectively, the values of these constants of motion, we obtain the first-order ODE

$$E = \frac{m}{2} l^2 \left(\dot{\theta}^2 + \frac{{L_3}^2}{m^2 l^4 \sin^2 \theta} \right) + mgl \cos \theta, \tag{2.32}$$

which gives θ as a function of t. The change of variable $u = \cos \theta$ leads to

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \pm \frac{1}{ml^2} \sqrt{2ml^2(E - mglu)(1 - u^2) - {L_3}^2}.$$

Thus, the solution can be expressed in terms of elliptic functions.

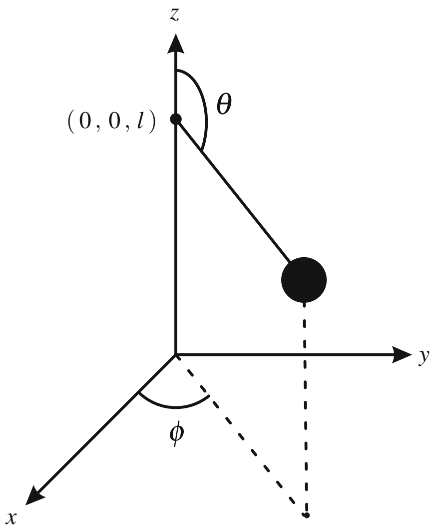


Fig. 2.4 One end of the rod is attached to the point with coordinates $(0, 0, l)$, where l is the length of the rod. At the other end there is a particle of mass m subject to a uniform gravitational field directed along the negative z-axis

In this example it is also interesting to find the Lagrangian in terms of another set of coordinates. If the motion of the pendulum bob is restricted to the lower hemisphere, the Cartesian coordinates, x, y, of the bob can be used as generalized coordinates (see Figure 2.4). From $x^2 + y^2 + (z - l)^2 = l^2$, we have $z = l - \sqrt{l^2 - x^2 - y^2}$ and, therefore, the standard Lagrangian takes the form

$$L = \frac{m}{2} \left[\dot{x}^2 + \dot{y}^2 + \frac{(x\dot{x} + y\dot{y})^2}{l^2 - x^2 - y^2} \right] - mg\left(l - \sqrt{l^2 - x^2 - y^2}\right),$$

which is not simpler than the expression in spherical coordinates. However, if we consider small oscillations about the stable equilibrium position $x = y = \dot{x} = \dot{y} = 0$,

$$L \simeq \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - \frac{mg}{2l}(x^2 + y^2), \tag{2.33}$$

which is the Lagrangian of a two-dimensional isotropic harmonic oscillator [cf. (2.29)]. Hence, for small oscillations, the projection on the xy-plane of the motion of the pendulum's bob is an ellipse with center at the origin (or a segment of a straight line).

Example 2.4 (The Foucault pendulum). This example is closely related to the previous one, but this time we shall consider a spherical pendulum hanging from a fixed point relative to the earth, assuming that the earth is rotating with constant angular velocity, ω, about a fixed axis with respect to an inertial frame whose origin is at the center of the earth. Making use of a set of Cartesian axes for this frame such that the z-axis is the axis of the earth's rotation, the position vector of a point P, fixed with respect to the earth surface, can be expressed in the form

$$\mathbf{r}_\mathrm{P} = R(\sin\theta_0 \cos\omega t\,\mathbf{i} + \sin\theta_0 \sin\omega t\,\mathbf{j} + \cos\theta_0\mathbf{k}),$$

where R is the distance from the earth center to P and θ_0 is the (constant) angle between the z-axis and the line going from the origin to P (see Figure 2.5). At $t = 0$, the point P crosses the xz-plane. The unit vectors

$$\hat{\theta} \equiv \cos\theta_0 \cos\omega t\,\mathbf{i} + \cos\theta_0 \sin\omega t\,\mathbf{j} - \sin\theta_0\mathbf{k}, \qquad \hat{\phi} \equiv -\sin\omega t\,\mathbf{i} + \cos\omega t\,\mathbf{j}$$

are tangent to the earth's surface at P (both are orthogonal to $\mathbf{r}_P$, $\hat{\theta}$ is tangent to the meridian passing through P, and $\hat{\phi}$ is tangent to the parallel passing through P). These two vectors together with

$$\hat{r} \equiv \sin\theta_0 \cos\omega t\,\mathbf{i} + \sin\theta_0 \sin\omega t\,\mathbf{j} + \cos\theta_0\mathbf{k}$$

form an orthonormal basis.

Now we consider a spherical pendulum hanging from a point a distance h above the point P. If l is the length of the pendulum, the position vector of the bob (*with respect to the inertial frame* mentioned above) can be expressed in the form

$$\mathbf{r} = (R + h - \sqrt{l^2 - x^2 - y^2})\hat{r} + x\hat{\theta} + y\hat{\phi}. \tag{2.34}$$

In this manner, x and y are Cartesian coordinates of the projection of the position vector $\mathbf{r}$ on the tangent plane to the earth at P (see Figure 2.6). From the definitions of the unit vectors $\hat{r}$, $\hat{\theta}$, and $\hat{\phi}$, one readily obtains

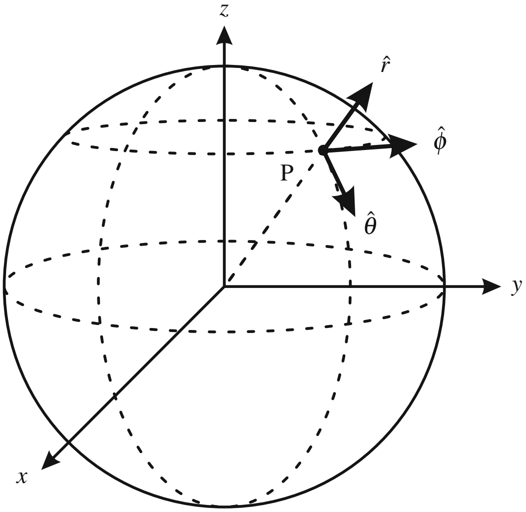


Fig. 2.5 The point P is fixed on the earth surface. We assume that the earth is rotating with a constant angular velocity ω with respect to the inertial frame with axes x, y, z. The vectors $\hat{\theta}$, $\hat{\phi}$, and $\hat{r}$ form an orthonormal basis at the point P on the surface of the earth. The unit vectors $\hat{\theta}$ and $\hat{\phi}$ are tangent to the surface of the earth; $\hat{\theta}$ points to the south and $\hat{\phi}$ points eastward

$$\frac{\mathrm{d}\hat{r}}{\mathrm{d}t} = \omega \sin\theta_0 \hat{\phi}, \qquad \frac{\mathrm{d}\hat{\theta}}{\mathrm{d}t} = \omega \cos\theta_0 \hat{\phi}, \qquad \frac{\mathrm{d}\hat{\phi}}{\mathrm{d}t} = -\omega \sin\theta_0 \hat{r} - \omega \cos\theta_0 \hat{\theta}. \tag{2.35}$$

Hence, the velocity of the pendulum bob (with respect to the inertial frame, expressed in terms of the orthonormal basis formed by $\hat{\theta}$, $\hat{\phi}$, and $\hat{r}$) is

$$\begin{aligned}\dot{\mathbf{r}} &= (\dot{x} - \omega y \cos\theta_0)\hat{\theta} + \left[(R + h - \sqrt{l^2 - x^2 - y^2})\omega \sin\theta_0 + \dot{y} + \omega x \cos\theta_0\right]\hat{\phi} \\ &\quad + \left(\frac{x\dot{x} + y\dot{y}}{\sqrt{l^2 - x^2 - y^2}} - \omega y \sin\theta_0\right)\hat{r}.\end{aligned}$$

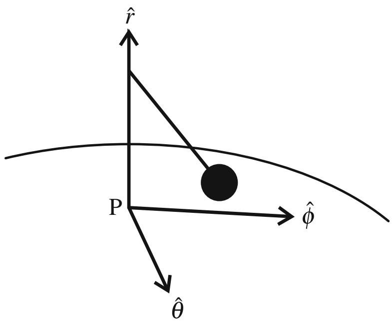


Fig. 2.6 The unit vectors $\hat{\theta}$, $\hat{\phi}$, and $\hat{r}$ form an orthonormal basis at the point P of the earth surface. The pendulum is hanging from a point at a height h above P

In order to simplify the calculations and the analysis of the solution of the equations of motion, we shall consider small oscillations (i.e., $|x|, |y| \ll l$) and we shall keep only the terms at most of first degree in ω, taking into account that the period of rotation of the earth is relatively long in comparison with the oscillation period of the pendulum. In this manner one obtains the approximate expression for the kinetic energy of the pendulum bob

$$T \simeq \frac{m}{2}\left[\dot{x}^2 + \dot{y}^2 + 2\omega\cos\theta_0(x\dot{y} - y\dot{x}) + 2(R + h - l)\omega\sin\theta_0\dot{y}\right]$$

One can readily verify that the last term, being a constant factor times $\dot{y}$, does not contribute in the Lagrange equations and we can omit it (see Proposition 1.25).

The potential is that of a particle of mass m in a uniform gravitational field, $V = mg(h - \sqrt{l^2 - x^2 - y^2})$, which, for small oscillations, is given approximately by

$$V \simeq mg\left[h - l\left(1 - \frac{1}{2}\frac{x^2 + y^2}{l^2}\right)\right] = \frac{mg}{2l}(x^2 + y^2) + mg(h - l).$$

The last term, which is a constant, can also be omitted. Thus, with the approximations already indicated,

$$L \simeq \frac{m}{2}\left[\dot{x}^2 + \dot{y}^2 + 2\omega\cos\theta_0(x\dot{y} - y\dot{x})\right] - \frac{mg}{2l}(x^2 + y^2) \tag{2.36}$$

[note that, when $\omega = 0$, this expression duly reduces to (2.33)]. Then, the Lagrange equations yield

$$\ddot{x} = 2\omega\cos\theta_0\,\dot{y} - \frac{gx}{l}, \qquad \ddot{y} = -2\omega\cos\theta_0\,\dot{x} - \frac{gy}{l}. \tag{2.37}$$

This is a system of two homogeneous, second-order linear ODEs, with constant coefficients and, therefore, can be easily solved. However, as we shall show, the equations of motion and their solution are more easily identified finding their expressions in a reference frame that rotates about $\hat{r}$. The coordinates x', y' measured with respect to Cartesian axes that rotate in the plane spanned by $\hat{\theta}$, $\hat{\phi}$ with a *constant* angular velocity $\varOmega$ are related to the coordinates x, y by means of

$$x = x'\cos\varOmega t - y'\sin\varOmega t, \qquad y = x'\sin\varOmega t + y'\cos\varOmega t.$$

Thus,

$$\begin{aligned}\dot{x} &= \dot{x}'\cos\varOmega t - \dot{y}'\sin\varOmega t - \varOmega(x'\sin\varOmega t + y'\cos\varOmega t),\\ \dot{y} &= \dot{x}'\sin\varOmega t + \dot{y}'\cos\varOmega t + \varOmega(x'\cos\varOmega t - y'\sin\varOmega t).\end{aligned}$$

Substituting these expressions into (2.36) we find

$$\begin{aligned}L \simeq \frac{m}{2}&\left[\dot{x}'^2 + \dot{y}'^2 + 2\varOmega(x'\dot{y}' - y'\dot{x}') + \varOmega^2(x'^2 + y'^2) + 2\omega\cos\theta_0(x'\dot{y}' - y'\dot{x}')\right]\\ &- \frac{mg}{2l}(x'^2 + y'^2).\end{aligned} \tag{2.38}$$

Choosing

$$\Omega = -\omega \cos\theta_0, \tag{2.39}$$

the linear terms in the velocities appearing in (2.38) are eliminated and, consistently with the approximations made above, we must neglect the term proportional to Ω^2. In this way we obtain

$$L \simeq \frac{m}{2}\left(\dot{x}'^2 + \dot{y}'^2\right) - \frac{mg}{2l}(x'^2 + y'^2),$$

which is the Lagrangian for a spherical pendulum in the approximation of small oscillations [cf. Equation (2.33)]. This means that if the pendulum is released at rest off the vertical position with respect to the $x'y'$-frame, it will oscillate in a vertical plane in this frame, but the plane of the oscillations will rotate with respect to the xy-frame with the angular velocity (2.39). In this way, the Foucault pendulum gives an experimental manner of demonstrating the rotation of the earth.

Exercise 2.5. Find the trajectory of a freely falling particle with respect to the Cartesian axes defined by $\hat{\theta}$, $\hat{\phi}$, $\hat{r}$, assuming that it is released from a height h above the surface of the earth, neglecting the terms of degree two and higher in ω.

2.3 The Lagrangians Corresponding to a Second-Order Ordinary Differential Equation

As we shall show in this section, for any given second-order ODE there exists an infinite number of Lagrangians that lead to that equation. We assume that

$$\ddot{q} = f(q, \dot{q}, t) \tag{2.40}$$

is a given a second-order ODE, which may correspond to a mechanical system or may have some other origin (the independent variable t does not have to be the time), and we want to find some function, $L(q, \dot{q}, t)$, such that

$$\frac{\partial^2 L}{\partial t \partial \dot{q}} + \dot{q}\frac{\partial^2 L}{\partial q \partial \dot{q}} + \ddot{q}\frac{\partial^2 L}{\partial \dot{q}^2} - \frac{\partial L}{\partial q} = 0 \tag{2.41}$$

be equivalent to Equation (2.40) [see Equation (1.51)]. This means that we are looking for a function $L(q, \dot{q}, t)$ that satisfies the second-order partial differential equation (PDE)

$$\frac{\partial^2 L}{\partial t \partial \dot{q}} + \dot{q}\frac{\partial^2 L}{\partial q \partial \dot{q}} + f(q, \dot{q}, t)\frac{\partial^2 L}{\partial \dot{q}^2} - \frac{\partial L}{\partial q} = 0 \tag{2.42}$$

(recall that q, $\dot{q}$, and t are here independent variables).

In order to solve (2.42), we take the partial derivative with respect to $\dot{q}$ on both sides of Equation (2.42). Letting

$$M \equiv \frac{\partial^2 L}{\partial \dot{q}^2}, \tag{2.43}$$

we have

$$\frac{\partial M}{\partial t} + \dot{q}\frac{\partial M}{\partial q} + f(q, \dot{q}, t)\frac{\partial M}{\partial \dot{q}} = -M\frac{\partial f}{\partial \dot{q}}, \tag{2.44}$$

which is a *first-order linear* PDE for $M(q, \dot{q}, t)$. Equation (2.44) can be solved using Lagrange's method of characteristics (see, e.g., Sneddon [14, Sect. 2.4]) but, in some simple cases, one can find a nontrivial solution by inspection. (Note that $M = 0$ is a solution of Equation (2.44), but this trivial solution is not useful because, when $M = 0$, Equation (2.41) is not a second-order ODE for $q(t)$.)

Making use of (2.40), Equation (2.44) can also be written as

$$\frac{\mathrm{d}M}{\mathrm{d}t} = -M\frac{\partial f}{\partial \dot{q}}. \tag{2.45}$$

From Equation (2.45) one readily sees that M is defined up to a multiplicative constant of motion; that is, if M_1 and M_2 are two solutions of Equation (2.45), then $\mathrm{d}(M_1/M_2)/\mathrm{d}t = 0$, which means that there are an infinite number of Lagrangians for Equation (2.40) since M_1/M_2 can be a trivial constant (i.e., a real number) or a constant of motion (see Example 2.11, below).

Once we have a solution of Equation (2.44), from Equation (2.43) we can find an expression for L, containing two indeterminate functions of q and t only. Substituting the expression for L thus obtained into Equation (2.42), the Lagrangian is determined up to an arbitrary function of q and t only (see the examples below).

Example 2.6 (A damped harmonic oscillator). A first example is given by the equation of motion

$$\ddot{q} + 2\gamma\dot{q} + \omega^2 q = 0, \tag{2.46}$$

which corresponds to a damped harmonic oscillator (here γ and ω are constants). In this case, Equation (2.44) takes the form

$$\frac{\partial M}{\partial t} + \dot{q}\frac{\partial M}{\partial q} - \left(2\gamma\dot{q} + \omega^2 q\right)\frac{\partial M}{\partial \dot{q}} = 2\gamma M,$$

and one can readily see that a solution of this equation is given by $M = m\mathrm{e}^{2\gamma t}$, where m is a constant, which is introduced in order to recover the standard Lagrangian of a harmonic oscillator when $\gamma = 0$. Then, using the definition of M [Equation (2.43)] we obtain

$$L = \frac{m}{2} e^{2\gamma t} \dot{q}^2 + g(q,t)\dot{q} + h(q,t), \tag{2.47}$$

where $g(q,t)$ and $h(q,t)$ are functions of q and t only. Substituting the expression (2.47) into (2.42), with $f(q,\dot{q},t) = -2\gamma\dot{q} - \omega^2 q$, we find that

$$\frac{\partial g}{\partial t} - \frac{\partial h}{\partial q} - m\omega^2 q\, e^{2\gamma t} = 0$$

or, equivalently,

$$\frac{\partial g}{\partial t} = \frac{\partial}{\partial q}\left(h + \tfrac{1}{2} m\omega^2 q^2 e^{2\gamma t}\right),$$

which implies the existence of a function $F(q,t)$ such that

$$g = \frac{\partial F}{\partial q}, \qquad h + \frac{1}{2} m\omega^2 q^2 e^{2\gamma t} = \frac{\partial F}{\partial t}.$$

Thus,

$$L(q,\dot{q},t) = \frac{m}{2} e^{2\gamma t}\left(\dot{q}^2 - \omega^2 q^2\right) + \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial t}, \tag{2.48}$$

where $F(q,t)$ is an arbitrary function, which can be taken equal to zero. (The last two terms in (2.48) correspond to the ambiguity found in Proposition 1.25.) Note that this Lagrangian does not have the standard form $T - V$.

Exercise 2.7. Show that the velocity-dependent force $-2m\gamma\dot{q} - m\omega^2 q$, considered in Example 2.6, cannot be derived from a generalized potential.

Example 2.8 (The Emden–Fowler equation). The Emden–Fowler equation

$$\ddot{q} + \frac{2}{t}\dot{q} + q^k = 0, \tag{2.49}$$

where k is a constant, arises in the study of the spherically symmetric equilibrium configurations of a self-gravitating polytropic fluid; t is proportional to the radial distance and q is related to the density of the fluid (this equation is also known as the Lane–Emden equation). In this case Equation (2.45) takes the form

$$\frac{dM}{dt} = M\frac{2}{t},$$

hence, we can choose $M = t^2$, i.e., $\partial^2 L/\partial\dot{q}^2 = t^2$, and therefore

$$L = \tfrac{1}{2} t^2 \dot{q}^2 + g(q,t)\dot{q} + h(q,t), \tag{2.50}$$

where g and h are some functions of two variables. Substituting this expression for L and $f(q, \dot{q}, t) = -2\dot{q}/t - q^k$ into Equation (2.42) we obtain

$$\frac{\partial g}{\partial t} - q^k t^2 = \frac{\partial h}{\partial q},$$

which can be written as

$$\frac{\partial g}{\partial t} = \frac{\partial}{\partial q}\left(h + \frac{q^{k+1}t^2}{k+1}\right).$$

Thus,

$$g = \frac{\partial F(q,t)}{\partial q}, \qquad h + \frac{q^{k+1}t^2}{k+1} = \frac{\partial F(q,t)}{\partial t},$$

where $F(q, t)$ is an *arbitrary* function of two variables and, substituting into Equation (2.50), we find the Lagrangian

$$L = \frac{t^2\dot{q}^2}{2} - \frac{t^2 q^{k+1}}{k+1} + \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial t}.$$

Choosing $F = 0$, we have

$$L = \frac{t^2\dot{q}^2}{2} - \frac{t^2 q^{k+1}}{k+1}. \tag{2.51}$$

Exercise 2.9. Find a Lagrangian for the Poisson–Boltzmann equation

$$\ddot{q} = -\frac{k}{t}\dot{q} - a\mathrm{e}^q,$$

where k and a are constants restricted by the conditions $k \neq 0$ and $a \in \{-1, 1\}$.

Exercise 2.10. Find a Lagrangian for the nonlinear equation

$$4\ddot{q} + 9t\dot{q}^{5/3} = 0.$$

(It may be noticed that this equation does not contain q and, therefore, it can be readily transformed into a first-order ODE that can be easily solved.)

Example 2.11. We consider the simple problem of a particle subject to a frictional force proportional to the velocity,

$$\ddot{q} = -2\gamma\dot{q}, \tag{2.52}$$

where γ is a constant [cf. Equation (2.46)]. Solving this linear equation one finds that two functionally independent constants of motion are $\dot{q}\mathrm{e}^{2\gamma t}$ and $\dot{q} + 2\gamma q$. (In fact, the general solution of (2.52) is of the form $q(t) = c_1 + c_2\mathrm{e}^{-2\gamma t}$, where c_1 and

c_2 are two arbitrary constants, hence, $\dot{q}(t) = -2\gamma c_2 e^{-2\gamma t}$. Expressing c_1 and c_2 in terms of q and $\dot{q}$, one identifies the two constants of motion previously indicated.)

In this case, Equation (2.45) takes the form

$$\frac{\mathrm{d}M}{\mathrm{d}t} = 2\gamma M$$

and a particular solution of this equation is $M = me^{2\gamma t}$. The general solution is $M = me^{2\gamma t} F(\dot{q}e^{2\gamma t}, \dot{q} + 2\gamma q)$, where F is an arbitrary function of two variables. Choosing $F(\dot{q}e^{2\gamma t}, \dot{q} + 2\gamma q) = 1/(\dot{q}e^{2\gamma t})$, we obtain $M = m/\dot{q}$ which leads to

$$L = m(\dot{q} \ln \dot{q} - \dot{q}) + g(q,t)\dot{q} + h(q,t),$$

where g and h are some functions of two variables. Proceeding as in the previous examples we find that the functions g and h can be chosen as $g = 0$ and $h(q,t) = -2m\gamma q$. Thus,

$$L = m(\dot{q} \ln \dot{q} - \dot{q}) - 2m\gamma q. \tag{2.53}$$

Exercise 2.12. Find the Lagrangians corresponding to the solution $M = me^{2\gamma t}$ obtained in Example 2.11.

Example 2.13 (Rocket motion). In the study of systems of variable mass in elementary mechanics one finds that the equation of motion of a rocket in a uniform gravitational field is given by

$$\ddot{q} = -g - u\frac{\mathrm{d}\ln m}{\mathrm{d}t}, \tag{2.54}$$

where g is the gravity acceleration, $m(t)$ is the mass of the rocket at time t, and u is the speed of the exhaust gases with respect to the rocket. Since the right-hand side of (2.54) is a function of t only, Equation (2.45) is satisfied if M is a trivial constant, e.g., $M = 1$. Then, from Equation (2.43) we see that $L = \frac{1}{2}\dot{q}^2 + b(q,t)\dot{q} + h(q,t)$, where $b(q,t)$ and $h(q,t)$ are some functions of q and t only.

Substituting the expression of L into (2.42), with $f(q,\dot{q},t) = -g - u\frac{\mathrm{d}\ln m}{\mathrm{d}t}$, we obtain

$$\frac{\partial b}{\partial t} - g - u\frac{\mathrm{d}\ln m}{\mathrm{d}t} - \frac{\partial h}{\partial q} = 0.$$

A simple solution of this equation is given by $b = 0$ and $h = -(g + u\frac{\mathrm{d}\ln m}{\mathrm{d}t})q$, which leads to the Lagrangian

$$L = \frac{1}{2}\dot{q}^2 - \left(g + u\frac{\mathrm{d}\ln m}{\mathrm{d}t}\right)q.$$

Taking into account that a second-order ODE requires the specification of a real-valued function of three variables [the function $f(q, \dot{q}, t)$ appearing in (2.40)], it is not surprising that any such equation can be expressed in terms of a Lagrangian, which is also a real-valued function of three variables. (This is related to the fact that when $n = 1$, any generalized force is derivable from a potential.) By contrast, giving a system of n second-order ODEs, with $n > 1$, requires the specification of n real-valued functions of $2n + 1$ variables, while a Lagrangian is a single real-valued function of $2n + 1$ variables, and therefore we can see that a system of ODEs has to be highly special in order to be expressible in terms a Lagrangian. Thus, for systems of two or more second-order ODEs, a Lagrangian may not exist.

2.4 Hamilton's Principle

As we shall show in this section, one finds systems of equations expressed in the form of the Lagrange equations in a wide class of problems where one is looking for curves that minimize certain integrals.

As we have shown in Chapter 1, for a mechanical system with holonomic constraints, each set of generalized coordinates, $(q_1, q_2, \ldots, q_n)$, constitutes a coordinate system of the configuration space of the mechanical system. Adding the time, t, to the set of coordinates $(q_1, q_2, \ldots, q_n)$, we obtain a system of coordinates $(q_1, q_2, \ldots, q_n, t)$ of the so-called *extended configuration space*. Each particular solution of the equations of motion is given by a curve in the extended configuration space. For instance, in the Kepler problem, discussed in Section 2.1, the solution of the equations of motion with $E < 0$, was found to be

$$r = a(1-e\cos\psi), \quad \theta = \theta_0 + \arccos\left(\frac{\cos\psi - e}{1 - e\cos\psi}\right), \quad t = \sqrt{\frac{ma^3}{k}}(\psi - e\sin\psi),$$

parameterized by the eccentric anomaly ψ. In most cases, the time itself is used as the parameter of the curve; for instance, in the case of the two-dimensional isotropic harmonic oscillator, each solution of the equations of motion has the form

$$x = c_1\cos\omega t + c_2\sin\omega t, \qquad y = c_3\cos\omega t + c_4\sin\omega t, \qquad t = t.$$

The Hamilton principle establishes that the curves described by a mechanical system in the extended configuration space are distinguished by the fact that the integral

$$\int_{t_0}^{t_1} L(q_i, \dot{q}_i, t)\,\mathrm{d}t$$

has a minimum (or stationary) value on the curve described by the system, compared with the curves in the extended configuration space with the same endpoints.

Let $L(q_i, \dot{q}_i, t)$, $i = 1, 2, \ldots, n$, be a given real-valued function of $2n + 1$ variables (which does not need to be related to a mechanical system, as in some of the examples considered in the preceding section), and let C be a curve, given by $C(t) = (q_i(t), t)$, where the $q_i(t)$ are real-valued functions of t, defined in some interval $[t_0, t_1]$. Starting from C, we define a second curve,

$$\overline{C}(t) \equiv (q_i(t), \mathrm{d}q_i(t)/\mathrm{d}t, t), \tag{2.55}$$

so that the line integral

$$I(C) \equiv \int_{\overline{C}} L \, \mathrm{d}t = \int_{t_0}^{t_1} L(q_i(t), \mathrm{d}q_i(t)/\mathrm{d}t, t) \, \mathrm{d}t$$

is some real number that depends on L and on the curve C. If t is expressed as a function of some parameter, τ, $t = t(\tau)$, then the line integral $I(C)$ is given by

$$I(C) = \int_{\tau_0}^{\tau_1} L\left(q_i(\tau), \frac{\mathrm{d}q_i(\tau)/\mathrm{d}\tau}{\mathrm{d}t(\tau)/\mathrm{d}\tau}, t(\tau)\right) \frac{\mathrm{d}t(\tau)}{\mathrm{d}\tau} \mathrm{d}\tau, \tag{2.56}$$

where τ_0 and τ_1 are defined by $t_0 = t(\tau_0)$ and $t_1 = t(\tau_1)$.

An illustrative example is given by $L(q, \dot{q}, t) = \sqrt{1 + \dot{q}^2}$. In this case, the integral (2.56) is

$$I(C) = \int_{\tau_0}^{\tau_1} \sqrt{1 + \left(\frac{\mathrm{d}q/\mathrm{d}\tau}{\mathrm{d}t/\mathrm{d}\tau}\right)^2} \frac{\mathrm{d}t}{\mathrm{d}\tau} \mathrm{d}\tau = \int_{\tau_0}^{\tau_1} \sqrt{\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 + \left(\frac{\mathrm{d}q}{\mathrm{d}\tau}\right)^2} \mathrm{d}\tau \tag{2.57}$$

and represents the (Euclidean) length of the curve $C(\tau) = (q(\tau), t(\tau))$ in the qt-plane, between the points $\mathrm{P}_0 = (q(\tau_0), t(\tau_0))$ and $\mathrm{P}_1 = (q(\tau_1), t(\tau_1))$ (see

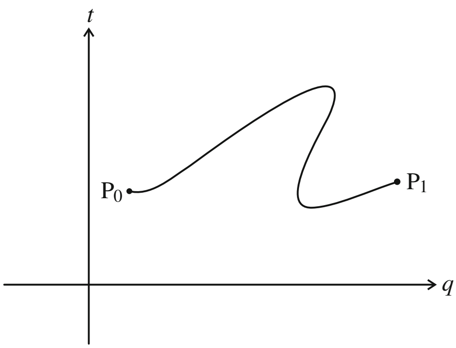


Fig. 2.7 The length of a curve $C(\tau) = (q(\tau), t(\tau))$ in the qt-plane is given by the integral (2.57). The entire curve shown *cannot* be parameterized by q or by t because there exist values of q for which there are two different values of t, and vice versa (in other words, the curve shown here is not the graph of a function $q = q(t)$, nor the graph of a function $t = t(q)$)

Figure 2.7). If the curve can be parameterized by t or q, the integral (2.57) reduces to the well-known formulas

$$\int_{t_0}^{t_1} \sqrt{1+\left(\frac{\mathrm{d}q}{\mathrm{d}t}\right)^2}\,\mathrm{d}t \qquad \text{or} \qquad \int_{q_0}^{q_1} \sqrt{1+\left(\frac{\mathrm{d}t}{\mathrm{d}q}\right)^2}\,\mathrm{d}q,$$

respectively, with t_0, t_1, q_0, and q_1 defined by $\mathrm{P}_0 = (q_0, t_0)$ and $\mathrm{P}_1 = (q_1, t_1)$.

The use of a parameter τ in place of t is especially important when we want to consider simultaneously several curves whose projections on the t-axis are not the same interval (see Figure 2.8).

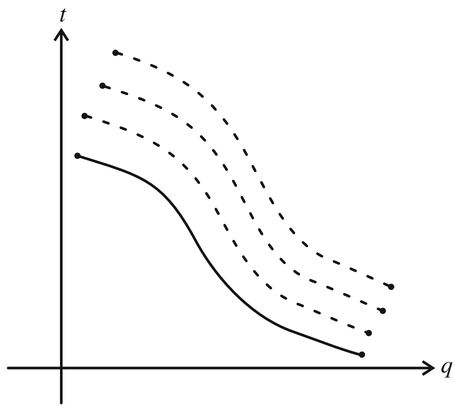


Fig. 2.8 The figure shows several curves belonging to a one-parameter family of curves. The projections of these curves on the t-axis need not be a single interval, but these curves can be parameterized by a parameter, τ, that in all cases takes values in the same interval $[\tau_0, \tau_1]$

For reasons that will become clear later, now we shall consider not a single curve but a one-parameter family of curves, labeled by a real parameter s, which can take values in some neighborhood of zero,

$$C^{(s)}(t) = (q_i^{(s)}(t), t).$$

The domain of these curves need not be the same interval, as shown in Figure 2.8; however, we can make use of a parameter τ such that all the curves $C^{(s)}$ correspond to the same interval $[\tau_0, \tau_1]$. In this way we are able to calculate the derivative with respect to s, at $s = 0$, of the real-valued function $I(C^{(s)})$. Employing the chain rule, (2.56), and the definitions

$$\eta_i(\tau) \equiv \left.\frac{\partial q_i^{(s)}(\tau)}{\partial s}\right|_{s=0}, \qquad \xi(\tau) \equiv \left.\frac{\partial t^{(s)}(\tau)}{\partial s}\right|_{s=0}, \tag{2.58}$$

we obtain

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}s} I(C^{(s)})\Big|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}s} \int_{\tau_0}^{\tau_1} L\left(q_i^{(s)}(\tau), \frac{\mathrm{d}q_i^{(s)}(\tau)/\mathrm{d}\tau}{\mathrm{d}t^{(s)}(\tau)/\mathrm{d}\tau}, t^{(s)}(\tau)\right) \frac{\mathrm{d}t^{(s)}}{\mathrm{d}\tau}\mathrm{d}\tau\Big|_{s=0} \\
&= \int_{\tau_0}^{\tau_1} \left\{\left[\frac{\partial L}{\partial q_i}\eta_i + \frac{\partial L}{\partial \dot{q}_i}\frac{\frac{\mathrm{d}t^{(0)}}{\mathrm{d}\tau}\frac{\mathrm{d}\eta_i}{\mathrm{d}\tau} - \frac{\mathrm{d}q_i^{(0)}}{\mathrm{d}\tau}\frac{\mathrm{d}\xi}{\mathrm{d}\tau}}{(\mathrm{d}t^{(0)}/\mathrm{d}\tau)^2} + \frac{\partial L}{\partial t}\xi\right]\frac{\mathrm{d}t^{(0)}}{\mathrm{d}\tau} + L\frac{\mathrm{d}\xi}{\mathrm{d}\tau}\right\}\mathrm{d}\tau, \\
&= \int_{\tau_0}^{\tau_1} \left[\frac{\partial L}{\partial q_i}\frac{\mathrm{d}t^{(0)}}{\mathrm{d}\tau}\eta_i + \frac{\partial L}{\partial \dot{q}_i}\left(\frac{\mathrm{d}\eta_i}{\mathrm{d}\tau} - \frac{\mathrm{d}q_i^{(0)}/\mathrm{d}\tau}{\mathrm{d}t^{(0)}/\mathrm{d}\tau}\frac{\mathrm{d}\xi}{\mathrm{d}\tau}\right) + \frac{\partial L}{\partial t}\frac{\mathrm{d}t^{(0)}}{\mathrm{d}\tau}\xi + L\frac{\mathrm{d}\xi}{\mathrm{d}\tau}\right]\mathrm{d}\tau, \quad (2.59)
\end{aligned}
$$

with L and its partial derivatives being evaluated on the curve $\overline{C^{(0)}}$. Integrating by parts the terms containing derivatives of ξ and η_i, we obtain the basic formula

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s} I(C^{(s)})\Big|_{s=0} &= \int_{\tau_0}^{\tau_1} \left(\frac{\partial L}{\partial q_i}\frac{\mathrm{d}t^{(0)}}{\mathrm{d}\tau} - \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial L}{\partial \dot{q}_i}\right)\left(\eta_i - \frac{\mathrm{d}q_i^{(0)}/\mathrm{d}\tau}{\mathrm{d}t^{(0)}/\mathrm{d}\tau}\xi\right)\mathrm{d}\tau \\
&\quad + \left[\frac{\partial L}{\partial \dot{q}_i}\eta_i + \left(L - \frac{\partial L}{\partial \dot{q}_i}\frac{\mathrm{d}q_i^{(0)}/\mathrm{d}\tau}{\mathrm{d}t^{(0)}/\mathrm{d}\tau}\right)\xi\right]\Big|_{\tau_0}^{\tau_1}. \qquad (2.60)
\end{aligned}
$$

As a first application of Equation (2.60) we shall assume that the integral $I(C)$ has an extreme value (a maximum or a minimum) at the curve $C^{(0)}$, compared with all the curves with the same endpoints (see Figure 2.9). From the definitions (2.58) we see that if all the curves $C^{(s)}$ have the same endpoints then

$$
0 = \xi(\tau_0) = \eta_i(\tau_0) = \xi(\tau_1) = \eta_i(\tau_1). \qquad (2.61)
$$

Hence, from (2.60) we have

$$
0 = \int_{\tau_0}^{\tau_1} \left(\frac{\partial L}{\partial q_i}\frac{\mathrm{d}t^{(0)}}{\mathrm{d}\tau} - \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial L}{\partial \dot{q}_i}\right)\left(\eta_i - \frac{\mathrm{d}q_i^{(0)}/\mathrm{d}\tau}{\mathrm{d}t^{(0)}/\mathrm{d}\tau}\xi\right)\mathrm{d}\tau.
$$

Since, apart from the conditions (2.61), the functions ξ and η_i are arbitrary, it follows that

$$
\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i}\frac{\mathrm{d}t^{(0)}}{\mathrm{d}\tau} = 0
$$

($i = 1, 2, \ldots, n$) on the curve $C^{(0)}$. In particular, if the curve $C^{(0)}$ can be parameterized by t, the last equations reduce to

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \tag{2.62}$$

which have the form of the Lagrange equations (1.49). In this context, not necessarily related to classical mechanics, Equations (2.62) are known as the Euler–Lagrange equations. Another application of (2.60) will be given in Section 2.5.

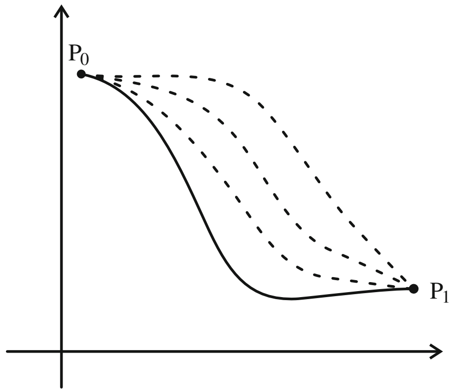


Fig. 2.9 The curves $C^{(s)}$ have the same endpoints for all values of s. The integral $I(C)$ has an extreme value on the curve $C^{(0)}$

The derivative $\mathrm{d}I(C^{(s)})/\mathrm{d}s|_{s=0}$ may be equal to zero even if $I(C^{(s)})$ does not have an extreme value at $s = 0$. We say that $I(C^{(s)})$ has a *stationary value* at $s = 0$ if $\mathrm{d}I(C^{(s)})/\mathrm{d}s|_{s=0} = 0$. Thus, we conclude that $I(C^{(s)})$ has a stationary value at $s = 0$ if and only if the curve $C^{(0)}$ satisfies the Euler–Lagrange equations. In particular, given two points of the extended configuration space, the solution of the Lagrange equations (1.49) passing through these points is the curve for which the integral $I(C)$ has a stationary value, compared with the curves passing through the given points. This result is known as *Hamilton's principle*.

Example 2.14 (Geodesic curves). A typical example of the application of the Euler–Lagrange equations arises in the search of the shortest curve joining two given points. Such curves are known as *geodesics*. The length of a curve in a surface of the Euclidean space or, more generally, in a Riemannian manifold, is locally given by an integral of the form

$$\int_{t_0}^{t_1} \sqrt{g_{ij}\frac{\mathrm{d}q_i}{\mathrm{d}t}\frac{\mathrm{d}q_j}{\mathrm{d}t}}\,\mathrm{d}t, \tag{2.63}$$

where $q_i(t_0)$ and $q_i(t_1)$ are the coordinates of the endpoints of the curve, the g_{ij} are some real-valued functions of the coordinates q_i (with expressions that depend on the surface or manifold of interest and on the coordinates), with $g_{ij} = g_{ji}$ and $\det(g_{ij}) \neq 0$ [see, e.g., Equation (2.57)].

According to the results derived above, the geodesics are determined by the Euler–Lagrange equations for

$$L = \sqrt{g_{ij} \frac{\mathrm{d}q_i}{\mathrm{d}t} \frac{\mathrm{d}q_j}{\mathrm{d}t}}. \tag{2.64}$$

Substituting this function into the Euler–Lagrange equations (2.62), making use of the symmetry of g_{ij} in its two indices (changing the names of the indices as necessary, in order to avoid that a summation index appears more than twice in a term) we obtain

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t} \left[\left(g_{rs} \dot{q}_r \dot{q}_s \right)^{-1/2} g_{ij} \dot{q}_j \right] - \frac{1}{2} \left(g_{rs} \dot{q}_r \dot{q}_s \right)^{-1/2} \frac{\partial g_{jk}}{\partial q_i} \dot{q}_j \dot{q}_k \\ &= -\frac{1}{2} \left(g_{rs} \dot{q}_r \dot{q}_s \right)^{-3/2} \left(2 g_{kl} \dot{q}_k \ddot{q}_l + \frac{\partial g_{kl}}{\partial q_m} \dot{q}_m \dot{q}_k \dot{q}_l \right) g_{ij} \dot{q}_j \\ &\quad + \left(g_{rs} \dot{q}_r \dot{q}_s \right)^{-1/2} \left(g_{ij} \ddot{q}_j + \frac{\partial g_{ij}}{\partial q_k} \dot{q}_k \dot{q}_j - \frac{1}{2} \frac{\partial g_{jk}}{\partial q_i} \dot{q}_j \dot{q}_k \right) \end{aligned} \tag{2.65}$$

($i = 1, 2, \ldots, n$). The n ODEs (2.65) are not independent since multiplying the right-hand side of (2.65) by $\dot{q}_i$ we obtain

$$\begin{aligned} &-\frac{1}{2} \left(g_{rs} \dot{q}_r \dot{q}_s \right)^{-1/2} \left(2 g_{kl} \dot{q}_k \ddot{q}_l + \frac{\partial g_{kl}}{\partial q_m} \dot{q}_m \dot{q}_k \dot{q}_l \right) \\ &+ \left(g_{rs} \dot{q}_r \dot{q}_s \right)^{-1/2} \left(g_{ij} \dot{q}_i \ddot{q}_j + \frac{\partial g_{ij}}{\partial q_k} \dot{q}_i \dot{q}_k \dot{q}_j - \frac{1}{2} \frac{\partial g_{jk}}{\partial q_i} \dot{q}_i \dot{q}_j \dot{q}_k \right), \end{aligned}$$

which is identically equal to zero. This result is related to the fact that the integral (2.63) is invariant under changes of parameter (as it should, since the length of a curve must not depend on the choice of the parametrization of the curve). In order to avoid this inconvenience, it is customary to impose the condition

$$g_{ij} \frac{\mathrm{d}q_i}{\mathrm{d}t} \frac{\mathrm{d}q_j}{\mathrm{d}t} = \text{const.} \tag{2.66}$$

on the parameter of the geodesics. Then, Equations (2.65) reduce to

$$g_{ij} \ddot{q}_j + \frac{\partial g_{ij}}{\partial q_k} \dot{q}_k \dot{q}_j - \frac{1}{2} \frac{\partial g_{jk}}{\partial q_i} \dot{q}_j \dot{q}_k = 0. \tag{2.67}$$

The constant on the right-hand side of (2.66) is equal to 1 if and only if t is the *arclength* of the curve.

We can readily verify that the so-called *geodesic equations*, (2.67), follow from any Lagrangian of the form

$$L = \tfrac{1}{2} \left(g_{ij} \dot{q}_i \dot{q}_j \right)^p, \tag{2.68}$$

where p is a real constant different from 0 and $1/2$ (the constant factor $1/2$ is introduced on the right-hand side of (2.68) for convenience). Indeed, following the same steps as in the case of the Lagrangian (2.64), we obtain

$$0 = p(p-1)\big(g_{rs}\dot{q}_r\dot{q}_s\big)^{p-2}\left(2g_{kl}\dot{q}_k\ddot{q}_l + \frac{\partial g_{kl}}{\partial q_m}\dot{q}_m\dot{q}_k\dot{q}_l\right)g_{ij}\dot{q}_j$$
$$+\,p\big(g_{rs}\dot{q}_r\dot{q}_s\big)^{p-1}\left(g_{ij}\ddot{q}_j + \frac{\partial g_{ij}}{\partial q_k}\dot{q}_k\dot{q}_j - \frac{1}{2}\frac{\partial g_{jk}}{\partial q_i}\dot{q}_j\dot{q}_k\right) \tag{2.69}$$

($i = 1, 2, \ldots, n$). Multiplying both sides of Equations (2.69) by $\dot{q}_i$ we have

$$0 = p(p-\tfrac{1}{2})\big(g_{rs}\dot{q}_r\dot{q}_s\big)^{p-1}\left(2g_{kl}\dot{q}_k\ddot{q}_l + \frac{\partial g_{kl}}{\partial q_m}\dot{q}_m\dot{q}_k\dot{q}_l\right).$$

Since we are assuming that p is different from 0 and $1/2$, it follows that the expression inside the last parenthesis must be equal to zero, which is equivalent to Equation (2.66). Hence, Equations (2.69) reduce to the geodesic equations (2.67).

Summarizing, if one uses the Lagrangian (2.64), then it is convenient to impose the condition (2.66) on the parameter of the geodesics and these curves are determined by the system of ODEs (2.67). However, if one employs the Lagrangian (2.68), with $p \neq \frac{1}{2}$, the Euler–Lagrange equations lead directly to (2.66) and (2.67).

Exercise 2.15. Show that Equation (2.66) follows from (2.67) and that Equation (2.66) expresses the conservation of the Jacobi integral for the Lagrangian (2.68).

Example 2.16 (Geodesics of the Poincaré half-plane). An example frequently considered in differential geometry is that of the so-called Poincaré half-plane, which is interesting because it provides a model of a non-Euclidean geometry. We start with the set of points of the Cartesian plane with $y > 0$, defining the length of a curve by means of

$$\int_{t_0}^{t_1}\sqrt{\frac{\dot{x}^2+\dot{y}^2}{y^2}}\,\mathrm{d}t.$$

(That is, $g_{11} = g_{22} = y^{-2}$, $g_{12} = g_{21} = 0$.) According to the discussion above, the geodesics corresponding to this definition of length can be obtained, e.g., from the Euler–Lagrange equations applied to the Lagrangian

$$L = \frac{1}{2}\frac{\dot{x}^2+\dot{y}^2}{y^2}.$$

Since this Lagrangian does not depend on x, the momentum conjugate to x,

$$\frac{\partial L}{\partial \dot{x}} = \frac{\dot{x}}{y^2}, \tag{2.70}$$

is conserved. Denoting by b the value of this constant, Equation (2.70) shows that if $b \neq 0$, then x can be used as a parameter of the curve, in place of t, and, by the chain rule,

$$\dot{y} = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t} = by^2\frac{\mathrm{d}y}{\mathrm{d}x}.$$

Furthermore,

$$\frac{1}{2}\frac{\dot{x}^2 + \dot{y}^2}{y^2}$$

is also a conserved quantity (see Equation (2.66) and Exercise 2.15). Denoting by E the value of this constant, we have

$$2Ey^2 = \dot{x}^2 + \dot{y}^2 = \left(by^2\right)^2 + \left(by^2\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2,$$

which leads to

$$\pm\int \mathrm{d}x = \int \frac{y\,\mathrm{d}y}{\sqrt{\frac{2E}{b^2} - y^2}}$$

and, therefore, $(x-a)^2+y^2 = 2E/b^2$, where a is an integration constant. In the case where $b = 0$, x is a constant. Thus, we conclude that the geodesics of the Poincaré half-plane are semicircles centered at points of the x-axis and lines parallel to the y-axis. (Note that it was not necessary to write down, and integrate, the second-order equations (2.67).)

As we shall see in Section 4.3, the orbits of certain mechanical systems are the geodesics of the configuration space, if the length of a curve is appropriately defined.

Exercise 2.17. Show that the geodesics of a sphere are arcs of great circles, that is, intersections of the sphere with planes passing through the center of the sphere.

Covariance of the Lagrange Equations Under Coordinate Transformations in the Extended Configuration Space

As an application of the Hamilton principle, we shall show that an *arbitrary* coordinate transformation in the extended configuration space, $q'_i = q'_i(q_j, t)$, $t' = t'(q_j, t)$, leaves the form of the Lagrange equations invariant, provided that the Lagrangian is appropriately transformed. Indeed, under such a coordinate change, the integral $I(C)$ takes the form

$$I(C) = \int_{t'_0}^{t'_1} L\left(q_i(q'_j, t'), \frac{\mathrm{d}q_i(q'_j, t')/\mathrm{d}t'}{\mathrm{d}t(q'_j, t')/\mathrm{d}t'}, t(q'_j, t')\right)\frac{\mathrm{d}t}{\mathrm{d}t'}\mathrm{d}t', \tag{2.71}$$

where

$$\frac{\mathrm{d}q_i}{\mathrm{d}t'} \equiv \frac{\partial q_i}{\partial t'} + \frac{\partial q_i}{\partial q'_j}\frac{\mathrm{d}q'_j}{\mathrm{d}t'}, \qquad \frac{\mathrm{d}t}{\mathrm{d}t'} \equiv \frac{\partial t}{\partial t'} + \frac{\partial t}{\partial q'_j}\frac{\mathrm{d}q'_j}{\mathrm{d}t'}. \tag{2.72}$$

Since the value of the integral $I(C)$ expressed in the form (2.71) is the same as that given by (2.56), the Euler–Lagrange equations (2.62) are equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}t'}\frac{\partial}{\partial \dot{q}_i'}\left(L\frac{\mathrm{d}t}{\mathrm{d}t'}\right) - \frac{\partial}{\partial q_i'}\left(L\frac{\mathrm{d}t}{\mathrm{d}t'}\right) = 0, \tag{2.73}$$

where it is understood that the Lagrangian, L, is expressed in terms of q_i', $\dot{q}_i'$ and t' as in Equation (2.71). In particular, if $t' = t$, we recover the result proved in Section 1.2 [see Equation (1.75)].

This means that if we write the ODEs obtained from the Lagrangian $L(q_i, \dot{q}_i, t)$ in terms of a new set of variables, q_i', t' (with the only condition that the transformation $q_i' = q_i'(q_j, t)$, $t' = t'(q_j, t)$, be invertible and differentiable), then the resulting equations are equivalent to the Euler–Lagrange equations corresponding to the Lagrangian $L(\mathrm{d}t/\mathrm{d}t')$, expressed in terms of q_i', $\mathrm{d}q_i'/\mathrm{d}t'$, t'. As in the case of a change of variable in an integral, what is desirable is to find a coordinate transformation leading to a simpler problem (see Example 2.18, below).

Example 2.18. A Lagrangian for the Poisson–Boltzmann equation (see Exercise 2.9), with $k = 1$, is given by

$$L = \frac{t}{2}\dot{q}^2 - at\mathrm{e}^q.$$

We shall study the effect on this Lagrangian of the coordinate transformation

$$t' = t^2\mathrm{e}^q, \qquad q' = \ln t,$$

or, equivalently,

$$t = \mathrm{e}^{q'}, \qquad q = \ln t' - 2q'.$$

From the last equations we find

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\mathrm{d}(\ln t' - 2q')}{\mathrm{d}\mathrm{e}^{q'}} = \frac{\mathrm{d}t'/t' - 2\mathrm{d}q'}{\mathrm{e}^{q'}\mathrm{d}q'} = \frac{1 - 2t'\mathrm{d}q'/\mathrm{d}t'}{t'\mathrm{e}^{q'}\mathrm{d}q'/\mathrm{d}t'}$$

and

$$\frac{\mathrm{d}t}{\mathrm{d}t'} = \mathrm{e}^{q'}\frac{\mathrm{d}q'}{\mathrm{d}t'},$$

hence, the Lagrangian for the new variables is

$$L\frac{\mathrm{d}t}{\mathrm{d}t'} = \left[\frac{\mathrm{e}^{q'}}{2}\frac{(1 - 2t'\mathrm{d}q'/\mathrm{d}t')^2}{t'^2\mathrm{e}^{2q'}(\mathrm{d}q'/\mathrm{d}t')^2} - a\mathrm{e}^{q'}\frac{t'}{\mathrm{e}^{2q'}}\right]\mathrm{e}^{q'}\frac{\mathrm{d}q'}{\mathrm{d}t'} = \frac{1}{2}\frac{(1 - 2t'\mathrm{d}q'/\mathrm{d}t')^2}{t'^2\mathrm{d}q'/\mathrm{d}t'} - at'\frac{\mathrm{d}q'}{\mathrm{d}t'},$$

which is a function of t' and $\mathrm{d}q'/\mathrm{d}t'$ only (that is, q' is ignorable).

2.5 Variational Symmetries

One of the advantages of expressing a system of ODEs in the form of the Lagrange equations is that the Lagrangian contains all the information that characterizes the system from the dynamical point of view. As we have seen in the foregoing sections, with the aid of the Lagrangian, in some cases, one can find constants of motion, which are very useful in the solution of the equations of motion.

Some constants of motion can be obtained by inspection when the Lagrangian does not depend on the time or has an ignorable coordinate. However, as we have seen in Section 1.2, the existence of an ignorable coordinate is not a property of the mechanical system alone, but it strongly depends on the coordinate system employed [compare, e.g., (1.79) with (1.80) or see Example 2.18]. As we shall see in this section, one can find certain constants of motion looking for one-parameter families of coordinate transformations that leave invariant the Lagrangian, in a sense to be defined below. One advantage of this approach is that the existence of these symmetries does not depend on the coordinate system chosen, but it is a property of the Lagrangian.

A first, natural, definition of invariance is this: we say that a one-parameter family of coordinate transformations $(q_i, t) \mapsto (q_i^{(s)}, t^{(s)})$ leaves (strictly) invariant the Lagrangian L if $I(C^{(s)})$ does not depend on s. Then $\mathrm{d}I(C^{(s)})/\mathrm{d}s = 0$ and, in particular, $\mathrm{d}I(C^{(s)})/\mathrm{d}s|_{s=0} = 0$, so that, from Equation (2.60) we find that, if the curve $C^{(0)}$ satisfies the Euler–Lagrange equations, then

$$\left[\frac{\partial L}{\partial \dot{q}_i}\eta_i + \left(L - \frac{\partial L}{\partial \dot{q}_i}\frac{\mathrm{d}q_i^{(0)}/\mathrm{d}\tau}{\mathrm{d}t^{(0)}/\mathrm{d}\tau}\right)\xi\right]\Bigg|_{\tau_0}^{\tau_1} = 0,$$

for all values of τ_0 and τ_1, which means that

$$\frac{\partial L}{\partial \dot{q}_i}\eta_i + \left(L - \frac{\partial L}{\partial \dot{q}_i}\frac{\mathrm{d}q_i^{(0)}/\mathrm{d}\tau}{\mathrm{d}t^{(0)}/\mathrm{d}\tau}\right)\xi$$

is a constant of motion, with L and its partial derivatives evaluated as in (2.59).

However, it is convenient to consider a less restrictive definition of invariance. A one-parameter family of coordinate transformations $(q_i, t) \mapsto (q_i^{(s)}, t^{(s)})$ is a *variational symmetry* of L if there exists a function $G(q_i, t)$ such that

$$\frac{\mathrm{d}}{\mathrm{d}s} I(C^{(s)})\bigg|_{s=0} = G\bigg|_{\tau_0}^{\tau_1}. \tag{2.74}$$

Then, from (2.60), we find that

$$\frac{\partial L}{\partial \dot{q}_i}\eta_i + \left(L - \frac{\partial L}{\partial \dot{q}_i}\frac{\mathrm{d}q_i^{(0)}/\mathrm{d}\tau}{\mathrm{d}t^{(0)}/\mathrm{d}\tau}\right)\xi - G \tag{2.75}$$

is a constant of motion. The functions η_i and ξ are determined by the condition

$$\frac{\partial L}{\partial q_i}\frac{\mathrm{d}t^{(0)}}{\mathrm{d}\tau}\eta_i + \frac{\partial L}{\partial \dot{q}_i}\left(\frac{\mathrm{d}\eta_i}{\mathrm{d}\tau} - \frac{\mathrm{d}q_i^{(0)}/\mathrm{d}\tau}{\mathrm{d}t^{(0)}/\mathrm{d}\tau}\frac{\mathrm{d}\xi}{\mathrm{d}\tau}\right) + \frac{\partial L}{\partial t}\frac{\mathrm{d}t^{(0)}}{\mathrm{d}\tau}\xi + L\frac{\mathrm{d}\xi}{\mathrm{d}\tau} = \frac{\mathrm{d}G}{\mathrm{d}\tau}, \tag{2.76}$$

which follows from (2.59) and (2.74) written in the form

$$\frac{\mathrm{d}}{\mathrm{d}s}I(C^{(s)})\bigg|_{s=0} = \int_{\tau_0}^{\tau_1}\frac{\mathrm{d}G}{\mathrm{d}\tau}\,\mathrm{d}\tau.$$

When t is used as the parameter, it follows from (2.75) (suppressing the superscript (0) from $q_i^{(0)}$ and $t^{(0)}$) that

$$\frac{\partial L}{\partial \dot{q}_i}\eta_i + \left(L - \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i\right)\xi - G \tag{2.77}$$

is a constant of motion, with the functions η_i and ξ determined by [see (2.76)]

$$\frac{\partial L}{\partial q_i}\eta_i + \frac{\partial L}{\partial \dot{q}_i}\left(\frac{\mathrm{d}\eta_i}{\mathrm{d}t} - \dot{q}_i\frac{\mathrm{d}\xi}{\mathrm{d}t}\right) + \frac{\partial L}{\partial t}\xi + L\frac{\mathrm{d}\xi}{\mathrm{d}t} = \frac{\mathrm{d}G}{\mathrm{d}t}. \tag{2.78}$$

The $n+1$ functions, η_i and ξ, of the $n+1$ variables q_i and t, are determined by the single PDE (2.76) (or Equation (2.78), if the curve is parameterized by t). The fact that this equation must hold for all values of q_i, $\dot{q}_i$, and t, and that its dependence on the $\dot{q}_i$ is given explicitly once a Lagrangian is chosen, leads to a set of PDEs that allows us to find η_i and ξ (see Examples 2.19 and 2.20, below). It may be remarked that in order to find the constants of motion associated with a variational symmetry we only have to know the functions ξ and η_i [see Equation (2.77)].

In most textbooks and articles, when the relationship between constants of motion and symmetries is presented, only "infinitesimal transformations" are considered as if they were the only ones that there exist, or that are relevant (an infinitesimal transformation is a transformation that "infinitesimally" differs from the identity). Actually, a transformation can be a symmetry even if it is not "close" to the identity (see, e.g., Example 2.19, below). Another common feature of the standard textbooks is that the examples given there are limited to translations, rotations, and Galilean transformations, when the symmetries are usually much more involved and interesting (see, e.g., Equations (2.84), below).

The "infinitesimal transformations" are equivalent to the functions η_i and ξ, which determine the constants of motion associated with the symmetry and, for this reason, it is enough to find the "infinitesimal transformations."

We can readily see that the results of this section reduce to those presented in Section 1.2. For instance, when the coordinate q_k (for a particular value k) is ignorable, then Equation (2.78) is satisfied with $\eta_k = 1$, $\xi = 0$, $\eta_i = 0$ for all $i \neq k$, and $G = 0$. According to (2.77), the conserved quantity associated with this symmetry is $\partial L/\partial \dot{q}_k$ (the momentum conjugate to q_k). Similarly, Equation (2.78) reduces to Equation (1.83) if $\eta_k = 1$, $\xi = 0$, $\eta_i = 0$ for all $i \neq k$, and (2.77) shows that the corresponding conserved quantity is $p_k - G$, as concluded in Section 1.2.

If $\partial L/\partial t = 0$, then Equation (2.78) is satisfied with $\eta_i = 0$, $\xi = 1$ and $G = 0$. Equation (2.77) shows that $L - (\partial L/\partial \dot{q}_i)\dot{q}_i$ is conserved, and this is, except for a minus sign, the Jacobi integral found in Equation (1.93).

Example 2.19 (Variational symmetries of a Lagrangian of the Poisson–Boltzmann equation). The Poisson–Boltzmann equation

$$\ddot{q} = -\frac{k}{t}\dot{q} - a\mathrm{e}^q,$$

where k and a are constants (with $k \neq 0$), can be obtained from the Lagrangian

$$L = \tfrac{1}{2}t^k\dot{q}^2 - at^k\mathrm{e}^q$$

(see Exercise 2.9). In order to find the variational symmetries of this Lagrangian, we substitute it into Equation (2.78), which gives

$$\begin{aligned}&-at^k\mathrm{e}^q\eta + t^k\dot{q}\left(\frac{\partial\eta}{\partial t} + \dot{q}\frac{\partial\eta}{\partial q} - \dot{q}\frac{\partial\xi}{\partial t} - \dot{q}^2\frac{\partial\xi}{\partial q}\right)\\ &+ \left(\tfrac{1}{2}kt^{k-1}\dot{q}^2 - kat^{k-1}\mathrm{e}^q\right)\xi + \left(\tfrac{1}{2}t^k\dot{q}^2 - at^k\mathrm{e}^q\right)\left(\frac{\partial\xi}{\partial t} + \dot{q}\frac{\partial\xi}{\partial q}\right) = \frac{\partial G}{\partial t} + \dot{q}\frac{\partial G}{\partial q}.\end{aligned}$$

Since the unknown functions, η, ξ, and G, depend on q and t only, the last equation is identically satisfied if and only if the corresponding coefficients of the powers of $\dot{q}$ on both sides of the equation coincide. This leads to the system of equations

$$-\frac{1}{2}t^k\frac{\partial\xi}{\partial q} = 0, \tag{2.79}$$

$$t^k\frac{\partial\eta}{\partial q} - \frac{t^k}{2}\frac{\partial\xi}{\partial t} + \frac{k}{2}t^{k-1}\xi = 0, \tag{2.80}$$

$$t^k\frac{\partial\eta}{\partial t} = \frac{\partial G}{\partial q}, \tag{2.81}$$

$$-at^k\mathrm{e}^q\eta - kat^{k-1}\mathrm{e}^q\xi - at^k\mathrm{e}^q\frac{\partial\xi}{\partial t} = \frac{\partial G}{\partial t}. \tag{2.82}$$

Equation (2.79) implies that $\xi = A(t)$, where A is a real-valued function of a single variable. Substituting this result into Equation (2.80) we obtain

$$\eta = \left(\frac{1}{2}\frac{\mathrm{d}A}{\mathrm{d}t} - \frac{k}{2}\frac{A}{t}\right)q + B(t),$$

where B is another real-valued function of one variable. Then, from Equations (2.81) and (2.82) we find

$$\frac{\partial G}{\partial q} = \frac{1}{2} t^k q \left(\frac{\mathrm{d}^2 A}{\mathrm{d}t^2} - \frac{k}{t}\frac{\mathrm{d}A}{\mathrm{d}t} + \frac{kA}{t^2} \right) + t^k \frac{\mathrm{d}B}{\mathrm{d}t},$$

$$\frac{\partial G}{\partial t} = -a t^k \mathrm{e}^q \left[\left(\frac{1}{2}\frac{\mathrm{d}A}{\mathrm{d}t} - \frac{k}{2}\frac{A}{t} \right) q + B + \frac{\mathrm{d}A}{\mathrm{d}t} + k\frac{A}{t} \right].$$

Making use of these equations we calculate the mixed second partial derivatives of G

$$\frac{\partial^2 G}{\partial t \partial q} = t^{k-1} \left(\frac{t}{2}\frac{\mathrm{d}^3 A}{\mathrm{d}t^3} - \frac{k^2}{2t}\frac{\mathrm{d}A}{\mathrm{d}t} + \frac{k}{t}\frac{\mathrm{d}A}{\mathrm{d}t} + \frac{k^2}{2}\frac{A}{t^2} - k\frac{A}{t^2} \right) q + \frac{\mathrm{d}}{\mathrm{d}t}\left(t^k \frac{\mathrm{d}B}{\mathrm{d}t} \right)$$

and

$$\frac{\partial^2 G}{\partial q \partial t} = -a t^{k-1} \mathrm{e}^q \left[\left(\frac{t}{2}\frac{\mathrm{d}A}{\mathrm{d}t} - \frac{kA}{2} \right) q + \frac{3t}{2}\frac{\mathrm{d}A}{\mathrm{d}t} + \frac{kA}{2} + tB \right].$$

Since these functions must coincide for all values of q, we obtain the conditions

$$\frac{t}{2}\frac{\mathrm{d}A}{\mathrm{d}t} - \frac{kA}{2} = 0, \qquad \frac{3t}{2}\frac{\mathrm{d}A}{\mathrm{d}t} + \frac{kA}{2} + tB = 0, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(t^k \frac{\mathrm{d}B}{\mathrm{d}t} \right) = 0,$$

among others. The first two of these equations give $A = c_1 t^k$, where c_1 is an arbitrary constant, and $B = -2c_1 k t^{k-1}$, which substituted into the third equation leads to $k = 1$ (excluding the trivial solution $c_1 = 0$). This means that, for $k \neq 1$, the Lagrangian under consideration *does not possess variational symmetries*.

Thus, in the case $k = 1$, $A(t) = c_1 t$, $B(t) = -2c_1$, and no further conditions are obtained from the equality of the mixed partial derivatives of G. In fact, we find that G is a trivial constant, which can be chosen equal to zero. Hence, the functions ξ and η are given by

$$\xi = c_1 t, \qquad \eta = -2c_1, \tag{2.83}$$

and substituting these expressions and that of the Lagrangian into Equation (2.77), we find that the constant of motion associated with the variational symmetry is

$$t\dot{q}(-2c_1) + \left(\tfrac{1}{2}t\dot{q}^2 - at\mathrm{e}^q - t\dot{q}\dot{q}\right)c_1 t = -c_1\left(2t\dot{q} + at^2\mathrm{e}^q + \tfrac{1}{2}t^2\dot{q}^2\right).$$

On the other hand, the knowledge of the functions ξ and η allows us, in principle at least, to find a family of coordinate transformations that constitute a variational symmetry of the Lagrangian. Noting first that from the definitions (2.58) and Equations (2.83) we have

$$\left.\frac{\partial t^{(s)}}{\partial s}\right|_{s=0} = \left. c_1 t^{(s)}\right|_{s=0}, \qquad \left.\frac{\partial q^{(s)}}{\partial s}\right|_{s=0} = -2c_1,$$

we *impose* the condition that these equations hold for all s, i.e., we assume that

$$\frac{\partial t^{(s)}}{\partial s} = c_1 t^{(s)}, \qquad \frac{\partial q^{(s)}}{\partial s} = -2c_1,$$

then, the solution of these equations is the one-parameter family of coordinate transformations

$$t^{(s)} = \mathrm{e}^{c_1 s} t^{(0)}, \qquad q^{(s)} = q^{(0)} - 2c_1 s, \tag{2.84}$$

which is defined for all $s \in \mathbb{R}$. In fact, a direct substitution shows that $I(C^{(s)}) = I(C^{(0)})$, for all $s \in \mathbb{R}$. (That is, we have a strict symmetry.)

As we have seen in Example 2.19, there exist Lagrangians that do not possess (nontrivial) variational symmetries (note that Equations (2.76) and (2.78) always have the trivial solution $\eta_i = 0$, $\xi = 0$). On the other hand, finding the variational symmetries of a Lagrangian with two or more degrees of freedom can become very cumbersome, since the number of equations to solve grows rapidly as the number of degrees of freedom increases.

Two different Lagrangians leading to a system of ODEs need not have the same variational symmetries (or even the same number of variational symmetries). However, if two Lagrangians are related as in Equation (1.103), then they possess the same variational symmetries.

Example 2.20 (Damped harmonic oscillator). As we have seen in Example 2.6, the Lagrangian

$$L = \frac{m}{2}\mathrm{e}^{2\gamma t}\big(\dot{q}^2 - \omega^2 q^2\big)$$

leads to the equation of motion of a damped harmonic oscillator. Substituting the Lagrangian into Equation (2.78) one obtains the PDE

$$-m\mathrm{e}^{2\gamma t}\omega^2 q\eta + m\mathrm{e}^{2\gamma t}\dot{q}\left(\frac{\partial \eta}{\partial t} + \dot{q}\frac{\partial \eta}{\partial q} - \dot{q}\frac{\partial \xi}{\partial t} - \dot{q}^2\frac{\partial \xi}{\partial q}\right)$$
$$+ m\gamma\mathrm{e}^{2\gamma t}\big(\dot{q}^2 - \omega^2 q^2\big)\xi + \frac{m}{2}\mathrm{e}^{2\gamma t}\big(\dot{q}^2 - \omega^2 q^2\big)\left(\frac{\partial \xi}{\partial t} + \dot{q}\frac{\partial \xi}{\partial q}\right) = \frac{\partial G}{\partial t} + \dot{q}\frac{\partial G}{\partial q}.$$

Equating the coefficients of the various powers of $\dot{q}$ on both sides of the last equation the following set of equations is obtained

$$-\frac{m}{2}\mathrm{e}^{2\gamma t}\frac{\partial\xi}{\partial q} = 0, \tag{2.85}$$

$$m\mathrm{e}^{2\gamma t}\left(\frac{\partial\eta}{\partial q} - \frac{1}{2}\frac{\partial\xi}{\partial t} + \gamma\xi\right) = 0, \tag{2.86}$$

$$m\mathrm{e}^{2\gamma t}\left(\frac{\partial\eta}{\partial t} - \frac{\omega^2}{2}q^2\frac{\partial\xi}{\partial q}\right) = \frac{\partial G}{\partial q}, \tag{2.87}$$

$$-m\mathrm{e}^{2\gamma t}\left(\omega^2 q\eta + \gamma\omega^2 q^2\xi + \frac{\omega^2}{2}q^2\frac{\partial\xi}{\partial t}\right) = \frac{\partial G}{\partial t}. \tag{2.88}$$

Equation (2.85) implies that $\xi = A(t)$, where $A(t)$ is a function of t only and Equation (2.86) gives

$$\eta = \left(\frac{1}{2}\frac{\mathrm{d}A}{\mathrm{d}t} - \gamma A\right)q + B(t),$$

where $B(t)$ is a function of t only. Hence, from Equations (2.87) and (2.88) it follows that

$$\frac{\partial G}{\partial q} = m\mathrm{e}^{2\gamma t}\left[\left(\frac{1}{2}\frac{\mathrm{d}^2A}{\mathrm{d}t^2} - \gamma\frac{\mathrm{d}A}{\mathrm{d}t}\right)q + \frac{\mathrm{d}B}{\mathrm{d}t}\right], \qquad \frac{\partial G}{\partial t} = -m\mathrm{e}^{2\gamma t}\omega^2\left(q^2\frac{\mathrm{d}A}{\mathrm{d}t} + qB\right)$$

and the equality of the mixed second partial derivatives of G leads to the condition

$$-2\gamma^2\frac{\mathrm{d}A}{\mathrm{d}t}q + 2\gamma\frac{\mathrm{d}B}{\mathrm{d}t} + \frac{1}{2}\frac{\mathrm{d}^3A}{\mathrm{d}t^3}q + \frac{\mathrm{d}^2B}{\mathrm{d}t^2} = -2\omega^2 q\frac{\mathrm{d}A}{\mathrm{d}t} - \omega^2 B,$$

which has to be satisfied for all values of q. Hence,

$$\frac{\mathrm{d}^3A}{\mathrm{d}t^3} + 4(\omega^2 - \gamma^2)\frac{\mathrm{d}A}{\mathrm{d}t} = 0, \qquad \frac{\mathrm{d}^2B}{\mathrm{d}t^2} + 2\gamma\frac{\mathrm{d}B}{\mathrm{d}t} + \omega^2 B = 0.$$

The general solutions of these equations contain three arbitrary constants in the case of $A(t)$ and two constants in the case of $B(t)$, which means that this Lagrangian possesses five one-parameter families of variational symmetries. The form of these general solutions depends on the value of the difference $\omega^2 - \gamma^2$. However, a particular nontrivial solution is given by $A(t) = 1$, $B(t) = 0$. Then $\xi = 1$, $\eta = -\gamma q$, and G is a constant that can be taken equal to zero. Substituting these expressions into (2.77) one finds the constant of motion

$$m\mathrm{e}^{2\gamma t}\dot{q}(-\gamma q) + \left[\frac{m}{2}\mathrm{e}^{2\gamma t}(\dot{q}^2 - \omega^2 q^2) - m\mathrm{e}^{2\gamma t}\dot{q}\dot{q}\right] = -\mathrm{e}^{2\gamma t}\left(\frac{m}{2}\dot{q}^2 + \frac{m}{2}\omega^2 q^2 + m\gamma q\dot{q}\right).$$

As in Example 2.19, with the aid of the functions ξ and η we can find a one-parameter family of variational symmetries of the Lagrangian. Substituting $\xi = 1$ and $\eta = -\gamma q$ into (2.58) we obtain

$$\left.\frac{\partial t^{(s)}}{\partial s}\right|_{s=0} = 1, \qquad \left.\frac{\partial q^{(s)}}{\partial s}\right|_{s=0} = -\left.\gamma q^{(s)}\right|_{s=0},$$

and we impose the condition that these equations hold for all s, that is

$$\frac{\partial t^{(s)}}{\partial s} = 1, \qquad \frac{\partial q^{(s)}}{\partial s} = -\gamma q^{(s)},$$

with the solution

$$t^{(s)} = t^{(0)} + s, \qquad q^{(s)} = q^{(0)} \mathrm{e}^{-\gamma s}.$$

Exercise 2.21. Find the variational symmetries of the Lagrangian

$$L = \frac{9}{2}\dot{q}^{1/3} + \frac{9}{4}tq$$

and find the associated constants of motion. Are they functionally independent? (This Lagrangian leads to the ODE considered in Exercise 2.10.)

Exercise 2.22. Find the variational symmetries of the Lagrangian

$$L = \tfrac{1}{2}m(\dot{r}^2 + \omega^2 r^2 \sin^2\theta) - mgr\cos\theta,$$

where m, ω, g, and θ are constants [see Equation (1.54)]. Show that there are two one-parameter families of variational symmetries with $\xi = 0$ and that the corresponding constants of motion are those given by Equation (1.20).

It can be shown that for a given Lagrangian, the existence of a variational symmetry is equivalent to the existence of coordinates in the extended configuration space such that one of the coordinates is ignorable [in the sense of Equation (1.83)] [20].

Chapter 3
Rigid Bodies

Another interesting application of the Lagrangian formalism is found in the motion of a rigid body. A rigid body can be defined as a collection of point particles such that the distances between them are constant. Even though, in essence, this example is similar to those already considered, the expression of the kinetic energy of a rigid body involves a more elaborate process and the definition of a new object (the inertia tensor).

This chapter differs from the other chapters of this book by the extensive use of objects with indices. A more elementary approach is based on the use of the vector algebra. The treatment given here highlights the use of the Lagrangian formalism.

3.1 The Configuration Space of a Rigid Body with a Fixed Point

We shall restrict ourselves to the study of the motion of a rigid body assuming that there exists a fixed point (with respect, of course, to some inertial frame). We shall also assume that the particles forming the rigid body are not all collinear (this means that there are at least three particles). Under these conditions, the system has three degrees of freedom (see below).

In order to study the motion of a rigid body with a fixed point, following a standard approach, we consider two sets of Cartesian axes, the first one, with coordinates x, y, z, assumed inertial, and the second one, with coordinates x', y', z', fixed in the rigid body. The origins of both sets of Cartesian axes coincide with the fixed point of the body (see Figure 3.1). It is convenient to denote x, y, z as x_1, x_2, x_3, and, similarly, x', y', z' as $x_{1'}, x_{2'}, x_{3'}$. Then, any point of the rigid body has a position vector

$$\mathbf{r} = x_{1'}\mathbf{e}_{1'} + x_{2'}\mathbf{e}_{2'} + x_{3'}\mathbf{e}_{3'},$$

G. F. Torres del Castillo, *An Introduction to Hamiltonian Mechanics*, Birkhäuser Advanced Texts Basler Lehrbücher, https://doi.org/10.1007/978-3-319-95225-3_3

where the unit vectors $\mathbf{e}_{i'}$ ($i' = 1', 2', 3'$) form an orthonormal basis fixed in the body, and, at the same time,

$$\mathbf{r} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3,$$

where the unit vectors $\mathbf{e}_i$ ($i = 1, 2, 3$) form an orthonormal basis associated with the inertial frame.

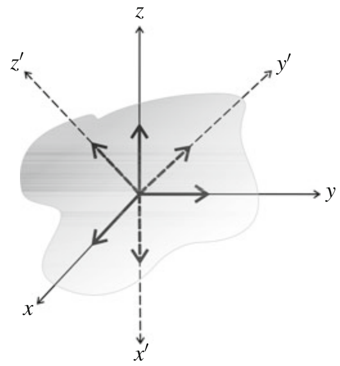


Fig. 3.1 The Cartesian axes with coordinates x', y', z' are fixed in the body and rotate about the origin of the Cartesian axes with coordinates x, y, z, which belong to an inertial frame. Both sets of axes are right-handed

Since the vectors $\mathbf{e}_i$ form a basis, there exist nine real numbers, $a_{ij'}$, which may depend on the time only, such that

$$\mathbf{e}_{i'} = a_{ji'}\mathbf{e}_j \tag{3.1}$$

(note the position of the indices, the order is purely conventional). With these numbers we can form a 3×3 matrix, $A = (a_{ij'})$, in the usual manner, using the first subscript to label rows and the second subscript to label columns, that is

$$A = \begin{pmatrix} a_{11'} & a_{12'} & a_{13'} \\ a_{21'} & a_{22'} & a_{23'} \\ a_{31'} & a_{32'} & a_{33'} \end{pmatrix},$$

so that the i-th column of this matrix contains the components of the vector $\mathbf{e}_{i'}$ with respect to the basis formed by the vectors $\mathbf{e}_j$. For instance, if the vectors $\mathbf{e}_{i'}$ are obtained from the vectors $\mathbf{e}_i$ by means of a rotation through an angle ϕ about the z-axis, then the matrix A is (see Figure 3.2)

$$A = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.2}$$

Using the fact that both bases are orthonormal we have

$$\delta_{i'j'} = \mathbf{e}_{i'} \cdot \mathbf{e}_{j'} = a_{ki'}\mathbf{e}_k \cdot a_{lj'}\mathbf{e}_l = a_{ki'}a_{lj'}\delta_{kl} = a_{ki'}a_{kj'},$$

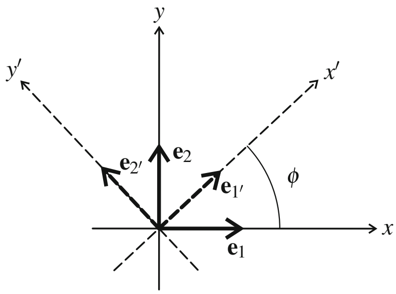


Fig. 3.2 The vectors $\mathbf{e}_{i'}$ are obtained from the vectors $\mathbf{e}_i$ by a rotation through an angle ϕ about the z-axis. The figure shows that $\mathbf{e}_{1'} = \cos\phi\,\mathbf{e}_1 + \sin\phi\,\mathbf{e}_2$, $\mathbf{e}_{2'} = -\sin\phi\,\mathbf{e}_1 + \cos\phi\,\mathbf{e}_2$, $\mathbf{e}_{3'} = \mathbf{e}_3$, which leads to the matrix (3.2)

i.e.,

$$\delta_{i'j'} = a_{ki'}a_{kj'}. \tag{3.3}$$

These conditions mean that the matrix $(a_{ij'})$ is orthogonal ($A^{\mathrm{t}}A = I$, where A^{t} is the transpose of A, and I is the unit matrix), that is, its inverse is equal to its transpose and, since any matrix commutes with its inverse, Equation (3.3) is equivalent to

$$\delta_{ij} = a_{ik'}a_{jk'}. \tag{3.4}$$

Once we have chosen the basis vectors $\mathbf{e}_i$ and $\mathbf{e}_{i'}$, the matrix $A = (a_{ij'})$ determines the configuration of the rigid body and, therefore, for a rigid body with a fixed point, the configuration space can be identified with the set of all the real 3×3 orthogonal matrices with positive determinant (so that the orientation, or handedness, of the basis vectors is not inverted). This set of matrices is, in fact, a group, which is denoted by SO(3). The set of equations (3.3) constitute six algebraically independent conditions on the nine entries of A (since both sides of the equation are symmetric in the indices i', j'); hence, the 3×3 orthogonal matrices can be parameterized by three coordinates (for instance, the three Euler angles presented in Section 3.3, below).

Usually, in the so-called tensor notation, the indices labeling the components of an object (e.g., a vector or a tensor) determine the way in which these components transform under a change of the basis vectors. For that reason, here we need two different kinds of indices: the unprimed and the primed ones, because we can perform two different kinds of changes of bases. We can replace the orthonormal basis $\mathbf{e}_i$ by another orthonormal basis (related to the first one by means of a constant orthogonal matrix) and, *independently*, we can replace the orthonormal basis $\mathbf{e}_{i'}$ by another orthonormal basis, also fixed in the rigid body (and the two orthonormal bases fixed with respect to the body are also related by some constant orthogonal matrix, see, e.g., Equation (3.19), below). The equations developed here must maintain their form under the independent changes of the two bases.

It may be remarked that in all the other examples in this book, we start by choosing some coordinates to represent the configuration of the mechanical system (which, in some cases, are replaced afterwards). By contrast, in the case of the motion of a rigid body we can postpone this choice and establish several results without having to write down the explicit expression of the Lagrangian in terms of coordinates.

3.2 The Instantaneous Angular Velocity and the Inertia Tensor

Assuming that the rigid body is made out of N point particles with position vectors $\mathbf{r}_\alpha$ ($\alpha = 1, 2, \ldots, N$), each of these vectors is represented by three real numbers, $x_i^{(\alpha)}$ ($i = 1, 2, 3$), with respect to the inertial frame defined by the basis vectors $\mathbf{e}_i$, and by three real numbers, $x_{i'}^{(\alpha)}$ ($i' = 1', 2', 3'$), with respect to the frame fixed in the body, defined by the basis vectors $\mathbf{e}_{i'}$, that is,

$$\mathbf{r}_\alpha = x_i^{(\alpha)} \mathbf{e}_i = x_{i'}^{(\alpha)} \mathbf{e}_{i'}.$$

According to Equation (3.1), using the fact that the inverse of the matrix $(a_{ij'})$ is its transpose, these sets of coordinates are related by

$$x_i^{(\alpha)} = a_{ij'} x_{j'}^{(\alpha)}, \qquad x_{i'}^{(\alpha)} = a_{ji'} x_j^{(\alpha)}, \tag{3.5}$$

with the *same matrix* $(a_{ij'})$ for all the particles of the body (that is, Equations (3.5) hold for $\alpha = 1, 2, \ldots, N$). (Note that the components of *any* vector with respect to the bases formed by the vectors $\mathbf{e}_i$ and $\mathbf{e}_{i'}$ are related in this form.)

Since the basis vectors $\mathbf{e}_{i'}$ are fixed with respect to the body, the coordinates $x_{i'}^{(\alpha)}$ cannot vary with the time. On the other hand, the coordinates $x_i^{(\alpha)}$ will vary with the time as a consequence of the rotation of the body, hence, making use of the first equation in (3.5),

$$\dot{\mathbf{r}}_\alpha = \dot{x}_j^{(\alpha)} \mathbf{e}_j = \dot{a}_{ji'} x_{i'}^{(\alpha)} \mathbf{e}_j$$

and, therefore, the kinetic energy of the body (with respect to the inertial frame) is

$$\begin{aligned} T &= \sum_{\alpha=1}^{N} \frac{1}{2} m_\alpha \dot{\mathbf{r}}_\alpha \cdot \dot{\mathbf{r}}_\alpha \\ &= \sum_{\alpha=1}^{N} \frac{1}{2} m_\alpha (\dot{a}_{ji'} x_{i'}^{(\alpha)})(\dot{a}_{jk'} x_{k'}^{(\alpha)}) \\ &= \dot{a}_{ji'} \dot{a}_{jk'} \sum_{\alpha=1}^{N} \frac{1}{2} m_\alpha x_{i'}^{(\alpha)} x_{k'}^{(\alpha)}. \end{aligned} \tag{3.6}$$

In this manner, the kinetic energy is expressed in terms of the time derivative of the matrix $(a_{ij'})$, which depends on how the body moves, and of the nine *constant* real numbers $\sum_{\alpha=1}^{N} \frac{1}{2} m_\alpha x_{i'}^{(\alpha)} x_{k'}^{(\alpha)}$, ($i', k' = 1', 2', 3'$) which are determined by the positions and masses of the particles forming the rigid body. As we shall see, the product of time derivatives $\dot{a}_{ji'} \dot{a}_{jk'}$, appearing in (3.6), can be written in terms of a single vector, which corresponds to the body's angular velocity and, instead of the sums $\sum_{\alpha=1}^{N} \frac{1}{2} m_\alpha x_{i'}^{(\alpha)} x_{k'}^{(\alpha)}$, it will be more convenient to use the components of the inertia tensor, to be defined below.

Differentiating both sides of Equation (3.3) with respect to the time we obtain

$$0 = a_{ki'} \dot{a}_{kj'} + \dot{a}_{ki'} a_{kj'}.$$

Since the factors appearing in the last equation are real-valued functions, we have $a_{ki'} \dot{a}_{kj'} = \dot{a}_{kj'} a_{ki'}$ and, therefore,

$$\dot{a}_{ki'} a_{kj'} = -\dot{a}_{kj'} a_{ki'}$$

which shows that the product $\dot{a}_{ki'} a_{kj'}$ is antisymmetric in the indices i' and j'. This is equivalent to say that there exist three real-valued functions of the time, $\omega_{i'}$, such that

$$\dot{a}_{ki'} a_{kj'} = \varepsilon_{i'j's'} \omega_{s'}, \tag{3.7}$$

where $\varepsilon_{i'j'k'}$ is the Levi-Civita symbol, defined by

$$\varepsilon_{i'j'k'} = \begin{cases} 1 \text{ if } i'j'k' \text{ is an even permutation of } 1'2'3' \\ -1 \text{ if } i'j'k' \text{ is an odd permutation of } 1'2'3' \\ 0 \text{ otherwise.} \end{cases} \tag{3.8}$$

The $\omega_{i'}$ are the components of the angular velocity of the rigid body in the basis $\mathbf{e}_{i'}$.

In terms of the matrix notation, Equation (3.7) is equivalent to

$$\dot{A}^{\mathrm{t}} A = \begin{pmatrix} 0 & \omega_{3'} & -\omega_{2'} \\ -\omega_{3'} & 0 & \omega_{1'} \\ \omega_{2'} & -\omega_{1'} & 0 \end{pmatrix}. \tag{3.9}$$

For instance, in the case of the matrix (3.2) we find that the product $\dot{A}^{\mathrm{t}} A$ is

$$\dot{\phi} \begin{pmatrix} -\sin\phi & \cos\phi & 0 \\ -\cos\phi & -\sin\phi & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} = \dot{\phi} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

which shows that the only nonzero component $\omega_{i'}$ is $\omega_{3'} = \dot{\phi}$, as one would expect.

The Levi-Civita symbol (3.8) is invariant under cyclic permutations of the indices, that is,

$$\varepsilon_{i'j'k'} = \varepsilon_{j'k'i'} = \varepsilon_{k'i'j'} \tag{3.10}$$

and satisfies the relation

$$\varepsilon_{i'j'k'} \varepsilon_{i'l'm'} = \delta_{j'l'} \delta_{k'm'} - \delta_{j'm'} \delta_{k'l'}. \tag{3.11}$$

The Levi-Civita symbol is very useful owing to its relationship with the determinant. If $B = (b_{ij})$ is a 3×3 matrix, then from the definition (3.8) it follows that

$$\varepsilon_{ijk} b_{ip} b_{jq} b_{kr} = (\det B) \varepsilon_{pqr}. \tag{3.12}$$

A related result is that the components of the vector product of two vectors can be conveniently expressed with the aid of the Levi-Civita symbol. If a_i and b_i are the components of two vectors, **a** and **b**, respectively, with respect to some right-handed orthonormal basis, then the components of the vector product $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ with respect to this basis are given by

$$c_i = \varepsilon_{ijk} a_j b_k. \tag{3.13}$$

Making use of Equations (3.4), from (3.7) we have

$$\varepsilon_{i'j's'} \omega_{s'} a_{rj'} = \dot{a}_{ki'} a_{kj'} a_{rj'} = \dot{a}_{ki'} \delta_{kr} = \dot{a}_{ri'}, \tag{3.14}$$

hence, with the aid of (3.3)

$$\begin{aligned} \dot{a}_{ji'} \dot{a}_{jk'} &= \varepsilon_{i'l's'} \omega_{s'} a_{jl'} \varepsilon_{k'n'r'} \omega_{r'} a_{jn'} = \varepsilon_{i'l's'} \omega_{s'} \varepsilon_{k'n'r'} \omega_{r'} \delta_{l'n'} = \varepsilon_{i'l's'} \omega_{s'} \varepsilon_{k'l'r'} \omega_{r'} \\ &= \varepsilon_{l's'i'} \omega_{s'} \varepsilon_{l'r'k'} \omega_{r'} = (\delta_{s'r'} \delta_{i'k'} - \delta_{s'k'} \delta_{i'r'})\, \omega_{s'} \omega_{r'}. \end{aligned}$$

Substituting this expression into (3.6) we find that the kinetic energy of the rigid body can also be written in the form

$$\begin{aligned} T &= (\delta_{s'r'} \delta_{i'k'} - \delta_{s'k'} \delta_{i'r'})\, \omega_{s'} \omega_{r'} \sum_{\alpha=1}^{N} \frac{1}{2} m_\alpha x_{i'}^{(\alpha)} x_{k'}^{(\alpha)} \\ &= \frac{1}{2} \omega_{s'} \omega_{r'} \sum_{\alpha=1}^{N} m_\alpha (\delta_{s'r'} \delta_{i'k'} - \delta_{s'k'} \delta_{i'r'})\, x_{i'}^{(\alpha)} x_{k'}^{(\alpha)} \\ &= \frac{1}{2} \omega_{s'} \omega_{r'} \sum_{\alpha=1}^{N} m_\alpha \big(\delta_{s'r'} \mathbf{r}_\alpha{}^2 - x_{s'}^{(\alpha)} x_{r'}^{(\alpha)} \big). \end{aligned}$$

The nine real numbers

$$I_{j'k'} \equiv \sum_{\alpha=1}^{N} m_\alpha \big(\delta_{j'k'} \mathbf{r}_\alpha{}^2 - x_{j'}^{(\alpha)} x_{k'}^{(\alpha)} \big) \tag{3.15}$$

are the components of the *inertia tensor* of the rigid body (with respect to the basis vectors $\mathbf{e}_{i'}$) so that the kinetic energy of the rigid body is expressed as

$$T = \tfrac{1}{2} I_{j'k'} \omega_{j'} \omega_{k'}. \tag{3.16}$$

If we consider a continuous distribution of matter, with a mass density ρ, the components of the inertia tensor are given by

$$I_{j'k'} \equiv \int \rho(\mathbf{r})(\delta_{j'k'}\mathbf{r}^2 - x_{j'}x_{k'})\,\mathrm{d}v', \tag{3.17}$$

with $\mathbf{r} = x_{i'}\mathbf{e}_{i'}$ and $\mathrm{d}v' = \mathrm{d}x_{1'}\mathrm{d}x_{2'}\mathrm{d}x_{3'}$. The definitions (3.15) and (3.17) show that the inertia tensor is symmetric, $I_{i'j'} = I_{j'i'}$, and therefore it has six independent components only.

Expression (3.16) can also be obtained in a more elementary manner. However, the procedure followed above yields a useful expression for the angular velocity of the rigid body in terms of the matrix $(a_{ji'})$ [Equation (3.7)].

Example 3.1 (Inertia tensor of a homogeneous cylinder). We shall calculate the inertia tensor of a homogeneous right circular cylinder. The height of the cylinder will be denoted by h, its radius by a, and its mass by M. Then its density, assumed constant, is $\rho = M/(\pi a^2 h)$. We take the fixed point, O, at the center of the cylinder and the $x_{3'}$-axis will coincide with the axis of the cylinder (see Figure 3.3). From Equation (3.17) we have, making use of cylindrical coordinates (that is, $x_{1'} = \rho\cos\phi$, $x_{2'} = \rho\sin\phi$, $x_{3'} = z$),

$$\begin{aligned}
I_{1'1'} &= \frac{M}{\pi a^2 h}\int_0^a \mathrm{d}\rho \int_{-h/2}^{h/2} \mathrm{d}z \int_0^{2\pi} \rho\mathrm{d}\phi\,({x_{2'}}^2 + {x_{3'}}^2) \\
&= \frac{M}{\pi a^2 h}\int_0^a \mathrm{d}\rho \int_{-h/2}^{h/2} \mathrm{d}z \int_0^{2\pi} \rho\mathrm{d}\phi\,(\rho^2\sin^2\phi + z^2) \\
&= \frac{M}{\pi a^2 h}\int_0^a \mathrm{d}\rho \int_{-h/2}^{h/2} \mathrm{d}z\,\left(\pi\rho^3 + 2\pi\rho z^2\right) \\
&= \frac{M}{a^2 h}\int_0^a \mathrm{d}\rho\left(\rho^3 h + \frac{1}{6}\rho h^3\right) \\
&= M\left(\frac{a^2}{4} + \frac{h^2}{12}\right).
\end{aligned}$$

Similarly we find that

$$I_{2'2'} = \frac{M}{\pi a^2 h}\int_0^a \mathrm{d}\rho \int_{-h/2}^{h/2} \mathrm{d}z \int_0^{2\pi} \rho\mathrm{d}\phi\,({x_{1'}}^2 + {x_{3'}}^2) = M\left(\frac{a^2}{4} + \frac{h^2}{12}\right),$$

and

$$I_{3'3'} = \frac{M}{\pi a^2 h}\int_0^a \mathrm{d}\rho \int_{-h/2}^{h/2} \mathrm{d}z \int_0^{2\pi} \rho\mathrm{d}\phi\,({x_{1'}}^2 + {x_{2'}}^2) = \frac{Ma^2}{2}.$$

Making use of the parity of the integrands one finds that $I_{1'2'}$, $I_{1'3'}$, and $I_{2'3'}$ are equal to zero. Hence,

$$(I_{i'j'}) = \frac{M}{12}\begin{pmatrix} 3a^2 + h^2 & 0 & 0 \\ 0 & 3a^2 + h^2 & 0 \\ 0 & 0 & 6a^2 \end{pmatrix}. \tag{3.18}$$

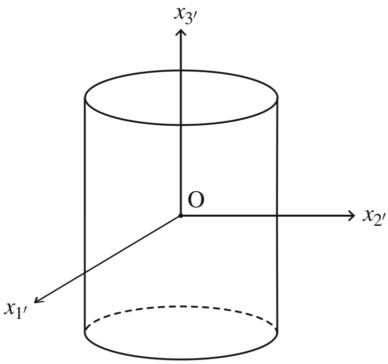


Fig. 3.3 The mass density of the cylinder is uniform. The origin is placed at the geometric center of the cylinder, which coincides with the center of mass. The $x_{3'}$-axis coincides with the axis of the cylinder

Principal Moments of Inertia

The components of the inertia tensor depend on the mass distribution of the body, but also on the choice of the coordinate axes fixed in the body. If we replace the orthonormal basis $\mathbf{e}_{i'}$ by another orthonormal basis $\tilde{\mathbf{e}}_{j'}$ also fixed in the body, then there exists a (time-independent) orthogonal matrix, $B = (b_{j'k'})$, such that

$$\tilde{\mathbf{e}}_{j'} = b_{k'j'}\mathbf{e}_{k'} \tag{3.19}$$

(again, note the position of the indices). According to this definition, the columns of the orthogonal matrix

$$B = \begin{pmatrix} b_{1'1'} & b_{1'2'} & b_{1'3'} \\ b_{2'1'} & b_{2'2'} & b_{2'3'} \\ b_{3'1'} & b_{3'2'} & b_{3'3'} \end{pmatrix}$$

are the components of the vectors $\tilde{\mathbf{e}}_{i'}$ with respect to the basis formed by the vectors $\mathbf{e}_{j'}$. Then, the Cartesian coordinates of the α-th particle with respect to the new axes fixed in the body are $\tilde{x}_{k'}^{(\alpha)} = b_{j'k'}x_{j'}^{(\alpha)}$ and, therefore, with respect to these new axes, the components of the inertia tensor are [see (3.15)]

$$\tilde{I}_{i'j'} = b_{r'i'}b_{s'j'}I_{r's'}. \tag{3.20}$$

In terms of matrices, this last equation amounts to

$$\tilde{\mathscr{I}} = B^{\mathrm{t}}\mathscr{I}B,$$

where $\tilde{\mathscr{I}} \equiv (\tilde{I}_{i'j'})$, $\mathscr{I} \equiv (I_{i'j'})$, and B is the orthogonal matrix defined above.

If the columns of the matrix B are three mutually orthogonal unit eigenvectors of the matrix $(I_{i'j'})$, then $(\tilde{I}_{i'j'})$ is diagonal (see below).

Since the matrix $(I_{r's'})$ is real and symmetric, we can always find three mutually orthogonal eigenvectors of $(I_{r's'})$. Recall that $v_{s'}$ is an eigenvector of $(I_{r's'})$, with eigenvalue λ, if

$$\begin{pmatrix} I_{1'1'} & I_{1'2'} & I_{1'3'} \\ I_{2'1'} & I_{2'2'} & I_{2'3'} \\ I_{3'1'} & I_{3'2'} & I_{3'3'} \end{pmatrix} \begin{pmatrix} v_{1'} \\ v_{2'} \\ v_{3'} \end{pmatrix} = \lambda \begin{pmatrix} v_{1'} \\ v_{2'} \\ v_{3'} \end{pmatrix} \tag{3.21}$$

or, equivalently,

$$I_{r's'} v_{s'} = \lambda v_{r'}. \tag{3.22}$$

The eigenvalue λ is a root of the characteristic polynomial of $(I_{r's'})$, that is, $\det(I_{r's'} - \lambda\delta_{r's'}) = 0$. Since the components $I_{r's'}$ are real, (3.22) is equivalent to

$$I_{r's'} \overline{v_{s'}} = \overline{\lambda}\, \overline{v_{r'}}, \tag{3.23}$$

where the bar denotes complex conjugation. By combining Equations (3.22) and (3.23), using the symmetry of $I_{r's'}$, we obtain

$$I_{i'j'} v_{j'} \overline{v_{i'}} = \lambda v_{i'} \overline{v_{i'}}$$

and

$$I_{i'j'} v_{j'} \overline{v_{i'}} = v_{j'} I_{j'i'} \overline{v_{i'}} = v_{j'} \overline{\lambda}\, \overline{v_{j'}} = \overline{\lambda} v_{i'} \overline{v_{i'}},$$

thus, $(\lambda - \overline{\lambda}) v_{i'} \overline{v_{i'}} = 0$, which means that λ is real ($v_{i'} \overline{v_{i'}}$ is equal to zero only if $v_{i'} = 0$, which is excluded from the definition of eigenvector). Furthermore, the eigenvectors corresponding to different eigenvalues are orthogonal to each other: if $w_{i'}$ is an eigenvector of $I_{i'j'}$ with eigenvalue μ, $I_{i'j'} w_{j'} = \mu w_{i'}$, then, proceeding as above,

$$I_{i'j'} v_{i'} w_{j'} = v_{i'} \mu w_{i'},$$

and

$$I_{i'j'} v_{i'} w_{j'} = w_{j'} I_{j'i'} v_{i'} = w_{j'} \lambda v_{j},$$

which leads to $(\lambda - \mu) v_{i'} w_{i'} = 0$, showing that if $\lambda \neq \mu$ then $v_{i'} w_{i'} = 0$, i.e., the vectors $v_{i'}$ and $w_{i'}$ are orthogonal to each other. Thus, if the three eigenvalues of $(I_{i'j'})$ are distinct, then the corresponding unit eigenvectors form an orthonormal basis.

When only two eigenvalues of $(I_{r's'})$ coincide, the corresponding eigenvectors form a two-dimensional plane, and any pair of orthogonal unit vectors of this plane will be part of an orthonormal basis formed by eigenvectors of $(I_{r's'})$. When the three eigenvalues of $(I_{r's'})$ coincide, then $(I_{r's'})$ is a multiple of the identity matrix and any orthonormal basis is formed by eigenvectors of $(I_{r's'})$.

In conclusion, in all cases we can find an orthonormal basis formed by eigenvectors of $(I_{r's'})$, and from (3.22) it follows that if the columns of the matrix B are three mutually orthogonal unit eigenvectors of the matrix $(I_{i'j'})$, then $(\tilde{I}_{i'j'})$ is diagonal.

If the matrix $(\tilde{I}_{i'j'})$ is diagonal, the directions defined by the basis vectors $\tilde{\mathbf{e}}_{i'}$ are called *principal axes* at O and the entries of $(\tilde{I}_{i'j'})$ along the diagonal are called *principal moments of inertia* [hence, the principal moments of inertia are the eigenvalues of $(I_{i'j'})$].

In Example 3.1 the matrix $(I_{r's'})$ is already diagonal, which means that the basis vectors $\mathbf{e}_{i'}$ point along the principal axes. The entries along the diagonal of $(I_{r's'})$ are the principal moments of inertia and, therefore, at least two principal moments of inertia coincide (the three principal moments of inertia coincide when $h = \sqrt{3}\,a$).

On the other hand, if we place the origin at the base of the cylinder, with the axes as shown in Figure 3.4, the matrix $(I_{i'j'})$ is given by (see Equation (3.28), below)

$$(I_{i'j'}) = \frac{M}{12}\begin{pmatrix} 3a^2+4h^2 & 0 & -6ah \\ 0 & 15a^2+4h^2 & 0 \\ -6ah & 0 & 18a^2 \end{pmatrix}.$$

The eigenvalues of $(I_{i'j'})$ (and, hence, the principal moments of inertia) are the roots of the polynomial

$$\begin{vmatrix} \frac{M}{12}(3a^2+4h^2)-\lambda & 0 & -\frac{M}{2}ah \\ 0 & \frac{M}{12}(15a^2+4h^2)-\lambda & 0 \\ -\frac{M}{2}ah & 0 & \frac{3M}{2}a^2-\lambda \end{vmatrix} = 0,$$

thus, the principal moments of inertia are

$$\frac{M}{12}(15a^2+4h^2), \qquad \frac{M}{24}\left(21a^2+4h^2 \pm \sqrt{225a^4+24a^2h^2+16h^4}\right).$$

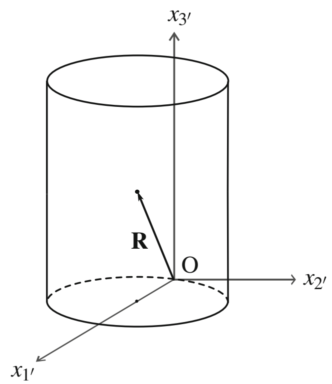


Fig. 3.4 The fixed point, O, is at the edge of the base of the cylinder. The $x_{1'}$-axis passes through the center of the base. The vector **R** goes from the origin to the center of mass

In order to simplify the expressions below, we shall consider the specific case where $h = \sqrt{3}\,a$, then, the principal moments of inertia are

$$\frac{9Ma^2}{4}, \qquad \frac{9Ma^2}{4}, \qquad \frac{Ma^2}{2}$$

and the matrix $(I_{i'j'})$ becomes

$$(I_{i'j'}) = \frac{Ma^2}{4}\begin{pmatrix} 5 & 0 & -2\sqrt{3} \\ 0 & 9 & 0 \\ -2\sqrt{3} & 0 & 6 \end{pmatrix}.$$

The eigenvectors of this matrix corresponding to the (repeated) eigenvalue $9Ma^2/4$ are determined by the homogeneous system of linear equations [see (3.21)]

$$\frac{Ma^2}{4}\begin{pmatrix} 5 & 0 & -2\sqrt{3} \\ 0 & 9 & 0 \\ -2\sqrt{3} & 0 & 6 \end{pmatrix}\begin{pmatrix} v_{1'} \\ v_{2'} \\ v_{3'} \end{pmatrix} = \frac{9Ma^2}{4}\begin{pmatrix} v_{1'} \\ v_{2'} \\ v_{3'} \end{pmatrix},$$

which gives $2v_{1'} + \sqrt{3}\,v_{3'} = 0$, with $v_{2'}$ arbitrary. These conditions define a two-dimensional plane, and two solutions of these conditions are $v_{1'} = -\sqrt{3/7}$, $v_{2'} = 0$, $v_{3'} = 2/\sqrt{7}$, and $v_{1'} = 0$, $v_{2'} = 1$, $v_{3'} = 0$, which correspond to two mutually orthogonal unit vectors. Hence, the two unit vectors

$$\tilde{\mathbf{e}}_{2'} \equiv \mathbf{e}_{2'}, \qquad \tilde{\mathbf{e}}_{3'} \equiv -\sqrt{\frac{3}{7}}\,\mathbf{e}_{1'} + \frac{2}{\sqrt{7}}\,\mathbf{e}_{3'}$$

point along principal axes (the labeling of the vectors $\tilde{\mathbf{e}}_{i'}$ is completely arbitrary).

According to the discussion above, the eigenvectors corresponding to the third eigenvalue, which is different from the first two, must be orthogonal to $\tilde{\mathbf{e}}_{2'}$ and $\tilde{\mathbf{e}}_{3'}$. Thus, we can find the third principal axis by means of the cross product $\tilde{\mathbf{e}}_{2'} \times \tilde{\mathbf{e}}_{3'}$. Letting $\tilde{\mathbf{e}}_{1'} \equiv \tilde{\mathbf{e}}_{2'} \times \tilde{\mathbf{e}}_{3'}$ we obtain the missing element of a positively oriented orthonormal basis such that the vectors $\tilde{\mathbf{e}}_{i'}$ point along principal axes. We find

$$\tilde{\mathbf{e}}_{1'} = \frac{2}{\sqrt{7}}\,\mathbf{e}_{1'} + \sqrt{\frac{3}{7}}\,\mathbf{e}_{3'}.$$

(One can readily verify that the vector given by $v_{1'} = 2/\sqrt{7}$, $v_{2'} = 0$, $v_{3'} = \sqrt{3/7}$ is indeed an eigenvector of $(I_{i'j'})$ with eigenvalue $Ma^2/2$.)

Exercise 3.2. Four particles of mass m are at the points $(a, 0, 0)$, $(0, a, 0)$, $(a, a, 0)$, and $(0, 0, 0)$, with respect to the Cartesian axes $x_{i'}$, where a is a positive constant. Find the principal axes and the principal moments of inertia.

Angular Momentum

From the elementary definition of the angular momentum of a particle ($\mathbf{L} = \mathbf{r} \times m\dot{\mathbf{r}}$) it follows that the Cartesian components of the angular momentum of the rigid body (with respect to the inertial frame) are given by

$$L_i = \sum_{\alpha=1}^{N} m_\alpha \varepsilon_{ijk} x_j^{(\alpha)} \dot{x}_k^{(\alpha)}.$$

According to (3.5) and (3.14), we have

$$\dot{x}_k^{(\alpha)} = \dot{a}_{ki'} x_{i'}^{(\alpha)} = \varepsilon_{i'j's'} \omega_{s'} a_{kj'} x_{i'}^{(\alpha)},$$

hence,

$$\begin{aligned} L_i &= \sum_{\alpha=1}^{N} m_\alpha \varepsilon_{ijk} x_j^{(\alpha)} \varepsilon_{i'j's'} \omega_{s'} a_{kj'} x_{i'}^{(\alpha)} \\ &= \sum_{\alpha=1}^{N} m_\alpha \varepsilon_{ijk} a_{jq'} x_{q'}^{(\alpha)} \varepsilon_{i'j's'} \omega_{s'} a_{kj'} x_{i'}^{(\alpha)}. \end{aligned}$$

Noting that, owing to (3.4) and (3.12),

$$\begin{aligned} \varepsilon_{ijk} a_{jq'} a_{kj'} &= \delta_{ip} \varepsilon_{pjk} a_{jq'} a_{kj'} \\ &= a_{ir'} a_{pr'} \varepsilon_{pjk} a_{jq'} a_{kj'} \\ &= a_{ir'} \varepsilon_{r'q'j'}, \end{aligned} \tag{3.24}$$

where we have used that the determinant of an orthogonal matrix that does not invert the orientation is equal to $+1$ (which follows from $1 = \det I = \det(A^{\mathrm{t}} A) = \det A^{\mathrm{t}} \det A = (\det A)^2$), with the aid of (3.10) and (3.11), we have

$$\begin{aligned} L_i &= \sum_{\alpha=1}^{N} m_\alpha a_{ir'} \varepsilon_{r'q'j'} \varepsilon_{i'j's'} \omega_{s'} x_{q'}^{(\alpha)} x_{i'}^{(\alpha)} \\ &= a_{ir'} \sum_{\alpha=1}^{N} m_\alpha (\delta_{r's'} \delta_{q'i'} - \delta_{r'i'} \delta_{q's'}) \omega_{s'} x_{q'}^{(\alpha)} x_{i'}^{(\alpha)} \\ &= a_{ir'} \sum_{\alpha=1}^{N} m_\alpha (\delta_{r's'} \mathbf{r}_\alpha{}^2 - x_{r'}^{(\alpha)} x_{s'}^{(\alpha)}) \, \omega_{s'} \\ &= a_{ir'} I_{r's'} \omega_{s'}. \end{aligned}$$

This means that $I_{r's'} \omega_{s'}$ is the r-th Cartesian component of the angular momentum of the body with respect to the basis vectors $\mathbf{e}_{i'}$,

$$L_{r'} = I_{r's'} \omega_{s'}. \tag{3.25}$$

Among other things, Equation (3.25) means that the angular velocity and the angular momentum may not be collinear, but when $\omega_{s'}$ is an eigenvector of the matrix $(I_{r's'})$ then the angular momentum and the angular velocity are collinear. Thus, the principal axes are the directions where the angular momentum and the angular velocity are collinear.

Exercise 3.3. Show that

$$A\dot{A}^{\rm t} = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix},$$

where the ω_i are the components of the angular velocity with respect to the basis vectors $\mathbf{e}_i$ (the inertial frame), that is,

$$a_{is'}\dot{a}_{js'} = \varepsilon_{ijk}\omega_k$$

[cf. Equation (3.7)].

Parallel Axes Theorem

Now we shall study the behavior of the inertia tensor under a parallel translation of the axes fixed in the body. To this end, it is convenient to consider two parallel sets of Cartesian axes fixed in the body, $x_{i'}$ and $y_{i'}$ (see Figure 3.5), defined in the following way. The origin of the axes $x_{i'}$ is located at an arbitrary point of the body, while the origin of the axes $y_{i'}$ is at the center of mass of the rigid body, this means that

$$\sum_{\alpha=1}^{N} m_\alpha y_{i'}^{(\alpha)} = 0, \qquad i' = 1', 2', 3', \tag{3.26}$$

where $(y_{1'}^{(\alpha)}, y_{2'}^{(\alpha)}, y_{3'}^{(\alpha)})$ are the Cartesian coordinates of the α-th particle of the body with respect to the axes with origin at the center of mass.

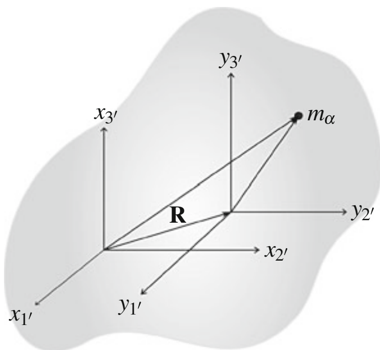


Fig. 3.5 The Cartesian axes $x_{i'}$ have their origin at an arbitrary point fixed in the body, and the origin of the Cartesian axes $y_{i'}$ is at the center of mass. The vector **R** is the position vector of the center of mass with respect to the axes $x_{i'}$. The axes $y_{i'}$ are parallel to the axes $x_{i'}$

If $(R_{1'}, R_{2'}, R_{3'})$ are the coordinates of the center of mass with respect to the axes $x_{i'}$, we have (see Figure 3.5)

$$x_{i'}^{(\alpha)} = R_{i'} + y_{i'}^{(\alpha)}, \qquad \alpha = 1, 2, \ldots, N;\ i' = 1', 2', 3'.$$

Hence, making use of the definition of the (components of the) inertia tensor (3.15) and Equation (3.26), we have

$$\begin{aligned}
I_{j'k'} &= \sum_{\alpha=1}^{N} m_\alpha \left(\delta_{j'k'} x_{i'}^{(\alpha)} x_{i'}^{(\alpha)} - x_{j'}^{(\alpha)} x_{k'}^{(\alpha)}\right) \\
&= \sum_{\alpha=1}^{N} m_\alpha \left[\delta_{j'k'} (R_{i'} + y_{i'}^{(\alpha)})(R_{i'} + y_{i'}^{(\alpha)}) - (R_{j'} + y_{j'}^{(\alpha)})(R_{k'} + y_{k'}^{(\alpha)})\right] \\
&= (R_{i'} R_{i'} \delta_{j'k'} - R_{j'} R_{k'}) \sum_{\alpha=1}^{N} m_\alpha + \sum_{\alpha=1}^{N} m_\alpha \left(\delta_{j'k'} y_{i'}^{(\alpha)} y_{i'}^{(\alpha)} - y_{j'}^{(\alpha)} y_{k'}^{(\alpha)}\right) \\
&= M(\mathbf{R}^2 \delta_{j'k'} - R_{j'} R_{k'}) + I_{j'k'}^{\text{CM}},
\end{aligned} \tag{3.27}$$

where M is the total mass of the body, $\mathbf{R}^2 = R_{i'} R_{i'}$ is the square of the norm of the vector $(R_{1'}, R_{2'}, R_{3'})$, and the $I_{j'k'}^{\text{CM}}$ are the components of the inertia tensor, taking the center of mass as the fixed point of the rigid body. This result is known as the parallel axes theorem.

For instance, considering again the homogeneous circular cylinder of Example 3.1, with the aid of Equation (3.27) we can readily obtain the components of the inertia tensor taking one point at the edge of the base of the cylinder as the fixed point (see Figure 3.4). If the $x_{1'}$-axis lies along a diameter of the base of the cylinder, then $(R_{1'}, R_{2'}, R_{3'}) = (a, 0, h/2)$ and from Equations (3.27) and (3.18) we find

$$\begin{aligned}
(I_{i'j'}) &= \frac{M}{4} \begin{pmatrix} h^2 & 0 & -2ah \\ 0 & 4a^2 + h^2 & 0 \\ -2ah & 0 & 4a^2 \end{pmatrix} + \frac{M}{12} \begin{pmatrix} 3a^2 + h^2 & 0 & 0 \\ 0 & 3a^2 + h^2 & 0 \\ 0 & 0 & 6a^2 \end{pmatrix} \\
&= \frac{M}{12} \begin{pmatrix} 3a^2 + 4h^2 & 0 & -6ah \\ 0 & 15a^2 + 4h^2 & 0 \\ -6ah & 0 & 18a^2 \end{pmatrix}.
\end{aligned} \tag{3.28}$$

Exercise 3.4. Show that if the line joining O and the center of mass is parallel to one of the principal axes at the center of mass, then this line is also parallel to a principal axis at O. Furthermore, any principal axis at the center of mass orthogonal to the line is parallel to a principal axis at O.

Coordinate-Free Expression of the Lagrange Equations. The Euler Equations
So far, we have not required the introduction of coordinates to parameterize the configuration of the rigid body, and as we shall see below and in Section 4.2, there are some results that can be obtained without giving an explicit expression for the matrix elements $a_{ij'}$ in terms of coordinates.

Assuming that the $a_{ij'}$ are parameterized by some coordinates q_s, from (3.7) and the chain rule we have

$$\frac{\partial a_{ki'}}{\partial q_r}\dot{q}_r a_{kj'} = \varepsilon_{i'j's'}\omega_{s'},$$

then, making use of (3.11),

$$\omega_{s'} = \frac{1}{2}\varepsilon_{i'j's'}\frac{\partial a_{ki'}}{\partial q_r}\dot{q}_r a_{kj'} = M_{s'r}\dot{q}_r, \tag{3.29}$$

where we have introduced the functions

$$M_{s'r} \equiv \frac{1}{2}\varepsilon_{i'j's'}\frac{\partial a_{ki'}}{\partial q_r}a_{kj'}, \tag{3.30}$$

which depend on the coordinates q_s only, and relate the angular velocity with the generalized velocities $\dot{q}_i$. The last equation is equivalent to the relation

$$\frac{\partial a_{ki'}}{\partial q_r} = a_{kj'}\varepsilon_{i'j's'}M_{s'r}. \tag{3.31}$$

According to Equations (3.31), (3.10), and (3.11) the second partial derivatives of $a_{ki'}$ are given by

$$\begin{aligned}
\frac{\partial^2 a_{ki'}}{\partial q_m \partial q_r} &= \frac{\partial a_{kj'}}{\partial q_m}\varepsilon_{i'j's'}M_{s'r} + a_{kj'}\varepsilon_{i'j's'}\frac{\partial M_{s'r}}{\partial q_m} \\
&= a_{kp'}\varepsilon_{j'p'n'}M_{n'm}\varepsilon_{i'j's'}M_{s'r} + a_{kj'}\varepsilon_{i'j's'}\frac{\partial M_{s'r}}{\partial q_m} \\
&= (\delta_{p's'}\delta_{n'i'} - \delta_{p'i'}\delta_{n's'})a_{kp'}M_{n'm}M_{s'r} + a_{kj'}\varepsilon_{i'j's'}\frac{\partial M_{s'r}}{\partial q_m} \\
&= a_{ks'}M_{i'm}M_{s'r} - a_{ki'}M_{s'm}M_{s'r} + a_{kj'}\varepsilon_{i'j's'}\frac{\partial M_{s'r}}{\partial q_m}.
\end{aligned}$$

Then, the commutativity of the partial derivatives of $a_{ki'}$ is equivalent to

$$\begin{aligned}
a_{kj'}\varepsilon_{i'j's'}\left(\frac{\partial M_{s'r}}{\partial q_m} - \frac{\partial M_{s'm}}{\partial q_r}\right) &= a_{ks'}(M_{i'r}M_{s'm} - M_{i'm}M_{s'r}) \\
&= a_{kj'}(M_{i'r}M_{j'm} - M_{i'm}M_{j'r}),
\end{aligned}$$

hence,

$$\varepsilon_{i'j's'}\left(\frac{\partial M_{s'r}}{\partial q_m} - \frac{\partial M_{s'm}}{\partial q_r}\right) = M_{i'r}M_{j'm} - M_{i'm}M_{j'r}.$$

With the aid of (3.11) one finds that the last equation is equivalent to

$$\frac{\partial M_{s'r}}{\partial q_m} - \frac{\partial M_{s'm}}{\partial q_r} = \varepsilon_{i'j's'} M_{i'r} M_{j'm}. \tag{3.32}$$

Note that these equations must hold for any choice of the coordinates q_i. (It turns out that Equations (3.32) are related to the structure of the rotation group itself.)

Assuming that the applied forces on the body are derivable from a potential $V(q_i)$, the equations of motion of the rigid body can be obtained substituting the Lagrangian [see Equation (3.16)]

$$L = \tfrac{1}{2} I_{j'k'} \omega_{j'} \omega_{k'} - V \tag{3.33}$$

into the Lagrange equations. Making use of the symmetry of $I_{j'k'}$, (3.25), (3.29), and (3.32), we obtain

$$\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial \dot{q}_i} \left(\tfrac{1}{2} I_{j'k'} \omega_{j'} \omega_{k'} \right) - \frac{\partial}{\partial q_i} \left(\tfrac{1}{2} I_{j'k'} \omega_{j'} \omega_{k'} - V \right) \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \left(I_{j'k'} \omega_{k'} \frac{\partial \omega_{j'}}{\partial \dot{q}_i} \right) - I_{j'k'} \omega_{k'} \frac{\partial \omega_{j'}}{\partial q_i} + \frac{\partial V}{\partial q_i} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \left(L_{j'} M_{j'i} \right) - L_{j'} \frac{\partial M_{j'k}}{\partial q_i} \dot{q}_k + \frac{\partial V}{\partial q_i} \\
&= M_{j'i} \frac{\mathrm{d}L_{j'}}{\mathrm{d}t} + L_{j'} \frac{\partial M_{j'i}}{\partial q_k} \dot{q}_k - L_{j'} \frac{\partial M_{j'k}}{\partial q_i} \dot{q}_k + \frac{\partial V}{\partial q_i} \\
&= M_{j'i} \frac{\mathrm{d}L_{j'}}{\mathrm{d}t} + L_{j'} \left(\frac{\partial M_{j'i}}{\partial q_k} - \frac{\partial M_{j'k}}{\partial q_i} \right) \dot{q}_k + \frac{\partial V}{\partial q_i} \\
&= M_{j'i} \frac{\mathrm{d}L_{j'}}{\mathrm{d}t} + L_{j'} \varepsilon_{r's'j'} M_{r'i} M_{s'k} \dot{q}_k + \frac{\partial V}{\partial q_i} \\
&= M_{r'i} \left(\frac{\mathrm{d}L_{r'}}{\mathrm{d}t} + \varepsilon_{r's'j'} \omega_{s'} L_{j'} \right) + \frac{\partial V}{\partial q_i},
\end{aligned}$$

that is,

$$\frac{\mathrm{d}L_{r'}}{\mathrm{d}t} + \varepsilon_{r's'j'} \omega_{s'} L_{j'} = -(M^{-1})_{ir'} \frac{\partial V}{\partial q_i}, \tag{3.34}$$

where the $(M^{-1})_{ir'}$ are the entries of the inverse of the matrix $(M_{r'i})$. The right-hand side of (3.34) is the r-th component of the torque on the rigid body, $\tau_{r'}$, with respect to the basis fixed in the body. In fact, with the aid of Equations (3.14) and (3.10) we find that

$$\begin{aligned}\frac{\mathrm{d}L_i}{\mathrm{d}t} &= \frac{\mathrm{d}(a_{ir'}L_{r'})}{\mathrm{d}t}\\ &= a_{ir'}\frac{\mathrm{d}L_{r'}}{\mathrm{d}t} + \varepsilon_{r'k's'}\omega_{s'}a_{ik'}L_{r'}\\ &= a_{ir'}\left(\frac{\mathrm{d}L_{r'}}{\mathrm{d}t} + \varepsilon_{r's'k'}\omega_{s'}L_{k'}\right),\end{aligned}$$

which shows that, indeed, the left-hand side of (3.34) is the r-th component of the torque on the rigid body with respect to the basis fixed in the body. Equations (3.34), written in the form

$$\frac{\mathrm{d}L_{r'}}{\mathrm{d}t} + \varepsilon_{r's'j'}\omega_{s'}L_{j'} = \tau_{r'},$$

are known as the Euler equations for a rigid body with a fixed point. As we have shown, these equations are equivalent to the Lagrange equations for the Lagrangian (3.33).

A particular case corresponds to the motion of the rigid body with the torque equal to zero. Then, the Euler equations reduce to

$$\frac{\mathrm{d}L_{r'}}{\mathrm{d}t} + \varepsilon_{r's'j'}\omega_{s'}L_{j'} = 0.$$

If the matrix $(I_{i'j'})$ is diagonal, these equations expressed in terms of the components of the angular velocity $\omega_{i'}$ take the form [see (3.25)]

$$\begin{aligned}I_1\frac{\mathrm{d}\omega_{1'}}{\mathrm{d}t} + (I_3 - I_2)\omega_{2'}\omega_{3'} &= 0,\\ I_2\frac{\mathrm{d}\omega_{2'}}{\mathrm{d}t} + (I_1 - I_3)\omega_{3'}\omega_{1'} &= 0,\\ I_3\frac{\mathrm{d}\omega_{3'}}{\mathrm{d}t} + (I_2 - I_1)\omega_{1'}\omega_{2'} &= 0,\end{aligned} \tag{3.35}$$

where the I_i are the principal moments of inertia ($I_1 \equiv I_{1'1'}$, $I_2 \equiv I_{2'2'}$, $I_3 \equiv I_{3'3'}$).

Exercise 3.5. A rigid body is symmetric if two of its principal moments of inertia coincide. Solve Equations (3.35) for a symmetric rigid body. Note that this solution only gives the angular velocity as a function of time; in order to find the configuration of the body we would still have to solve another system of ODEs [e.g., (3.14) or (3.38)].

Exercise 3.6. Making use of the Euler equations (3.35), show that the kinetic energy and the total angular momentum of the rigid body are constants of motion, that is,

$$\tfrac{1}{2}(I_1{\omega_{1'}}^2 + I_2{\omega_{2'}}^2 + I_3{\omega_{3'}}^2) \qquad \text{and} \qquad {I_1}^2{\omega_{1'}}^2 + {I_2}^2{\omega_{2'}}^2 + {I_3}^2{\omega_{3'}}^2$$

are conserved [see (3.16) and (3.25)]. (With the aid of these two constants of motion, Equations (3.35) can be reduced to a single first-order ODE whose solution, in general, involves elliptic functions.)

3.3 The Euler Angles

A usual and convenient set of generalized coordinates for a rigid body with a fixed point is given by the so-called Euler angles. The matrix $A = (a_{ij'})$, representing the configuration of a rigid body, is expressed in the form

$$A = R_z(\phi) R_x(\theta) R_z(\psi), \tag{3.36}$$

where $R_z(\phi)$ is the 3×3 orthogonal matrix (3.2), corresponding to a rotation about the z-axis through an angle ϕ and, similarly,

$$R_x(\theta) \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

corresponds to a rotation about the x-axis through an angle θ. The angles ϕ, θ, ϕ are called Euler angles and are restricted by $0 \leqslant \phi \leqslant 2\pi$, $0 \leqslant \theta \leqslant \pi$, $0 \leqslant \psi \leqslant 2\pi$. (A slightly different definition, which is especially convenient in the study of the rotation group in quantum mechanics, is given by $A = R_z(\phi) R_y(\theta) R_z(\psi)$, where $R_y(\theta)$ corresponds to a rotation about the y-axis through an angle θ.)

According to Equation (3.16), in order to write the kinetic energy in terms of the Euler angles and their time derivatives, we need the explicit expression of the components of the angular velocity, $\omega_{i'}$, in terms of those variables. These expressions can be readily obtained with the aid of Equation (3.9) by calculating the product $\dot{A}^{\mathrm{t}} A$ (without having to resort to a geometrical image or to the consideration of "infinitesimal rotations"). Using the fact that $(AB)^{\mathrm{t}} = B^{\mathrm{t}} A^{\mathrm{t}}$ and that each matrix appearing in (3.36) is orthogonal, we find

$$\begin{aligned} \dot{A}^{\mathrm{t}} A &= \left[\dot{R}_z(\phi) R_x(\theta) R_z(\psi) + R_z(\phi) \dot{R}_x(\theta) R_z(\psi) + R_z(\phi) R_x(\theta) \dot{R}_z(\psi)\right]^{\mathrm{t}} \\ &\quad \times R_z(\phi) R_x(\theta) R_z(\psi) \\ &= R_z(\psi)^{\mathrm{t}} R_x(\theta)^{\mathrm{t}} \dot{R}_z(\phi)^{\mathrm{t}} R_z(\phi) R_x(\theta) R_z(\psi) + R_z(\psi)^{\mathrm{t}} \dot{R}_x(\theta)^{\mathrm{t}} R_x(\theta) R_z(\psi) \\ &\quad + \dot{R}_z(\psi)^{\mathrm{t}} R_z(\psi). \end{aligned} \tag{3.37}$$

The last term in Equation (3.37) was already calculated in Section 3.2; the result is

$$\dot{R}_z(\psi)^{\mathrm{t}} R_z(\psi) = \dot{\psi} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

A similar computation gives

$$\dot{R}_x(\theta)^{\mathrm{t}} R_x(\theta) = \dot{\theta} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

and, therefore,

$$R_z(\psi)^{\mathrm{t}} \dot{R}_x(\theta)^{\mathrm{t}} R_x(\theta) R_z(\psi) = \dot{\theta} \begin{pmatrix} 0 & 0 & \sin\psi \\ 0 & 0 & \cos\psi \\ -\sin\psi & -\cos\psi & 0 \end{pmatrix}.$$

A more lengthy computation gives

$$\begin{aligned} & R_z(\psi)^{\mathrm{t}} R_x(\theta)^{\mathrm{t}} \dot{R}_z(\phi)^{\mathrm{t}} R_z(\phi) R_x(\theta) R_z(\psi) \\ & \qquad = \dot{\phi} \begin{pmatrix} 0 & \cos\theta & -\sin\theta\cos\psi \\ -\cos\theta & 0 & \sin\theta\sin\psi \\ \sin\theta\cos\psi & -\sin\theta\sin\psi & 0 \end{pmatrix}. \end{aligned}$$

Adding these expressions and comparing the result with (3.9) we conclude that the components of the angular velocity of the body, with respect to the axes fixed in the body, are

$$\begin{aligned} \omega_{1'} &= \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi, \\ \omega_{2'} &= \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi, \\ \omega_{3'} &= \dot{\phi}\cos\theta + \dot{\psi}. \end{aligned} \tag{3.38}$$

Thus, assuming that the matrix $(I_{i'j'})$ is diagonal, from (3.16) and (3.38) we have the expression for the kinetic energy in terms of the Euler angles

$$\begin{aligned} T = {} & \tfrac{1}{2}[I_1(\dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi)^2 + I_2(\dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi)^2 \\ & + I_3(\dot{\phi}\cos\theta + \dot{\psi})^2]. \end{aligned} \tag{3.39}$$

When two principal moments of inertia coincide, it is convenient to select the axes in such a way that $I_1 = I_2$ because then (3.39) reduces to

$$T = \tfrac{1}{2}[I_1(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2) + I_3(\dot{\phi}\cos\theta + \dot{\psi})^2]. \tag{3.40}$$

Example 3.7 (Symmetric top in a uniform gravitational field). A commonly studied example is that of a symmetric top in a uniform gravitational field. This problem consists of an axially symmetric top with a fixed point, in a uniform gravitational field. Assuming that the x_3-axis points upwards and taking the $x_{3'}$-axis as the symmetry axis of the top, we have $I_1 = I_2$. Then, making use of (3.40) we find that the standard Lagrangian is

$$L = \tfrac{1}{2}[I_1(\dot{\phi}^2 \sin^2\theta + \dot{\theta}^2) + I_3(\dot{\phi}\cos\theta + \dot{\psi})^2] - Mgl\cos\theta, \qquad (3.41)$$

where M is the mass of the top and l is the distance between the fixed point of the body (which is placed at the origin) and the center of mass. (The product $l\cos\theta$ is the height of the center of mass with respect to the origin since, according to (3.1), the components of the vector $\mathbf{e}_{3'}$ with respect to the basis formed by the vectors $\mathbf{e}_i$ appear in the third column of the matrix A [see (3.36)], which is given by

$$R_z(\phi)R_x(\theta)R_z(\psi)\begin{pmatrix}0\\0\\1\end{pmatrix} = R_z(\phi)R_x(\theta)\begin{pmatrix}0\\0\\1\end{pmatrix} = R_z(\phi)\begin{pmatrix}0\\-\sin\theta\\\cos\theta\end{pmatrix}$$

$$= \begin{pmatrix}\cos\phi & -\sin\phi & 0\\ \sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0\\-\sin\theta\\\cos\theta\end{pmatrix} = \begin{pmatrix}\sin\theta\sin\phi\\-\sin\theta\cos\phi\\\cos\theta\end{pmatrix}$$

and the position vector of the center of mass is $l\mathbf{e}_{3'}$.)

As in previous examples, it is not convenient to obtain the equations of motion by substituting the Lagrangian (3.41) into the Lagrange equations and then try to solve them. It is preferable to use the fact that the coordinates ϕ and ψ are ignorable and that the Lagrangian does not depend on t; this implies that

$$\frac{\partial L}{\partial\dot{\phi}} = I_1\dot{\phi}\sin^2\theta + I_3\cos\theta(\dot{\phi}\cos\theta + \dot{\psi}), \qquad \frac{\partial L}{\partial\dot{\psi}} = I_3(\dot{\phi}\cos\theta + \dot{\psi})$$

as well as

$$\tfrac{1}{2}[I_1(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2) + I_3(\dot{\phi}\cos\theta + \dot{\psi})^2] + Mgl\cos\theta$$

are constants of motion. Denoting as a, b, and E, respectively, the values of these constants of motion, the combination of the foregoing expressions leads to the first-order ODE

$$E = \frac{1}{2}\left[I_1\dot{\theta}^2 + \frac{(a - b\cos\theta)^2}{I_1\sin^2\theta} + \frac{b^2}{I_3}\right] + Mgl\cos\theta,$$

which determines θ as a function of the time [cf. Equation (2.32)]. The substitution $u = \cos\theta$ yields the equivalent equation

$$\frac{1}{2}I_1\dot{u}^2 + \frac{(a - bu)^2}{2I_1} + \left(\frac{b^2}{2I_3} - E\right)(1 - u^2) + Mglu(1 - u^2) = 0. \qquad (3.42)$$

Since the "effective potential,"

$$\frac{(a - bu)^2}{2I_1} + \left(\frac{b^2}{2I_3} - E\right)(1 - u^2) + Mglu(1 - u^2),$$

is a third-degree polynomial in u, the solution of (3.42) involves elliptic functions. Alternatively, one can find the qualitative behavior of the solutions with the aid of the graph of the effective potential.

Chapter 4
The Hamiltonian Formalism

In this chapter it is shown that, for a regular Lagrangian, the Lagrange equations can be translated into a set of first-order ODEs, known as the Hamilton, or canonical, equations, which turn out to be more useful than the Lagrange equations, as we shall see in this chapter and in the following ones. In the same manner as the Lagrange equations are defined by a single function (the Lagrangian), the Hamilton equations are defined by a single function, known as the Hamiltonian.

In Section 4.1 we derive the Hamilton equations and we give some few examples of their application. In Section 4.2 we start presenting the advantages of the Hamiltonian formalism, introducing the Poisson bracket and, in Section 4.3, showing that, for certain systems, the use of an appropriate parameter, in place of the time, allows us to simplify the integration of the equations of motion.

4.1 The Hamilton Equations

As we have seen in the preceding chapters, various systems of second-order ODEs, especially the equations of motion of holonomic conservative mechanical systems, can be expressed in the form of the Lagrange equations for some Lagrangian.

On the other hand, any system of n second-order ODEs can be transformed into a system of $2n$ first-order ODEs, in *infinitely many* different ways, by introducing n auxiliary variables. For example, by defining

$$y_i \equiv \dot{x}_i, \tag{4.1}$$

any given system of n second-order ODEs, $\ddot{x}_i = f_i(x_j, \dot{x}_j, t)$, can be written in the equivalent form

$$\dot{x}_i = y_i, \qquad \dot{y}_i = f_i(x_j, y_j, t). \tag{4.2}$$

G. F. Torres del Castillo, *An Introduction to Hamiltonian Mechanics*, Birkhäuser Advanced Texts Basler Lehrbücher, https://doi.org/10.1007/978-3-319-95225-3_4

In the specific case of the Emden–Fowler equation (see Example 2.8)

$$\ddot{x} + \frac{2}{t}\dot{x} + x^k = 0,$$

letting $y \equiv \dot{x}$, we have the equivalent system of two first-order ODEs

$$\dot{x} = y, \qquad \dot{y} = -\frac{2}{t}y - x^k, \tag{4.3}$$

while the definition $z \equiv t^2\dot{x}$, leads to the system

$$\dot{x} = \frac{z}{t^2}, \qquad \dot{z} = -t^2 x^k. \tag{4.4}$$

(Note that the first equation in (4.2), (4.3), and (4.4) comes from the *definition* of the new auxiliary variable.)

As we shall show in this section, if we have a system of second-order ODEs obtained from a regular Lagrangian, then the Lagrangian itself gives us a systematic way of arriving at a system of first-order ODEs (known as the Hamilton equations), which has several useful properties.

Generalized Momentum

The basic ingredient of the Hamiltonian formalism is the definition of the *generalized momenta*. Recall that given a Lagrangian, $L(q_i, \dot{q}_i, t)$, the generalized momenta are defined by

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i} \tag{4.5}$$

($i = 1, 2, \ldots, n$) [cf. Equation (1.78)] and p_i is called the momentum conjugate (or canonically conjugate) to q_i. Each p_i thus defined is some function of q_j, $\dot{q}_j$, and t, and the purpose is to replace the $\dot{q}_i$ by the p_i.

Example 4.1. The equations of motion of a particle of mass m moving in the plane, subject to forces derivable from a potential V, in polar coordinates, are

$$m\ddot{r} - mr\dot{\theta}^2 = -\frac{\partial V}{\partial r}, \qquad mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta} = -\frac{\partial V}{\partial \theta} \tag{4.6}$$

and these equations can be obtained from the standard Lagrangian

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) - V. \tag{4.7}$$

The definition (4.5) yields the canonical momenta

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}, \qquad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}, \tag{4.8}$$

and, making use of these variables, the system (4.6) can be rewritten as the following system of four first-order ODEs

$$\dot{r} = \frac{p_r}{m}, \qquad \dot{\theta} = \frac{p_\theta}{mr^2}, \qquad \dot{p}_r = \frac{{p_\theta}^2}{mr^3} - \frac{\partial V}{\partial r}, \qquad \dot{p}_\theta = -\frac{\partial V}{\partial \theta}. \tag{4.9}$$

The first two of these equations amount to the definition of the canonical momenta (4.8), while the last two equations come from (4.6), eliminating $\dot{r}$ and $\dot{\theta}$ in favor of p_r and p_θ. Equations (4.9) give the derivative with respect to the time of r, θ, p_r, and p_θ as functions of themselves.

The relevant feature of the system of equations (4.9) is that the right-hand sides are the partial derivatives of the function

$$\frac{{p_r}^2}{2m} + \frac{{p_\theta}^2}{2mr^2} + V$$

with respect to p_r, p_θ, $-r$ and $-\theta$, respectively, as one can verify. As we shall see now, a similar result applies to any system of equations given by a regular Lagrangian.

We start from a Lagrangian $L(q_i, \dot{q}_i, t)$ ($i = 1, 2, \ldots, n$). Taking into account the definition (4.5), the total differential of L is given by

$$\mathrm{d}L = \frac{\partial L}{\partial q_i}\mathrm{d}q_i + p_i \mathrm{d}\dot{q}_i + \frac{\partial L}{\partial t}\mathrm{d}t,$$

hence, expressing the term $p_i \mathrm{d}\dot{q}_i$ in the equivalent form $\mathrm{d}(p_i \dot{q}_i) - \dot{q}_i \mathrm{d}p_i$, we have

$$\mathrm{d}\big(p_i \dot{q}_i - L\big) = -\frac{\partial L}{\partial q_i}\mathrm{d}q_i + \dot{q}_i \mathrm{d}p_i - \frac{\partial L}{\partial t}\mathrm{d}t.$$

Thus, defining the Hamiltonian function

$$H \equiv p_i \dot{q}_i - L, \tag{4.10}$$

we have

$$\mathrm{d}H = -\frac{\partial L}{\partial q_i}\mathrm{d}q_i + \dot{q}_i \mathrm{d}p_i - \frac{\partial L}{\partial t}\mathrm{d}t. \tag{4.11}$$

This equation shows that H can be expressed as a function of (q_i, p_i, t) and that *if* the set (q_i, p_i, t) is functionally independent [which amounts to say that the $\dot{q}_i$ can be expressed as functions of (q_j, p_j, t)], then we have the *identities*

$$\left(\frac{\partial H}{\partial q_i}\right)_{q,p,t} = -\left(\frac{\partial L}{\partial q_i}\right)_{q,\dot{q},t}, \quad \left(\frac{\partial H}{\partial p_i}\right)_{q,p,t} = \dot{q}_i, \quad \left(\frac{\partial H}{\partial t}\right)_{q,p} = -\left(\frac{\partial L}{\partial t}\right)_{q,\dot{q}}. \tag{4.12}$$

The subscripts in the parentheses above indicate which variables are kept fixed in the partial differentiation. This notation may be unnecessary when one is using a single set of coordinates, but in this case, as well as in the discussion in Chapter 5, we are making use simultaneously of more than one set of coordinates and, in order to avoid errors, it is convenient to indicate explicitly the coordinates being employed. (Note that, without this clarification, from the first and the last equation in (4.12) one might conclude that $H+L$ is always independent of the q_i and t.) This notation may be familiar from Thermodynamics, where one employs various sets of variables as coordinates. The relationship (4.10), between L and H, is an example of a Legendre transformation.

Making use of the Lagrange equations (1.49) and of the definition (4.5), we see that the first equation in (4.12) amounts to

$$\left(\frac{\partial H}{\partial q_i}\right)_{q,p,t} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)_{q,\dot{q},t} = -\dot{p}_i,$$

so that the Lagrange equations are equivalent to the system of $2n$ first-order ODEs

$$\dot{q}_i = \left(\frac{\partial H}{\partial p_i}\right)_{q,p,t}, \qquad \dot{p}_i = -\left(\frac{\partial H}{\partial q_i}\right)_{q,p,t} \tag{4.13}$$

($i = 1, 2, \ldots, n$), which are known as the *Hamilton equations* (or *Hamilton's canonical equations*). The variables q_i together with their conjugate momenta, p_i, are called *canonical coordinates*.

In many textbooks it is asserted that the Hamiltonian *must be* expressed as a function of (q_i, p_i, t) and that an expression for H involving the $\dot{q}_i$ *is wrong*. Actually, these claims are incorrect; with the aid of the chain rule we could calculate the partial derivatives appearing in (4.13), making use of the expression for H in terms of any set of variables. It is *convenient* to express H in terms of (q_i, p_i, t) because, in that way, the partial derivatives involved in Hamilton's equations can be directly calculated and yield a system of ODEs that determine the time derivatives of the coordinates q_i and p_i in terms of (q_i, p_i, t). In what follows the parentheses and the subscripts in (4.13) will be suppressed unless some confusion might arise.

Equations (4.13) were obtained assuming that Equations (4.5) can be inverted to find the $\dot{q}_i$ in terms of (q_j, p_j, t). According to the inverse function theorem, this is locally possible if

$$\det\left(\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j}\right) \neq 0,$$

that is, if L is a regular Lagrangian (see Section 1.2). Thus, we conclude that if the Lagrange equations obtained from a Lagrangian L amount to an authentic system of n second-order ODEs, $\ddot{q}_i = f_i(q_j, \dot{q}_j, t)$, then this system is equivalent to the Hamilton equations (4.13).

When the set (q_i, p_i, t) is not functionally independent, the expression of H as a function of q_i, p_i, and t need not be unique (see Example 4.2, below), and Equations (4.13) make no sense.

Example 4.2. The function

$$L = \tfrac{1}{2}a\dot{q}_1{}^2 + bq_1\dot{q}_2 - q_1{}^2 q_2,$$

where a and b are constants, is a singular Lagrangian. The definition (4.5) gives the relations

$$p_1 = a\dot{q}_1, \qquad p_2 = bq_1$$

which show that, in effect, it is impossible to express the $\dot{q}_i$ as functions of (q_j, p_j, t) or, equivalently, that the set (q_i, p_i, t) is not functionally independent (e.g., p_2 can be expressed as a function of q_1).

Nevertheless, we can apply the definition of the Hamiltonian, which gives

$$H = \tfrac{1}{2}a\dot{q}_1{}^2 + q_1{}^2 q_2.$$

As shown above, this function must be expressible in terms of (q_j, p_j, t). In fact, there are an infinite number of expressions of H in terms of (q_j, p_j, t), some of them are

$$H = \frac{1}{2a}p_1{}^2 + q_1{}^2 q_2 = \frac{1}{2a}p_1{}^2 + \frac{p_2{}^2 q_2}{b^2} = \frac{1}{2a}p_1{}^2 + \frac{q_1 q_2 p_2}{b},$$

and one can readily verify that in each case Equation (4.11) duly holds.

Another example of a singular Lagrangian is

$$L = \sqrt{g_{ij}\frac{\mathrm{d}q_i}{\mathrm{d}t}\frac{\mathrm{d}q_j}{\mathrm{d}t}},$$

which leads to the equations for the geodesic curves [see (2.64)]. In fact, a straightforward computation gives

$$\frac{\partial^2 L}{\partial\dot{q}_i\partial\dot{q}_j} = \frac{(g_{lm}\dot{q}_l\dot{q}_m)g_{ij} - g_{il}\dot{q}_l g_{jm}\dot{q}_m}{(g_{rs}\dot{q}_r\dot{q}_s)^{3/2}}$$

and using this expression one verifies that

$$\frac{\partial^2 L}{\partial\dot{q}_i\partial\dot{q}_j}\dot{q}_j = 0,$$

identically. This, in turn, implies that $\det\left(\partial^2 L/\partial\dot{q}_i\partial\dot{q}_j\right) = 0$. Furthermore, the definition (4.10) gives $H = 0$. (This fact does not contradict Equations (4.13), because these equations are not applicable in this case.)

Fortunately, the Lagrangians found in almost all examples considered in classical mechanics are regular. In what follows we shall consider only systems defined by regular Lagrangians.

Example 4.3 (Hamiltonian for a charged particle in an electromagnetic field). The standard Lagrangian for a charged particle in an electromagnetic field, in Cartesian coordinates, is

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \frac{e}{c}(A_1\dot{x} + A_2\dot{y} + A_3\dot{z}) - e\varphi,$$

where $A_1, A_2, , A_3$ are the Cartesian components of the vector potential, $\mathbf{A}$ [see (1.69)]. According to the definition (4.5), the momenta conjugate to x, y, z are

$$p_1 = m\dot{x} + \frac{e}{c}A_1, \qquad p_2 = m\dot{y} + \frac{e}{c}A_2, \qquad p_3 = m\dot{z} + \frac{e}{c}A_3,$$

respectively, and these equations can be inverted to express $\dot{x}, \dot{y}, \dot{z}$ in terms of the canonical momenta,

$$\dot{x} = \frac{1}{m}(p_1 - \frac{e}{c}A_1), \qquad \dot{y} = \frac{1}{m}(p_2 - \frac{e}{c}A_2), \qquad \dot{z} = \frac{1}{m}(p_3 - \frac{e}{c}A_3).$$

Then, the definition of the Hamiltonian gives

$$\begin{aligned} H &= p_1\dot{x} + p_2\dot{y} + p_3\dot{z} - \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - \frac{e}{c}(A_1\dot{x} + A_2\dot{y} + A_3\dot{z}) + e\varphi \\ &= (p_1 - \frac{e}{c}A_1)\dot{x} + (p_2 - \frac{e}{c}A_2)\dot{y} + (p_3 - \frac{e}{c}A_3)\dot{z} - \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + e\varphi \\ &= \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + e\varphi \\ &= \frac{1}{2m}\left[\left(p_1 - \frac{e}{c}A_1\right)^2 + \left(p_2 - \frac{e}{c}A_2\right)^2 + \left(p_3 - \frac{e}{c}A_3\right)^2\right] + e\varphi. \end{aligned} \tag{4.14}$$

Exercise 4.4. Find the Hamiltonian corresponding to the Lagrangian

$$L = m(\dot{q}\ln\dot{q} - \dot{q}) - 2m\gamma q,$$

where m and γ are constants [see (2.53)]. Write down the Hamilton equations and solve them. (This is a rather unusual example where the Hamilton equations can be solved without raising the order of the equations.)

In many examples, the Lagrangian is the sum of a homogeneous function of the $\dot{q}_i$ of degree two and a function of the q_i only [see, e.g., Equations (1.54), (1.56), (2.2), (2.9), and (2.51)]. In order to simplify the computation of the corresponding Hamiltonian, it is convenient to establish a formula applicable to such cases.

Proposition 4.5. *For a Lagrangian of the form*

$$L = \tfrac{1}{2} a_{ij} \dot{q}_i \dot{q}_j - U(q_i, t),$$

where the a_{ij} are functions of (q_k, t) only, with $a_{ij} = a_{ji}$ and $\det(a_{ij}) \neq 0$*, the Hamiltonian is given by*

$$H = \tfrac{1}{2} b_{ij} p_i p_j + U(q_i, t),$$

where (b_{ij}) is the inverse of the matrix (a_{ij}). (The matrix (b_{ij}), being the inverse of a symmetric matrix, is also symmetric.)

The function U need not be the potential [see, e.g., Equation (1.54)].

Proof. From the definition of the generalized momenta (4.5) and the symmetry of the matrix (a_{ij}) we have

$$p_k = \frac{\partial}{\partial \dot{q}_k} \left(\frac{1}{2} a_{ij} \dot{q}_i \dot{q}_j \right) = \frac{1}{2} a_{ij} (\delta_{ik} \dot{q}_j + \delta_{jk} \dot{q}_i) = a_{kj} \dot{q}_j,$$

hence,

$$\dot{q}_i = b_{ij} p_j.$$

Then, making use of the definition of the Hamiltonian (4.10), we obtain

$$\begin{aligned} H &= p_i \dot{q}_i - \left[\tfrac{1}{2} a_{ij} \dot{q}_i \dot{q}_j - U(q_i, t) \right] \\ &= p_i \dot{q}_i - \left[\tfrac{1}{2} p_i \dot{q}_i - U(q_i, t) \right] \\ &= \tfrac{1}{2} p_i \dot{q}_i + U(q_i, t) \\ &= \tfrac{1}{2} b_{ij} p_i p_j + U(q_i, t). \end{aligned}$$

□

Example 4.6. In Example 1.11 we obtained the Lagrangian

$$L = \tfrac{1}{2} \left[(m_1 + m_2) \dot{x}^2 + 2 m_2 \cot \theta_0 \dot{x} \dot{y} + m_2 \csc^2 \theta_0 \dot{y}^2 \right] - m_2 g y$$

[see Equation (1.56)], which has the form considered in Proposition 4.5. In order to identify the matrix (a_{ij}) we note that the term $2m_2 \cot\theta_0 \dot{x}\dot{y}$, inside the brackets, is equal to $m_2 \cot\theta_0 \dot{x}\dot{y} + m_2 \cot\theta_0 \dot{y}\dot{x}$ and, therefore, with $q_1 = x$ and $q_2 = y$,

$$(a_{ij}) = \begin{pmatrix} m_1 + m_2 & m_2 \cot\theta_0 \\ m_2 \cot\theta_0 & m_2 \csc^2\theta_0 \end{pmatrix}.$$

Then, we find that

$$(b_{ij}) = \frac{1}{m_2(m_1 \csc^2\theta_0 + m_2)} \begin{pmatrix} m_2 \csc^2\theta_0 & -m_2 \cot\theta_0 \\ -m_2 \cot\theta_0 & m_1 + m_2 \end{pmatrix}$$

and, according to Proposition 4.5, the Hamiltonian is given by

$$H = \frac{1}{2m_2(m_1 \csc^2\theta_0 + m_2)} \left[m_2 \csc^2\theta_0 \, {p_x}^2 - 2m_2 \cot\theta_0 \, p_x p_y + (m_1 + m_2) {p_y}^2 \right] + m_2 g y.$$

Substituting H into the Hamilton equations (4.13) we obtain a system of four first-order ODEs equivalent (by construction) to the equations of motion given by the Lagrangian L.

Two additional examples of the application of Proposition 4.5 are given by the Lagrangians (1.57) and (2.29), corresponding to the one-dimensional harmonic oscillator and the two-dimensional isotropic harmonic oscillator, respectively. In the first case we obtain

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 q^2, \tag{4.15}$$

and in the second case

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + \frac{1}{2} m\omega^2 (x^2 + y^2). \tag{4.16}$$

The fact that the Hamilton equations are of first order usually does not signify a simplification of the solution of the equations of motion, because several canonical coordinates may be mixed in each equation. For instance, in the case of the Lagrangian (1.54),

$$L = \tfrac{1}{2} m(\dot{r}^2 + \omega^2 r^2 \sin^2\theta_0) - mgr\cos\theta_0,$$

making use of Proposition 4.5 we find that the Hamiltonian is

$$H = \frac{{p_r}^2}{2m} - \frac{m}{2}\omega^2 r^2 \sin^2\theta_0 + mgr\cos\theta_0,$$

where we have denoted by p_r the momentum conjugate to r. From the Hamilton equations we obtain the system of coupled equations

$$\dot{r} = \frac{p_r}{m}, \qquad \dot{p}_r = m\omega^2 r \sin^2\theta_0 - mg\cos\theta_0,$$

and in order to solve it we combine these equations to obtain a second-order ODE containing only one of the unknowns,

$$\ddot{p}_r = \omega^2 p_r \sin^2\theta_0 \qquad \text{or} \qquad \ddot{r} = \omega^2 r \sin^2\theta_0 - g\cos\theta_0.$$

The last of these equations is precisely the one that we obtained directly from the Lagrange formalism.

Taking into account the fact that in order to write down the Hamilton equations we have to find the Lagrangian as a first step, at this point the use of the Hamilton equations may not seem advantageous; however, in the rest of this book we shall be able to appreciate the many benefits of the Hamiltonian formalism.

As we have seen in Section 1.2, if we have obtained the expression for the Lagrangian in some set of generalized coordinates, q_i, we can find its expression in another set of generalized coordinates, $q'_i = q'_i(q_j, t)$, by simply substituting the variables q_i and $\dot{q}_i$, appearing in the original Lagrangian, by their expressions in terms of q'_i and $\dot{q}'_i$. A similar result holds in the case of the Hamiltonian, provided that we limit ourselves to coordinate transformations that do not involve the time, $q'_i = q'_i(q_j)$. (More general coordinate transformations will be considered in Chapter 5.)

Indeed, making use of the chain rule and (1.73) we have

$$p'_i = \frac{\partial L}{\partial \dot{q}'_i} = \frac{\partial L}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial \dot{q}'_i} = p_j \frac{\partial q_j}{\partial q'_i} \tag{4.17}$$

hence

$$p'_i\dot{q}'_i - L = p_j \frac{\partial q_j}{\partial q'_i}\dot{q}'_i - L = p_j\dot{q}_j - L$$

which means that, in order to find the Hamiltonian in terms of q'_i, p'_i, we only have to replace the coordinates q_i, p_i appearing in H by their expressions in terms of q'_i, p'_i, with the new and the old canonical momenta related by (4.17) or, equivalently, by

$$p_i = \frac{\partial q'_j}{\partial q_i} p'_j. \tag{4.18}$$

Ignorable Coordinates

In Chapter 1, we defined a (nontrivial) constant of motion as a function of $(q_i, \dot{q}_i, t)$ whose total derivative with respect to time is equal to zero as a consequence of the equations of motion. Since we have replaced the variables $\dot{q}_i$ by the generalized momenta, in the Hamiltonian formalism, a constant of motion will be a function of (q_i, p_i, t) whose total derivative with respect to time is equal to zero as a consequence of the equations of motion. (A trivial constant of motion is a function that does not depend on any of the variables (q_i, p_i, t) and, therefore, its total derivative with respect to the time is trivially equal to zero, without having to use the equations of motion.)

As in the case of the Lagrangian formalism, the absence of a coordinate q_i, or of t, in $H(q_i, p_i, t)$ allows us to identify immediately a constant of motion. In effect, if the generalized coordinate, q_k, does not appear in $H(q_i, p_i, t)$, then from (4.13) we have

$$\dot{p}_k = -\frac{\partial H}{\partial q_k} = 0,$$

which means that the momentum conjugate to q_k is a constant of motion. According to (4.12), q_k does not appear in $H(q_i, p_i, t)$ if and only if it does not appear in $L(q_i, \dot{q}_i, t)$. That is, q_k is an ignorable coordinate for H if and only if it is ignorable (or cyclic) for L, and in both cases one finds that the momentum conjugate to the ignorable coordinate is conserved. (The only difference is that the conserved momentum, p_k, can be substituted by a constant in the Hamiltonian and, in that manner, the Hamiltonian depends on the remaining $n - 1$ coordinates and their conjugate momenta only.)

Similarly, when the function $H(q_i, p_i, t)$ does not contain t, making use of the chain rule and the Hamilton equations one finds that the total time derivative of $H(q_i, p_i, t)$ is equal to zero

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q_i} = 0, \tag{4.19}$$

which means that H is a constant of motion. Equation (4.12) shows that $H(q_i, p_i, t)$ does not depend on t if and only if $L(q_i, \dot{q}_i, t)$ does not depend on t, and we know that when $L(q_i, \dot{q}_i, t)$ does not depend on t, the Jacobi integral (1.93) is a constant of motion, but, by comparing Equations (1.93) and (4.10) we see that these functions are equivalent to each other. Thus, also in this case, through the Lagrangian or the Hamiltonian we obtain the same constant of motion.

In some cases the following result is useful to discover constants of motion not related to ignorable coordinates.

Proposition 4.7. *If the Hamiltonian is a sum of the form*

$$H(q_i, p_i, t) = h_1(q_1, \ldots, q_k, p_1, \ldots, p_k) + h_2(q_{k+1}, \ldots, q_n, p_{k+1}, \ldots, p_n, t),$$

for some k, with $1 \leqslant k \leqslant n$, then h_1 is a constant of motion.

Note that we are only assuming that h_1 does not depend on t.

Proof. The chain rule and the Hamilton equations give

$$\frac{\mathrm{d}h_1}{\mathrm{d}t} = \sum_{i=1}^{k} \left(\frac{\partial h_1}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial h_1}{\partial p_i}\frac{\partial H}{\partial q_i} \right) = \sum_{i=1}^{k} \left(\frac{\partial h_1}{\partial q_i}\frac{\partial h_1}{\partial p_i} - \frac{\partial h_1}{\partial p_i}\frac{\partial h_1}{\partial q_i} \right) = 0.$$

□

A simple example is given by

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy,$$

which is the sum of

$$h_1 \equiv \frac{{p_x}^2}{2m} \qquad \text{and} \qquad h_2 \equiv \frac{{p_y}^2}{2m} + mgy$$

and, according to Proposition 4.7, h_1 and h_2 are constants of motion. The conservation of h_1 also follows from the fact that H does not depend on x and, therefore, p_x is conserved, and, since H is conserved (because it does not depend on t), h_2, being equal to $H - h_1$, must be also conserved.

However, if we perform the coordinate transformation

$$x = u + v, \qquad y = u - v,$$

according to (4.18) we have $p_x = \frac{1}{2}(p_u + p_v)$, $p_y = \frac{1}{2}(p_u - p_v)$ and, in the new coordinates, H has the expression

$$H = \frac{{p_u}^2 + {p_v}^2}{4m} + mg(u - v),$$

which is the sum of

$$h_1 \equiv \frac{{p_u}^2}{4m} + mgu \qquad \text{and} \qquad h_2 \equiv \frac{{p_v}^2}{4m} - mgv.$$

Now, Proposition 4.7 implies that H, h_1, and h_2 are constants of motion, but only two of them are functionally independent. Going back to the original coordinates, making use of the relations $p_u = p_x + p_y$, $p_v = p_x - p_y$, we have

$$h_1 = \frac{(p_x + p_y)^2}{4m} + mg\frac{x + y}{2}, \qquad h_2 = \frac{(p_x - p_y)^2}{4m} - mg\frac{x - y}{2}.$$

In particular, the difference

$$h_1 - h_2 = \frac{p_x p_y}{m} + mgx \tag{4.20}$$

is a constant of motion, which does not follow from the existence of ignorable coordinates.

Exercise 4.8. Show that if the Hamiltonian has the form $H(q_i, p_i, t) = f(t)\Phi(q_i, p_i)$, then Φ is a constant of motion. (Applications of this result are given, e.g., in Examples 5.17 and 5.18.)

4.2 The Poisson Bracket

In the same way as a set of generalized coordinates, q_i, serves to parameterize the configurations of a mechanical system (Section 1.1), a set of canonical coordinates, q_i, p_i parameterizes the *states* of the system (since the values of q_i and p_i at some initial time uniquely define the solution of the Hamilton equations) and, in the same way as the generalized coordinates q_i are coordinates of a certain space, called the configuration space, the canonical coordinates q_i, p_i are coordinates of the so-called *phase space*, and the $2n + 1$ variables q_i, p_i, and t are coordinates of the *extended phase space*. (The term "extended phase space" is also employed with other meanings in the literature.)

One of the advantages of the Hamilton equations over the Lagrange equations is that the former give directly the time derivatives of the variables q_i, p_i in terms of q_i, p_i, and t, which, among other things, facilitates the identification of the constants of motion. (The Lagrange equations contain the second derivatives of the q_i with respect to the time through the combinations $(\partial^2 L/\partial \dot{q}_i \partial \dot{q}_j)\ddot{q}_j$, see Equations (1.51).)

If $f(q_i, p_i, t)$ is a real-valued function defined on the extended phase space, according to the Hamilton equations, its total derivative with respect to the time is given by

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial p_i}\dot{p}_i = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i}.$$

Introducing the Poisson bracket of f and g, defined by

$$\{f, g\} \equiv \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}, \tag{4.21}$$

we can write

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \{f, H\}. \tag{4.22}$$

Thus, $f(q_i, p_i, t)$ is a constant of motion if

$$\frac{\partial f}{\partial t} + \{f, H\} = 0. \tag{4.23}$$

(The Poisson bracket is also denoted by $[f, g]$ and the notation (f, g) is sometimes employed. Moreover, the definition given by some authors amounts to the negative of the one used here.)

The definition (4.21) implies that this operation is antisymmetric,

$$\{f, g\} = -\{g, f\},$$

bilinear,

$$\{f, ag + bh\} = a\{f, g\} + b\{f, h\}, \qquad \{af + bg, h\} = a\{f, h\} + b\{g, h\},$$

where a, b are real numbers, and a derivation on each argument, which means that,

$$\{f, gh\} = g\{f, h\} + \{f, g\}h, \qquad \{fg, h\} = f\{g, h\} + \{f, h\}g.$$

Another important property of the Poisson bracket is the Jacobi identity.

Proposition 4.9 (The Jacobi identity). *For any three functions $f(q_i, p_i, t)$, $g(q_i, p_i, t)$, and $h(q_i, p_i, t)$,*

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0. \tag{4.24}$$

Proof. We shall give a direct proof of this proposition, which will give us the opportunity of introducing some notation that will be useful later. First, the canonical coordinates, q_i, p_i, will be denoted by a single symbol, x_α ($\alpha = 1, 2, \ldots, 2n$), with

$$(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{2n}) \equiv (q_1, \ldots, q_n, p_1, \ldots, p_n) \tag{4.25}$$

and introducing the constant, antisymmetric, $2n \times 2n$ block matrix

$$(\epsilon_{\alpha\beta}) \equiv \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{4.26}$$

($\alpha, \beta = 1, 2, \ldots, 2n$), where I is the $n \times n$ unit matrix (note that only one entry on each row, or column, is different from zero), that is

$$\varepsilon_{ij} = 0, \qquad \varepsilon_{i,n+j} = \delta_{ij}, \qquad \varepsilon_{n+i,j} = -\delta_{ij}, \qquad \varepsilon_{n+i,n+j} = 0 \tag{4.27}$$

$(i, j = 1, 2, \ldots, n)$, the Poisson bracket can be written as

$$\{f, g\} = \varepsilon_{\alpha\beta} \frac{\partial f}{\partial x_\alpha} \frac{\partial g}{\partial x_\beta} \tag{4.28}$$

(with sum over repeated indices).

Making use of (4.28) and the Leibniz rule we have

$$\begin{aligned} \{\{f, g\}, h\} &= \varepsilon_{\alpha\beta} \frac{\partial}{\partial x_\alpha} \left(\varepsilon_{\gamma\delta} \frac{\partial f}{\partial x_\gamma} \frac{\partial g}{\partial x_\delta} \right) \frac{\partial h}{\partial x_\beta} \\ &= \varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta} \frac{\partial f}{\partial x_\gamma} \frac{\partial^2 g}{\partial x_\alpha \partial x_\delta} \frac{\partial h}{\partial x_\beta} + \varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta} \frac{\partial^2 f}{\partial x_\alpha \partial x_\gamma} \frac{\partial g}{\partial x_\delta} \frac{\partial h}{\partial x_\beta}, \end{aligned}$$

hence,

$$\begin{aligned} &\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} \\ &\quad = \varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta} \frac{\partial f}{\partial x_\gamma} \frac{\partial^2 g}{\partial x_\alpha \partial x_\delta} \frac{\partial h}{\partial x_\beta} + \varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta} \frac{\partial^2 f}{\partial x_\alpha \partial x_\gamma} \frac{\partial g}{\partial x_\delta} \frac{\partial h}{\partial x_\beta} \\ &\qquad + \varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta} \frac{\partial g}{\partial x_\gamma} \frac{\partial^2 h}{\partial x_\alpha \partial x_\delta} \frac{\partial f}{\partial x_\beta} + \varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta} \frac{\partial^2 g}{\partial x_\alpha \partial x_\gamma} \frac{\partial h}{\partial x_\delta} \frac{\partial f}{\partial x_\beta} \\ &\qquad + \varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta} \frac{\partial h}{\partial x_\gamma} \frac{\partial^2 f}{\partial x_\alpha \partial x_\delta} \frac{\partial g}{\partial x_\beta} + \varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta} \frac{\partial^2 h}{\partial x_\alpha \partial x_\gamma} \frac{\partial f}{\partial x_\delta} \frac{\partial g}{\partial x_\beta}. \end{aligned}$$

Replacing the indices appearing in the first term on the right-hand side of the last equation according to $\alpha \mapsto \gamma, \beta \mapsto \delta, \gamma \mapsto \beta, \delta \mapsto \alpha$, we have the equivalent expression

$$\begin{aligned} \varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta} \frac{\partial f}{\partial x_\gamma} \frac{\partial^2 g}{\partial x_\alpha \partial x_\delta} \frac{\partial h}{\partial x_\beta} &= \varepsilon_{\gamma\delta}\varepsilon_{\beta\alpha} \frac{\partial f}{\partial x_\beta} \frac{\partial^2 g}{\partial x_\gamma \partial x_\alpha} \frac{\partial h}{\partial x_\delta} \\ &= -\varepsilon_{\gamma\delta}\varepsilon_{\alpha\beta} \frac{\partial f}{\partial x_\beta} \frac{\partial^2 g}{\partial x_\alpha \partial x_\gamma} \frac{\partial h}{\partial x_\delta}, \end{aligned}$$

where, in the last step, we made use of the antisymmetry of $\varepsilon_{\alpha\beta}$ (i.e., $\varepsilon_{\alpha\beta} = -\varepsilon_{\beta\alpha}$) and the commutativity of the second partial derivatives. Therefore, the terms with second partial derivatives of g cancel. In a similar manner we conclude that the terms containing second partial derivatives of f and h cancel, thus proving the Proposition. □

The Poisson brackets can be computed using directly the definition, but very often the computation can be abbreviated making use of the properties presented above and the formulas

$$\{q_i, f\} = \frac{\partial f}{\partial p_i}, \qquad \{p_i, f\} = -\frac{\partial f}{\partial q_i}, \tag{4.29}$$

which follow from the definition (4.21) and are applicable to any function f. In particular, we have

$$\{q_i, q_j\} = 0, \qquad \{q_i, p_j\} = \delta_{ij}, \qquad \{p_i, p_j\} = 0. \tag{4.30}$$

Note also that the Poisson bracket of two functions of the q_i only is equal to zero.

Exercise 4.10. When a charged particle interacts with an electromagnetic field, it is convenient to distinguish between the components of its *kinematic momentum*, $m\dot{\mathbf{r}}$, and its canonical momenta p_i. In terms of the Cartesian coordinates of the particle, the standard Lagrangian is given by

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \frac{e}{c}(A_1\dot{x} + A_2\dot{y} + A_3\dot{z}) - e\varphi,$$

where e is the electric charge of the particle, the A_i are the Cartesian components of the vector potential, $\mathbf{A}$, and φ is the scalar potential [see Equation (1.69)]. Hence, the Cartesian components of the kinematic momentum, $\pi_i \equiv m\dot{q}_i$, are related to the canonical momenta, p_i, by

$$p_i = \pi_i + \frac{e}{c}A_i. \tag{4.31}$$

(Since the vector potential is defined up to a gauge transformation, the canonical momenta are gauge-dependent.) Show that the Poisson brackets of the Cartesian components of the kinematic momentum of a charged particle in an electromagnetic field are given by

$$\{\pi_1, \pi_2\} = \frac{e}{c}B_3, \qquad \{\pi_2, \pi_3\} = \frac{e}{c}B_1, \qquad \{\pi_3, \pi_1\} = \frac{e}{c}B_2, \tag{4.32}$$

where the functions B_i are the Cartesian components of the magnetic field (i.e., $\mathbf{B} = \nabla \times \mathbf{A}$).

Example 4.11 (Poisson brackets between the Cartesian components of the angular momentum). A standard example related to the computation of Poisson brackets corresponds to the Poisson brackets between the Cartesian components of the usual angular momentum of a particle, $\mathbf{L} = \mathbf{r} \times m\dot{\mathbf{r}}$. If the Lagrangian of the particle has the usual form

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V(x, y, z),$$

then the conjugate momenta to the coordinates x, y, z are

$$p_x = m\dot{x}, \qquad p_y = m\dot{y}, \qquad p_z = m\dot{z},$$

respectively, and the Cartesian components of the angular momentum in terms of these canonical variables are given by

$$L_1 = yp_z - zp_y, \qquad L_2 = zp_x - xp_z, \qquad L_3 = xp_y - yp_x. \tag{4.33}$$

(Note that it is necessary to specify the form of the Lagrangian in order to determine the expression of the angular momentum in terms of the canonical momenta. Cf. Equation (4.31).) Then, making use of the properties of the Poisson bracket we have, for instance,

$$\begin{aligned}\{L_1, L_2\} &= \{yp_z - zp_y, zp_x - xp_z\}\\ &= \{yp_z, zp_x\} - \{yp_z, xp_z\} - \{zp_y, zp_x\} + \{zp_y, xp_z\}\\ &= y\{p_z, z\}p_x + p_y\{z, p_z\}x\\ &= -yp_x + xp_y\\ &= L_3\end{aligned}$$

and, by cyclic permutation of the indices we obtain, $\{L_2, L_3\} = L_1$ and $\{L_3, L_1\} = L_2$. With the aid of the Levi-Civita symbol (3.8), these relations can be expressed in the abbreviated form

$$\{L_i, L_j\} = \varepsilon_{ijk} L_k. \tag{4.34}$$

In a similar manner, with the aid of (4.29), one finds that

$$\begin{aligned}\{L_1, x^2 + y^2 + z^2\} &= 2x\{L_1, x\} + 2y\{L_1, y\} + 2z\{L_1, z\}\\ &= -2x\frac{\partial L_1}{\partial p_x} - 2y\frac{\partial L_1}{\partial p_y} - 2z\frac{\partial L_1}{\partial p_z}\\ &= -2y(-z) - 2zy\\ &= 0,\end{aligned}$$

and

$$\begin{aligned}\{L_1, {p_x}^2 + {p_y}^2 + {p_z}^2\} &= 2p_x\{L_1, p_x\} + 2p_y\{L_1, p_y\} + 2p_z\{L_1, p_z\}\\ &= 2p_x\frac{\partial L_1}{\partial x} + 2p_y\frac{\partial L_1}{\partial y} + 2p_z\frac{\partial L_1}{\partial z}\\ &= 2p_y p_z + 2p_z(-p_y).\\ &= 0,\end{aligned}$$

It should be clear that similar results are obtained considering L_2 and L_3 in place of L_1. Then, using the chain rule, one concludes that $\{L_i, f(r)\} = 0$, where $f(r)$ is an arbitrary function of $r = \sqrt{x^2 + y^2 + z^2}$, and that $\{L_i, f(p)\} = 0$, for any function of $p = \sqrt{{p_x}^2 + {p_y}^2 + {p_z}^2}$ only.

In particular, in the case of a particle in a central field of force, the Hamiltonian can be taken as

$$H = \frac{1}{2m}({p_x}^2 + {p_y}^2 + {p_z}^2) + V(r),$$

and $\{L_i, H\} = 0$, which shows that each Cartesian component of the angular momentum is a constant of motion.

Poisson Brackets of the Components of the Angular Momentum of a Rigid Body

In the case of a rigid body, it is possible to compute the Poisson brackets between the Cartesian components of the angular momentum without specifying a system of generalized coordinates.

The components of the angular momentum of a rigid body with a fixed point, with respect to Cartesian axes fixed in the body, $L_{i'}$, are related to the components of the angular velocity by means of $L_{r'} = I_{r's'}\omega_{s'}$, where the $I_{r's'}$ are the components of the inertia tensor [see (3.25)]. The components of the angular velocity, in turn, are related to the generalized velocities by means of $\omega_{s'} = M_{s'r}\dot{q}_r$ [see (3.29)], where the $M_{s'r}$ are functions of the q_i only, defined by Equation (3.30). Assuming that the generalized forces on the rigid body are derivable from a potential $V(q_i, t)$, the canonical momenta are given by [see (3.16)]

$$p_i = \frac{\partial T}{\partial \dot{q}_i} = \frac{\partial}{\partial \dot{q}_i}\left(\frac{1}{2} I_{j'k'}\omega_{j'}\omega_{k'}\right) = I_{j'k'}\omega_{j'}\frac{\partial \omega_{k'}}{\partial \dot{q}_i} = L_{k'}M_{k'i}.$$

Hence,

$$L_{k'} = (M^{-1})_{ik'}p_i, \tag{4.35}$$

where the $(M^{-1})_{ik'}$ are the entries of the inverse of the matrix $(M_{k'i})$ and they are functions of the q_i only [if the coordinates q_i are the standard Euler angles, the entries $M_{k'i}$ can be read from (3.38)]. Thus,

$$\begin{aligned}
\{L_{i'}, L_{j'}\} &= \frac{\partial L_{i'}}{\partial q_k}\frac{\partial L_{j'}}{\partial p_k} - \frac{\partial L_{j'}}{\partial q_k}\frac{\partial L_{i'}}{\partial p_k} \\
&= \frac{\partial (M^{-1})_{ri'}}{\partial q_k} p_r (M^{-1})_{kj'} - \frac{\partial (M^{-1})_{rj'}}{\partial q_k} p_r (M^{-1})_{ki'} \\
&= \frac{\partial (M^{-1})_{ri'}}{\partial q_k} L_{s'}M_{s'r}(M^{-1})_{kj'} - \frac{\partial (M^{-1})_{rj'}}{\partial q_k} L_{s'}M_{s'r}(M^{-1})_{ki'}.
\end{aligned}$$

Noting that

$$\frac{\partial (M^{-1})_{ri'}}{\partial q_k} M_{s'r} = \frac{\partial [M_{s'r}(M^{-1})_{ri'}]}{\partial q_k} - \frac{\partial M_{s'r}}{\partial q_k}(M^{-1})_{ri'} = -\frac{\partial M_{s'r}}{\partial q_k}(M^{-1})_{ri'},$$

with the aid of (3.32), we obtain

$$\begin{aligned}
\{L_{i'}, L_{j'}\} &= -(M^{-1})_{ri'} \frac{\partial M_{s'r}}{\partial q_k} L_{s'} (M^{-1})_{kj'} + (M^{-1})_{rj'} \frac{\partial M_{s'r}}{\partial q_k} L_{s'} (M^{-1})_{ki'} \\
&= -(M^{-1})_{ki'} \frac{\partial M_{s'k}}{\partial q_r} L_{s'} (M^{-1})_{rj'} + (M^{-1})_{rj'} \frac{\partial M_{s'r}}{\partial q_k} L_{s'} (M^{-1})_{ki'} \\
&= (M^{-1})_{ki'} (M^{-1})_{rj'} L_{s'} \left(\frac{\partial M_{s'r}}{\partial q_k} - \frac{\partial M_{s'k}}{\partial q_r} \right) \\
&= (M^{-1})_{ki'} (M^{-1})_{rj'} L_{s'} \varepsilon_{l'm's'} M_{l'r} M_{m'k} \\
&= \delta_{m'i'} \delta_{l'j'} \varepsilon_{l'm's'} L_{s'} \\
&= -\varepsilon_{i'j's'} L_{s'},
\end{aligned} \tag{4.36}$$

which are similar to Equations (4.34). The additional minus sign appearing in (4.36) comes from the fact that in this last equation we are dealing with components of the angular momentum with respect to axes fixed in the body. With the aid of (4.36) we can prove that

$$\{L_i, L_j\} = \varepsilon_{ijk} L_k \qquad \text{and} \qquad \{L_i, L_{j'}\} = 0, \tag{4.37}$$

where the L_i are the components of the angular momentum with respect to the inertial frame.

Exercise 4.12. Show that Equations (4.37) indeed hold.

Proposition 4.13 (Poisson's Theorem). *If f and g are two differentiable functions defined on the extended phase space, then*

$$\frac{\mathrm{d}\{f, g\}}{\mathrm{d}t} = \{f, \mathrm{d}g/\mathrm{d}t\} + \{\mathrm{d}f/\mathrm{d}t, g\}. \tag{4.38}$$

Hence, if f and g are constants of motion (that may depend explicitly on the time), then $\{f, g\}$ is also a constant of motion.

Proof. According to (4.22), the definition of the Poisson bracket, the Jacobi identity, the antisymmetry, and the bilinearity of the Poisson bracket, we have

$$\begin{aligned}
\frac{\mathrm{d}\{f,g\}}{\mathrm{d}t} &= \frac{\partial\{f,g\}}{\partial t} + \{\{f,g\},H\} \\
&= \frac{\partial}{\partial t}\left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right) - \{\{g,H\},f\} - \{\{H,f\},g\} \\
&= \frac{\partial f}{\partial q_i}\frac{\partial^2 g}{\partial t\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial^2 g}{\partial t\partial q_i} + \frac{\partial^2 f}{\partial t\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial^2 f}{\partial t\partial p_i}\frac{\partial g}{\partial q_i} \\
&\quad + \{f,\{g,H\}\} + \{\{f,H\},g\} \\
&= \{f,\partial g/\partial t\} + \{\partial f/\partial t, g\} + \{f,\{g,H\}\} + \{\{f,H\},g\} \\
&= \{f, \mathrm{d}g/\mathrm{d}t\} + \{\mathrm{d}f/\mathrm{d}t, g\}.
\end{aligned}$$

□

Unfortunately, the Poisson bracket of two nontrivial constants of motion may be a trivial constant of motion, or may be a function of the initial constants of motion. For instance, in the case of a particle in a uniform gravitational field, with the Hamiltonian

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy,$$

p_x and $p_x p_y/m + mgx$ are constants of motion [see (4.20)], but their Poisson bracket is a trivial constant [see (4.29)].

A similar result, applicable in the case where H does not depend explicitly on the time, is that if f is a constant of motion, then $\partial f/\partial t$ is also a constant of motion, as can be readily seen by calculating the partial derivative with respect to the time of both sides of (4.23). Also in this case, the resulting constant of motion may be trivial or a function of the constants of motion already known.

Exercise 4.14. Consider a charged particle in a magnetic field whose Cartesian components, B_1, B_2, B_3, are functions of z only (this means that the magnetic field is invariant under translations parallel to the xy-plane). Then, the condition $\nabla\cdot\mathbf{B} = 0$ implies that B_3 is constant. Show that the functions

$$\mathscr{P}_1 = \pi_1 + \frac{e}{c}\int (B_2\mathrm{d}z - B_3\mathrm{d}y), \qquad \mathscr{P}_2 = \pi_2 + \frac{e}{c}\int (B_3\mathrm{d}x - B_1\mathrm{d}z),$$

are well defined (that is, the integrands are exact differentials) and are constants of motion. Find their Poisson bracket. Is it a nontrivial constant of motion?

4.2.1 Hamilton's Principle in the Phase Space

In the same manner as the Lagrange equations determine the stationary values of certain integral (see Section 2.4), the Hamilton equations determine the stationary values of the line integral

$$I(C) \equiv \int_C (p_i \mathrm{d}q_i - H \mathrm{d}t) = \int_{t_0}^{t_1} \left[p_i(t) \frac{\mathrm{d}q_i(t)}{\mathrm{d}t} - H(q_i(t), p_i(t), t) \right] \mathrm{d}t, \qquad (4.39)$$

in the space of curves in the extended phase space that share the same endpoints with C. In order to prove this assertion it is enough to consider one-parameter families of curves in the extended phase space of the form

$$C^{(s)}(t) = (q_i^{(s)}(t), p_i^{(s)}(t), t),$$

where s is a real parameter that takes values in some neighborhood of zero. That is, all the curves $C^{(s)}$ are parameterized by t, which takes values in some interval $[t_0, t_1]$.

Making use of the definitions

$$\eta_i(t) \equiv \left. \frac{\partial q_i^{(s)}(t)}{\partial s} \right|_{s=0}, \qquad \xi_i(t) \equiv \left. \frac{\partial p_i^{(s)}(t)}{\partial s} \right|_{s=0}$$

and the chain rule we obtain

$$\begin{aligned} \left. \frac{\mathrm{d}I(C^{(s)})}{\mathrm{d}s} \right|_{s=0} &= \frac{\mathrm{d}}{\mathrm{d}s} \int_{t_0}^{t_1} \left[p_i^{(s)}(t) \frac{\mathrm{d}q_i^{(s)}(t)}{\mathrm{d}t} - H\big(q_i^{(s)}(t), p_i^{(s)}(t), t\big) \right] \mathrm{d}t \bigg|_{s=0} \\ &= \int_{t_0}^{t_1} \left[\xi_i \frac{\mathrm{d}q_i^{(0)}}{\mathrm{d}t} + p_i^{(0)} \frac{\mathrm{d}\eta_i}{\mathrm{d}t} - \frac{\partial H}{\partial q_i} \eta_i - \frac{\partial H}{\partial p_i} \xi_i \right] \mathrm{d}t \\ &= p_i^{(0)} \eta_i \Big|_{t_0}^{t_1} + \int_{t_0}^{t_1} \left[\xi_i \left(\frac{\mathrm{d}q_i^{(0)}}{\mathrm{d}t} - \frac{\partial H}{\partial p_i} \right) - \eta_i \left(\frac{\mathrm{d}p_i^{(0)}}{\mathrm{d}t} + \frac{\partial H}{\partial q_i} \right) \right] \mathrm{d}t, \end{aligned}$$

where, in the last equality, we have integrated by parts. The integrand in the last integral is evaluated on the curve $C^{(0)}$.

If we now assume that all the curves $C^{(s)}$ have the same endpoints in the extended phase space, then the functions $\eta_i(t)$ and $\xi_i(t)$ are equal to zero at $t = t_0$ and $t = t_1$, and if the integral (4.39) has a stationary value on the curve $C^{(0)}$, compared with the curves in the extended phase space that have the same endpoints as $C^{(0)}$, from the last equation we have

$$\int_{t_0}^{t_1} \left[\xi_i \left(\frac{\mathrm{d}q_i^{(0)}}{\mathrm{d}t} - \frac{\partial H}{\partial p_i} \right) - \eta_i \left(\frac{\mathrm{d}p_i^{(0)}}{\mathrm{d}t} + \frac{\partial H}{\partial q_i} \right) \right] \mathrm{d}t = 0,$$

for all functions $\eta_i(t)$ and $\xi_i(t)$ whose value is zero at $t = t_0$ and $t = t_1$, which implies that $q_i^{(0)}$ and $p_i^{(0)}$ satisfy the Hamilton equations.

It may be noticed that it is not necessary to assume that the functions $\xi_i(t)$ be equal to zero at $t = t_0$ and $t = t_1$. This means that the Hamilton equations determine the stationary values of the integral (4.39) in the space of curves in the extended phase space that have the same coordinates q_i at $t = t_0$ and $t = t_1$ (without restriction on the coordinates p_i).

4.3 Equivalent Hamiltonians

In this section we shall show that, apart from the fact that the form of the Hamiltonian can be modified by means of a change of coordinates, a given time-independent Hamiltonian can be replaced by other functions, at the expense of replacing the time by another independent variable. We shall begin by applying this procedure to the Kepler problem in two dimensions, and then we will state the general result (Proposition 4.16, below). (A similar method, though more restricted, is employed in Pars [11, Sects. 17.3, 18.1, 18.3, and 26.7], and in Perelomov [13, Sect. 2.3]. Equivalent results are obtained by means of the Hamilton–Jacobi equation, see, e.g., Example 6.4.) As we shall see, in some cases, this method also allows us to find constants of motion (not related to ignorable coordinates), without having to solve the equations of motion, and to reduce the solution of the equations of motion to quadratures.

The *parabolic coordinates* in the Euclidean plane, u, v, can be defined by

$$x = \tfrac{1}{2}(u^2 - v^2), \qquad y = uv, \tag{4.40}$$

in terms of the Cartesian coordinates, x, y. Hence, the distance, r, from a point of the plane to the origin can be expressed as

$$r = \sqrt{x^2 + y^2} = \tfrac{1}{2}(u^2 + v^2). \tag{4.41}$$

With the aid of these expressions one finds that the standard Lagrangian for the Kepler problem in two dimensions,

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + \frac{k}{\sqrt{x^2 + y^2}},$$

takes the form

$$L = \frac{m}{2}(u^2 + v^2)(\dot{u}^2 + \dot{v}^2) + \frac{2k}{u^2 + v^2} \tag{4.42}$$

and, therefore (see Proposition 4.5),

$$H = \frac{1}{2m} \frac{{p_u}^2 + {p_v}^2}{u^2 + v^2} - \frac{2k}{u^2 + v^2}. \tag{4.43}$$

Even though, at first sight, this expression may not seem more convenient than the Hamiltonian in Cartesian coordinates,

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} - \frac{k}{\sqrt{x^2 + y^2}}, \tag{4.44}$$

as we shall see, with the introduction of an auxiliary parameter, (4.43) leads to simple equations of motion. To this end we start by pointing out that, since H does not depend explicitly on the time, H is a constant of motion and that if $k > 0$, H can take any real value E (by contrast, if $k < 0$, corresponding to a repulsive force, the Hamiltonian H takes positive values only). Noting that the equation

$$\frac{1}{2m} \frac{{p_u}^2 + {p_v}^2}{u^2 + v^2} - \frac{2k}{u^2 + v^2} = E$$

amounts to

$$\frac{{p_u}^2 + {p_v}^2}{2m} - E(u^2 + v^2) = 2k,$$

we introduce the auxiliary function

$$h_E \equiv \frac{{p_u}^2 + {p_v}^2}{2m} - E(u^2 + v^2), \tag{4.45}$$

so that the hypersurface of the phase space $H = E$ is equivalently given by $h_E = 2k$. From (4.43) and (4.45) we see that the functions H and h_E are related by

$$H - E = \frac{h_E - 2k}{u^2 + v^2} \tag{4.46}$$

and, as we shall prove now, h_E is, in a sense to be specified below, a Hamiltonian equivalent to H.

In fact, if w denotes any of the parabolic coordinates, u, v, or their conjugate momenta, p_u, p_v, making use of (4.46) and the Leibniz rule, we have

$$\left.\frac{\partial H}{\partial w}\right|_{H=E} = \left.\frac{\partial H}{\partial w}\right|_{h_E=2k} = \left.\frac{\partial}{\partial w}\left(\frac{h_E - 2k}{u^2 + v^2}\right)\right|_{h_E=2k} = \frac{1}{u^2 + v^2} \left.\frac{\partial h_E}{\partial w}\right|_{h_E=2k}.$$

Thus, introducing the auxiliary parameter τ by means of

$$\mathrm{d}\tau = \frac{\mathrm{d}t}{u^2 + v^2} \tag{4.47}$$

we find that, on the hypersurface $H = E$,

$$\frac{\mathrm{d}u}{\mathrm{d}\tau} = (u^2 + v^2)\frac{\mathrm{d}u}{\mathrm{d}t} = (u^2 + v^2)\frac{\partial H}{\partial p_u} = \frac{\partial h_E}{\partial p_u} \tag{4.48}$$

and, similarly,

$$\frac{\mathrm{d}p_u}{\mathrm{d}\tau} = -\frac{\partial h_E}{\partial u}, \qquad \frac{\mathrm{d}v}{\mathrm{d}\tau} = \frac{\partial h_E}{\partial p_v}, \qquad \frac{\mathrm{d}p_v}{\mathrm{d}\tau} = -\frac{\partial h_E}{\partial v}. \tag{4.49}$$

That is, the equations of motion can be written in the form of the Hamilton equations with the function h_E as the Hamiltonian and the parameter τ in place of the time.

Note that the right-hand side of (4.47) is not an exact differential and, therefore, τ is not a function defined on the extended phase space; when u and v are given as functions of t, or τ, then (4.47) allows us to find a relationship between t and τ (see the examples below). Equation (4.47) is analogous to the well-known expression for the arclength of a curve in the Euclidean plane $\mathrm{d}s = \sqrt{1 + (\mathrm{d}y/\mathrm{d}x)^2}\,\mathrm{d}x$. (The parameter τ is sometimes called *fictitious time* or *local time*.)

The function h_E, being a polynomial of second degree in the canonical coordinates [see (4.45)], seems more convenient than the Hamiltonian (4.43) and leads to a linear system of first-order ODEs. We postpone the solution of this system to the end of this chapter. Right now, we note that the function h_E, defined by Equation (4.45), is the sum of the two functions

$$h_1 \equiv \frac{{p_u}^2}{2m} - Eu^2 \qquad \text{and} \qquad h_2 \equiv \frac{{p_v}^2}{2m} - Ev^2.$$

h_1 depends on u and p_u, while h_2 is a function of v and p_v only. Hence, according to Proposition 4.7, on the hypersurface $H = E$, h_1, and h_2 are constants of motion (but they are not independent because $h_1 + h_2 = h_E = 2k$ on the hypersurface $H = E$). Thus,

$$\frac{{p_u}^2}{2m} - Eu^2 = k - D, \qquad \frac{{p_v}^2}{2m} - Ev^2 = k + D, \tag{4.50}$$

where D is a constant of motion. (Cf. Perelomov [13, Sect. 2.3].) Since $p_u = m\mathrm{d}u/\mathrm{d}\tau$ [see Equation (4.48)] and, similarly, $p_v = m\mathrm{d}v/\mathrm{d}\tau$, Equations (4.50) are equivalent to the independent first-order ODEs

$$\frac{m}{2}\left(\frac{\mathrm{d}u}{\mathrm{d}\tau}\right)^2 - Eu^2 = k - D, \qquad \frac{m}{2}\left(\frac{\mathrm{d}v}{\mathrm{d}\tau}\right)^2 - Ev^2 = k + D,$$

which can be readily solved. (The explicit form of the solution depends on whether E is positive, negative, or equal to zero.)

In order to eliminate E from the expression for the constant of motion D, we multiply Equations (4.50) by v^2 and u^2, respectively, and subtracting the results, we obtain

$$\frac{v^2 {p_u}^2 - u^2 {p_v}^2}{2m} = k(v^2 - u^2) - (v^2 + u^2)D,$$

which gives an expression for D in terms of the coordinates u, v and their conjugate momenta. In order to facilitate the identification of this constant, it is convenient to express it in terms of the Cartesian coordinates of the particle and their conjugate momenta. We will make use of the relations

$$p_u = up_x + vp_y, \qquad p_v = -vp_x + up_y, \tag{4.51}$$

which follow from (4.17), and (4.40). For instance,

$$p_u = p_u \frac{\partial x}{\partial u} + p_y \frac{\partial y}{\partial u} = p_x u + p_y v.$$

Then, with the aid of Equations (4.40), (4.41), and (4.51), we see that

$$\begin{aligned}
D &= \frac{u^2 {p_v}^2 - v^2 {p_u}^2}{2m(u^2 + v^2)} + k\frac{v^2 - u^2}{u^2 + v^2} \\
&= \frac{u^2(-vp_x + up_y)^2 - v^2(up_x + vp_y)^2}{2m(u^2 + v^2)} - k\frac{x}{r} \\
&= \frac{x{p_y}^2 - yp_x p_y}{m} - k\frac{x}{r} \\
&= \frac{L_3 p_y}{m} - k\frac{x}{r},
\end{aligned}$$

where L_3 is the z-component of the angular momentum of the particle about the origin [see (4.33)]. A straightforward computation shows that, except for a constant factor $1/m$, D is the x-component of the *Laplace–Runge–Lenz vector* (also known as *Runge–Lenz vector*), defined by

$$\mathbf{A} = \mathbf{p} \times \mathbf{L} - \frac{mk}{r}\mathbf{r}, \tag{4.52}$$

identifying a vector (a_1, a_2) in the plane with the vector $(a_1, a_2, 0)$, and $\mathbf{L} = (0, 0, L_3)$.

It should be remarked that in spite of the fact that initially we considered a hypersurface $H = E$, the expression for the constant of motion, D, derived above is applicable in all the phase space since all reference to the value of E was removed. In fact, a direct computation shows that D is conserved, regardless of the value of the energy (see Exercise 4.34, below).

Owing to the rotational symmetry of the potential, we may expect that also the y-component of the Laplace–Runge–Lenz vector be conserved (see Exercise 4.15).

Exercise 4.15. Find the Poisson bracket between the constants of motion L_3 and D which, according to Poisson's theorem (Proposition 4.13), must be also a constant of motion. Is it directly related to the Laplace–Runge–Lenz vector?

Now we shall establish the general results, considering an arbitrary Hamiltonian, H, that does not depend explicitly on the time, so that $H(q_i, p_i)$ is a constant of motion [see Equation (4.19)], which means that, for any initial condition, the curve $(q_i(t), p_i(t))$, representing the evolution of the system, is contained in one of the level surfaces $H(q_i, p_i) = \text{const.}$ (see Figure 4.1).

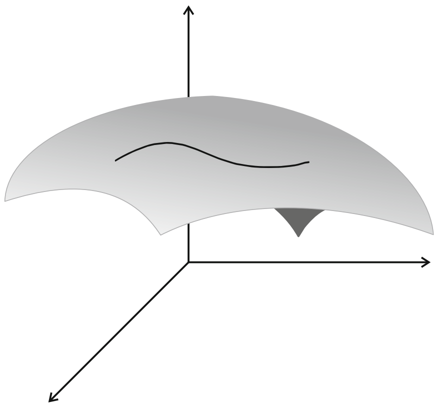


Fig. 4.1 The level surfaces of the Hamiltonian $H(q_i, p_i)$, which are defined by $H(q_i, p_i) = \text{const.}$, are hypersurfaces in the phase space (provided that $\mathrm{d}H$ is different from zero at the points of $H(q_i, p_i) = \text{const.}$). The conservation of H means that, for any initial condition, the curve $(q_i(t), p_i(t))$ lies on the hypersurface $H(q_i, p_i) = \text{const.}$ containing the initial condition

We are going to prove that, on each hypersurface $H = E$, where E is a possible value of H, the Hamiltonian H can be replaced by *any* real-valued function defined on the phase space, h_E, provided that the level surface $H = E$ coincides with some level surface $h_E = \varepsilon$, where ε is a constant. In a more precise sense, we have the following Proposition.

Proposition 4.16. *Let $H(q_i, p_i)$ be a given time-independent Hamiltonian and let E be a possible value of H, if there exist a nonvanishing real-valued function $f(q_i, p_i)$ and a one-to-one function $g : \mathbb{R} \to \mathbb{R}$ such that*

$$H - E = f(q_i, p_i)[g(h_E - \varepsilon) - g(0)], \tag{4.53}$$

so that the level surface $H = E$ coincides with the level surface $h_E = \varepsilon$, then, on the hypersurface $H = E$, the equations of motion are given by

$$\frac{\mathrm{d}q_i}{\mathrm{d}\tau} = \frac{\partial h_E}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}\tau} = -\frac{\partial h_E}{\partial q_i}, \tag{4.54}$$

with the parameter τ *defined by*

$$\mathrm{d}\tau \equiv f(q_i, p_i)|_{h_E=\varepsilon}\, g'(0)\, \mathrm{d}t. \tag{4.55}$$

Proof. Taking the partial derivatives of both sides of Equation (4.53) with respect to p_i, at the points of the hypersurface $h_E = \varepsilon$, we obtain

$$\left.\frac{\partial H}{\partial p_i}\right|_{h_E=\varepsilon} = f(q_i, p_i)|_{h_E=\varepsilon}\, g'(0) \left.\frac{\partial h_E}{\partial p_i}\right|_{h_E=\varepsilon},$$

hence, making use of the Hamilton equations and (4.55),

$$\left.\frac{\partial h_E}{\partial p_i}\right|_{h_E=\varepsilon} = \frac{1}{f(q_i, p_i)|_{h_E=\varepsilon}\, g'(0)} \frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\mathrm{d}q_i}{\mathrm{d}\tau}.$$

In a similar manner, taking the partial derivative of both sides of Equation (4.53) with respect to q_i, at the points of the hypersurface $h_E = \varepsilon$, we have

$$\left.\frac{\partial H}{\partial q_i}\right|_{h_E=\varepsilon} = f(q_i, p_i)|_{h_E=\varepsilon}\, g'(0) \left.\frac{\partial h_E}{\partial q_i}\right|_{h_E=\varepsilon},$$

hence, by virtue of the Hamilton equations and (4.55),

$$-\left.\frac{\partial h_E}{\partial q_i}\right|_{h_E=\varepsilon} = \frac{1}{f(q_i, p_i)|_{h_E=\varepsilon}\, g'(0)} \frac{\mathrm{d}p_i}{\mathrm{d}t} = \frac{\mathrm{d}p_i}{\mathrm{d}\tau}.$$

□

As one may guess, even though we can replace a given Hamiltonian, H, by other functions, h_E, as indicated, not every function h_E will lead to simpler equations of motion than those obtained from H.

Example 4.17 (Particle in a central potential). The standard Hamiltonian for a particle of mass m in a central potential $V(r)$, in spherical coordinates, is given by

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

(see Equation (2.2) and Proposition 4.5). Since this expression does not contain ϕ or t, p_ϕ and H are two constants of motion. The number of degrees of freedom is three and therefore, in order to find the general solution of the equations of motion, we require six (functionally independent) constants of motion.

The equation $H = E$ is equivalent to $h_E = 0$, where

$$h_E \equiv \frac{1}{2m}\left(r^2 {p_r}^2 + {p_\theta}^2 + \frac{{p_\phi}^2}{\sin^2\theta}\right) + r^2 V(r) - E r^2,$$

which is the sum of the two functions

$$h_1 \equiv \frac{r^2 {p_r}^2}{2m} + r^2 V(r) - Er^2 \qquad \text{and} \qquad h_2 \equiv \frac{1}{2m}\left({p_\theta}^2 + \frac{{p_\phi}^2}{\sin^2\theta}\right).$$

h_1 is a function of r and p_r only, and h_2 is a function of the angles θ, ϕ and their conjugate momenta. Hence, h_1 and h_2 are constants of motion (with $h_1 + h_2 = h_E = 0$ on the hypersurface $H = E$) (see Proposition 4.7).

Taking into account that

$$ {p_\theta}^2 + \frac{{p_\phi}^2}{\sin^2\theta} = \mathbf{L}^2, \tag{4.56}$$

where $\mathbf{L}$ is the angular momentum of the particle about the origin [see (2.7)], it follows that the square of the angular momentum is conserved. In fact, not only the magnitude of $\mathbf{L}$ is a constant of motion, the three Cartesian components of $\mathbf{L}$ are constants of motion (see Section 4.2). The function h_2 has the form of the standard Hamiltonian of a free particle in a sphere of radius 1 (a spherical pendulum without the gravitational field). Hence, the intersection of the position vector of the particle with the unit sphere is a circle centered at the origin, which is traversed with a constant speed according to the parameter τ.

On the other hand,

$$\frac{\mathrm{d}r}{\mathrm{d}\tau} = \frac{\partial h}{\partial p_r} = \frac{r^2 p_r}{m}.$$

Hence, using the fact that h_1 is conserved and its value is equal to $-\mathbf{L}^2/2m$, we have

$$\frac{r^2}{2m}\left(\frac{m}{r^2}\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 + r^2 V(r) - Er^2 = -\frac{\mathbf{L}^2}{2m},$$

hence,

$$\pm \mathrm{d}\tau = \frac{\mathrm{d}r}{r^2\sqrt{\dfrac{2E}{m} - \dfrac{\mathbf{L}^2}{m^2 r^2} - \dfrac{2V(r)}{m}}},$$

which is equivalent to the equation for the orbit (2.15).

Exercise 4.18 (Plane motion of a particle attracted by two fixed centers). The potential energy of a particle of mass m moving on the plane under the gravitational attraction produced by two fixed centers at the points $(c, 0)$ and $(-c, 0)$, has the form

$$V = -\frac{k_1}{r_1} - \frac{k_2}{r_2},$$

where k_1 and k_2 are constants, and $r_1 = \sqrt{(x-c)^2+y^2}$, $r_2 = \sqrt{(x+c)^2+y^2}$. In this case it is convenient to employ the *elliptic coordinates* (also called confocal coordinates), u, v, defined by

$$x = c\,\cosh u\,\cos v, \qquad y = c\,\sinh u\,\sin v,$$

because the distances from an arbitrary point of the plane to the points $(c, 0)$ and $(-c, 0)$ are given by $r_1 = c\,(\cosh u - \cos v)$ and $r_2 = c\,(\cosh u + \cos v)$, respectively (see Figure 4.2). Hence, the standard Lagrangian is given by

$$\begin{aligned} L &= \frac{mc^2}{2}(\cosh^2 u - \cos^2 v)(\dot{u}^2 + \dot{v}^2) + \frac{k_1}{c\,(\cosh u - \cos v)} + \frac{k_2}{c\,(\cosh u + \cos v)} \\ &= \frac{mc^2}{2}(\cosh^2 u - \cos^2 v)(\dot{u}^2 + \dot{v}^2) + \frac{(k_1+k_2)\cosh u + (k_1-k_2)\cos v}{c\,(\cosh^2 u - \cos^2 v)}, \end{aligned}$$

and the Hamiltonian is (see Proposition 4.5)

$$H = \frac{{p_u}^2 + {p_v}^2}{2mc^2(\cosh^2 u - \cos^2 v)} - \frac{(k_1+k_2)\cosh u + (k_1-k_2)\cos v}{c\,(\cosh^2 u - \cos^2 v)}. \tag{4.57}$$

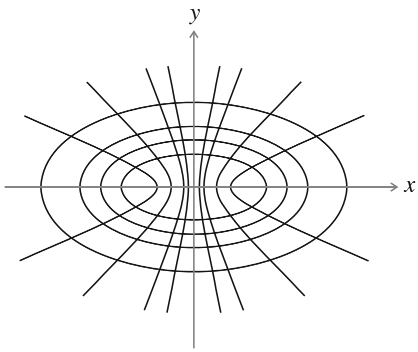


Fig. 4.2 The distances from an arbitrary point of the Euclidean plane with elliptic coordinates u, v to $(c, 0)$ and $(-c, 0)$ are $r_1 = c(\cosh u - \cos v)$ and $r_2 = c(\cosh u + \cos v)$, respectively, hence $r_1 + r_2 = 2c\,\cosh u$, and $r_2 - r_1 = 2c\,\cos v$, which show that the curves $u =$ const. are ellipses with foci $(\pm c, 0)$ and eccentricity $1/\cosh u$, while the curves $v =$ const. are branches of hyperbolas with foci $(\pm c, 0)$ and eccentricity $1/|\cos v|$

Find an equivalent Hamiltonian for this problem, h_E, such that the equations of motion for u and p_u in terms of τ are independent of the equations of motion for v and p_v, and vice versa, identify a constant of motion in addition to H, and reduce to quadratures the problem of finding u and v in terms of τ (cf. Example 4.17).

Exercise 4.19 (Two-dimensional isotropic harmonic oscillator in confocal coordinates). Show that the standard Hamiltonian for the two-dimensional isotropic harmonic oscillator in confocal coordinates, defined in Exercise 4.18, is given by

$$H = \frac{{p_u}^2 + {p_v}^2}{2mc^2(\cosh^2 u - \cos^2 v)} + \frac{m\omega^2 c^2}{2}(\cosh^2 u + \cos^2 v - 1). \tag{4.58}$$

Show that we can define an equivalent Hamiltonian, h_E, such that the equations of motion for u and p_u in terms of the parameter τ are independent of the equations of motion for v and p_v and vice versa. Reduce the solution of the equations of motion to quadratures.

Exercise 4.20. Show that if the Hamiltonian has the form

$$H = \frac{1}{2}\frac{P{p_x}^2 + Q{p_y}^2}{X + Y} + \frac{\xi + \eta}{X + Y}, \tag{4.59}$$

where P, X, ξ are functions of x only, and Q, Y, η are functions of y only, then by defining a suitable equivalent Hamiltonian, h_E, the equations of motion for x and p_x in terms of τ are independent of the equations of motion for y and p_y, and vice versa, and identify a constant of motion in addition to H. Note that the Hamiltonians (4.43), (4.57), and (4.58) are of the form (4.59).

Exercise 4.21 (Liouville's system). As an extension of the problem treated in Exercise 4.20, we consider the Hamiltonian

$$H = \frac{1}{2}\frac{P_1{p_1}^2 + P_2{p_2}^2 + \cdots + P_n{p_n}^2}{X_1 + X_2 + \cdots + X_n} + \frac{\xi_1 + \xi_2 + \cdots + \xi_n}{X_1 + X_2 + \cdots + X_n},$$

where P_i, X_i, and ξ_i are functions of q_i only (see, e.g., Pars [11, Sect. 18.1]). By defining a suitable equivalent Hamiltonian, h_E, show that, for each value of the index i, the equations of motion for q_i and p_i in terms of τ are independent of the remaining coordinates q_j and p_j, and identify $n - 1$ constants of motion in addition to H.

Exercise 4.22 (Charged particle in the field of a point electric dipole). The standard Hamiltonian of a charged particle in the field of a point electric dipole placed at the origin, in spherical coordinates (r, θ, ϕ), is

$$H = \frac{1}{2m}\left({p_r}^2 + \frac{{p_\theta}^2}{r^2} + \frac{{p_\phi}^2}{r^2 \sin^2\theta}\right) + \frac{k\cos\theta}{r^2},$$

where k is a constant. Find a convenient equivalent Hamiltonian for this problem such that the equations of motion for r and p_r, expressed in terms of a parameter τ, are independent of those for the angular variables. Find r in terms of τ and show that there exists solutions of the equations of motion with constant r. Identify a constant of motion, in addition to p_ϕ and H.

Exercise 4.23. The standard Hamiltonian of a charged particle in the field of a point charge placed at the origin, and a constant field in the z-direction, expressed in parabolic coordinates (u, v, ϕ) [defined by $x = uv\cos\phi$, $y = uv\sin\phi$, $z = \frac{1}{2}(u^2 - v^2)$], is

$$H = \frac{1}{2m}\left(\frac{{p_u}^2 + {p_v}^2}{u^2 + v^2} + \frac{{p_\phi}^2}{u^2 v^2}\right) - \frac{2k}{u^2 + v^2} - \frac{\gamma}{2}(u^2 - v^2),$$

where k and γ are constants. Find a convenient equivalent Hamiltonian and use it to deduce that $A_3 + \frac{1}{2}m\gamma(x^2 + y^2)$ is a constant of motion, where A_3 is the z-component of the Laplace–Runge–Lenz vector (4.52).

Example 4.24 (The Morse potential). In the preceding examples and exercises of this section we have considered systems with two or more degrees of freedom only. In this example we shall consider the one-dimensional motion of a particle in the Morse potential. This system can be defined by the Lagrangian

$$L = \tfrac{1}{2}m\dot{x}^2 - V_0(1 - \mathrm{e}^{-\alpha x})^2,$$

where m, V_0, and α are positive constants. First, we replace the coordinate x by

$$q = \mathrm{e}^{-\alpha x/2}$$

(that is, $x = -(2/\alpha)\ln q$), so that the Lagrangian takes the form

$$L = \frac{m}{2}\left(\frac{2}{\alpha q}\right)^2 \dot{q}^2 - V_0(1 - q^2)^2$$

and the corresponding Hamiltonian is

$$H = \frac{1}{2m}\left(\frac{\alpha q}{2}\right)^2 p^2 + V_0(1 - q^2)^2.$$

The Hamiltonian H is conserved and one readily finds that the equation $H = E$ can be rewritten as

$$\frac{p^2}{2m} + \frac{4(V_0 - E)}{\alpha^2 q^2} + \frac{4V_0 q^2}{\alpha^2} = \frac{8V_0}{\alpha^2}, \tag{4.60}$$

which suggests the definition

$$h_E \equiv \frac{p^2}{2m} + \frac{4(V_0 - E)}{\alpha^2 q^2} + \frac{4V_0 q^2}{\alpha^2}.$$

Then, a straightforward computation gives

$$H - E = \frac{\alpha^2 q^2}{4}\left(h_E - \frac{8V_0}{\alpha^2}\right),$$

which is of the form (4.53) with $f(q, p) = \alpha^2 q^2/4$, $g(x) = x$, and $\varepsilon = 8V_0/\alpha^2$. Hence on the *curve* $H = E$ of the phase space, the time evolution is given by the Hamilton equations (4.54) with $\mathrm{d}\tau = (\alpha^2 q^2/4)\,\mathrm{d}t$ [see (4.55)].

Using the fact that

$$\frac{\mathrm{d}q}{\mathrm{d}\tau} = \frac{\partial h_E}{\partial p} = \frac{p}{m},$$

from (4.60) we obtain the first-order ODE

$$\frac{m}{2}\left(\frac{\mathrm{d}q}{\mathrm{d}\tau}\right)^2 + \frac{4(V_0 - E)}{\alpha^2 q^2} + \frac{4V_0 q^2}{\alpha^2} = \frac{8V_0}{\alpha^2},$$

which allows us to find q as a function of τ [cf. Equation (2.25)].

Exercise 4.25. Making use of the equations derived in Example 4.24, find q as a function of τ, and the relation between τ and t in the case where $E = V_0$.

Example 4.26. Given a time-independent Hamiltonian, $H(q_i, p_i)$, the equations of motion can be written in the form (4.54) with the parameter τ being one of the coordinates q_i, provided that its conjugate momentum, p_i, appears in H.

Indeed, assuming that the equation $H(q_i, p_i) = E$ can be solved for p_1, say, we would obtain $p_1 = F(q_1, \ldots, q_n, p_2, \ldots, p_n, E)$, then letting

$$h_E \equiv p_1 - F(q_1, \ldots, q_n, p_2, \ldots, p_n, E),$$

the hypersurface $H = E$ is also given by $h_E = 0$ and, on this hypersurface [see (4.54)],

$$\frac{\mathrm{d}q_1}{\mathrm{d}\tau} = \frac{\partial h_E}{\partial p_1} = 1,$$

which shows that we can take $\tau = q_1$.

For instance, if

$$H = \frac{1}{2m}\left({p_r}^2 + \frac{{p_\theta}^2}{r^2}\right) + V(r),$$

which corresponds to a particle in a central force field in polar coordinates [see (2.9)], from $H = E$ we obtain $p_\theta = \pm r\sqrt{2m[E - V(r)] - {p_r}^2}$, hence, on the hypersurface $H = E$, the equations of motion can be obtained making use of

$$h_E = p_\theta - r\sqrt{2m[E - V(r)] - {p_r}^2},$$

using θ in place of the time. Thus,

$$\frac{\mathrm{d}r}{\mathrm{d}\theta} = \frac{\partial h_E}{\partial p_r} = \frac{r p_r}{\sqrt{2m[E - V(r)] - {p_r}^2}} = \frac{r^2 p_r}{p_\theta} = \frac{r^2}{p_\theta}\sqrt{2m[E - V(r)] - \frac{{p_\theta}^2}{r^2}}. \tag{4.61}$$

On the other hand,

$$\frac{\mathrm{d}p_\theta}{\mathrm{d}\theta} = -\frac{\partial h_E}{\partial \theta} = 0,$$

that is, p_θ is a constant of motion, and from (4.61) we obtain the equation of the orbit

$$\mathrm{d}\theta = \frac{p_\theta \, \mathrm{d}r}{r^2 \sqrt{2m[E - V(r)] - \dfrac{{p_\theta}^2}{r^2}}}$$

[cf. Equation (2.15)].

Jacobi's Principle

As another application of Proposition 4.16, we consider a Hamiltonian of the form

$$H = \tfrac{1}{2} b_{ij} p_i p_j + V(q_i), \tag{4.62}$$

where the b_{ij} and V are functions of the coordinates q_i only ($i, j = 1, 2, \ldots, n$). One readily sees that $H = E$ is equivalent to $h_E = 1$, if we let

$$h_E = \frac{b_{ij} p_i p_j}{2(E - V)}. \tag{4.63}$$

In fact,

$$H - E = (E - V) \left[\frac{b_{ij} p_i p_j}{2(E - V)} - 1 \right],$$

which is of the form (4.53), with $f = E - V$ and $g(x) = x$.

The Hamiltonian (4.62) corresponds to the Lagrangian

$$L = \tfrac{1}{2} a_{ij} \dot{q}_i \dot{q}_j - V(q_i), \tag{4.64}$$

where the matrix (a_{ij}) is the inverse of (b_{ij}) (see Proposition 4.5) and, in a similar manner, the Hamiltonian (4.63) corresponds to the Lagrangian

$$L_E = \tfrac{1}{2} (E - V) \, a_{ij} q'_i q'_j,$$

where $q'_i \equiv \mathrm{d}q_i/\mathrm{d}\tau$. According to the results of Example 2.14, the last Lagrangian leads to the geodesics of the configuration space, if the length of a curve is defined by

$$\int_{\tau_0}^{\tau_1} \sqrt{(E - V) \, a_{ij} q'_i q'_j} \, \mathrm{d}\tau. \tag{4.65}$$

Thus, the orbits in the configuration space defined by the Hamiltonian (4.62) [or, equivalently, by the Lagrangian (4.64)] are the geodesics of the configuration space with the length of a curve defined by (4.65). This result is known as Jacobi's principle.

Exercise 4.27. Show that by direct substitution of the metric tensor $g_{ij} = (E - V)a_{ij}$ into the geodesic equations

$$g_{ij} q_j'' + \frac{\partial g_{ij}}{\partial q_k} q_k' q_j' - \frac{1}{2} \frac{\partial g_{jk}}{\partial q_i} q_j' q_k' = 0$$

[see Equation (2.67)], where $q_i' \equiv \mathrm{d}q_i/\mathrm{d}\tau$, one obtains equations equivalent to those given by the Lagrangian (4.64).

Further applications of Proposition 4.16 are given in Section 5.2 (see Examples 5.38 and 5.39).

4.3.1 The Kepler Problem Revisited

In this last subsection of this chapter we shall make use of the equivalent Hamiltonian (4.45) to study in some detail the Kepler problem, which was already considered in Section 2.1. We will analyze separately the cases where E is zero, negative, or positive.

In the case where $E = 0$, we have

$$h_E = \frac{{p_u}^2 + {p_v}^2}{2m} \tag{4.66}$$

(which has the *form* of the standard Hamiltonian of a free particle in Cartesian coordinates). Therefore [from Equations (4.48)–(4.49)], p_u and p_v are constant, and

$$u = \frac{p_u}{m}\tau + u_0, \qquad v = \frac{p_v}{m}\tau + v_0, \tag{4.67}$$

where u_0 and v_0 are constants. These are parametric equations of a straight line in the uv-plane passing through (u_0, v_0), in the direction (p_u, p_v). Alternatively, by eliminating the parameter τ from these equations we find that the equation of the orbit in the parabolic coordinates is $p_v u - p_u v =$ const. (which is the equation of a straight line in the uv-plane). The meaning of this constant can be found making use of the relations (4.51). We find

$$p_v u - p_u v = 2(x p_y - y p_x) = 2L_3, \tag{4.68}$$

where L_3 is the z-component of the angular momentum about the origin. (We already knew that L_3 is conserved because the potential is central.) As we shall see now, this means that the orbit (in the xy-plane) is a parabola with its focus at the origin, since [see Equations (4.40) and (4.41)]

$$\begin{aligned} {L_3}^2 &= \tfrac{1}{4}(p_v u - p_u v)^2 \\ &= \tfrac{1}{4}\big[{p_v}^2(r+x) - 2p_u p_v\, y + {p_u}^2(r-x)\big] \\ &= \tfrac{1}{4}\big[({p_u}^2 + {p_v}^2)r - ({p_u}^2 - {p_v}^2)x - 2p_u p_v\, y\big]. \end{aligned} \tag{4.69}$$

The last two terms correspond to the dot product of the position vector $\mathbf{r} = (x, y)$ with the *constant* vector

$$\mathbf{A} \equiv -\tfrac{1}{4}\big({p_u}^2 - {p_v}^2, 2p_u p_v\big), \tag{4.70}$$

whose norm is given by $|\mathbf{A}| = \frac{1}{4}({p_u}^2 + {p_v}^2) = mk$ [see Equation (4.43)]. Hence, (4.69) amounts to

$${L_3}^2 = |\mathbf{A}|r + \mathbf{A}\cdot\mathbf{r}. \tag{4.71}$$

This is the equation of a parabola with its focus at the origin; the constant vector $\mathbf{A}$ defines the axis of the parabola and ${L_3}^2/|\mathbf{A}|$ is the distance between the focus and the directrix of the parabola (see Figure 4.3). (If θ is the angle between $\mathbf{A}$ and $\mathbf{r}$, Equation (4.71) is equivalent to ${L_3}^2 = |\mathbf{A}|r(1+\cos\theta)$, which is the equation of a conic with eccentricity equal to 1, cf. Equation (2.18).)

The shortest distance from the orbit to the center of force (located at the origin) occurs when $\mathbf{r}$ is parallel to $\mathbf{A}$ and has the value

$$r_{\text{min}} = \frac{{L_3}^2}{2|\mathbf{A}|} = \frac{{L_3}^2}{2mk}.$$

(See also Exercise 4.29, below.)

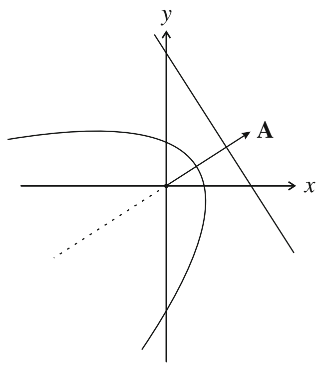


Fig. 4.3 A parabola can be defined with the aid of a point of the plane (the focus of the parabola) and a straight line that does not contain the focus (the directrix). A point of the plane belongs to the parabola if its distance to the focus is equal to its distance to the directrix

Exercise 4.28. Show that a rotation in the uv-plane by an angle θ produces a rotation by an angle 2θ in the xy-plane. (A simple procedure consists in noting that the definition of the parabolic coordinates (4.40) amounts to $x + \mathrm{i}y = \frac{1}{2}(u + \mathrm{i}v)^2$ and that a (passive) rotation in the uv-plane,

$$\begin{aligned} u' &= u\cos\theta + v\sin\theta \\ v' &= -u\sin\theta + v\cos\theta \end{aligned} \tag{4.72}$$

is equivalent to $u' + \mathrm{i}v' = \mathrm{e}^{-\mathrm{i}\theta}(u + \mathrm{i}v)$.)

Exercise 4.29. By means of a rotation in the uv-plane (which, according to Exercise 4.28, corresponds to a rotation in the xy-plane), we can make $p_u = 0$. Show that the orbit is then given by the parametric equations

$$x = \frac{{L_3}^2}{2mk} - \frac{2k\tau^2}{m}, \qquad y = \frac{2L_3\tau}{m}.$$

(Thus, τ is proportional to the distance from the axis of the parabola to the particle.)

From Equations (4.47) and (4.67), we find that the time, t, is related to the parameter τ by

$$\mathrm{d}t = (u^2 + v^2)\,\mathrm{d}\tau = \left[\left(\frac{p_u}{m}\tau + u_0\right)^2 + \left(\frac{p_v}{m}\tau + v_0\right)^2\right]\mathrm{d}\tau, \tag{4.73}$$

where u_0, v_0 are the parabolic coordinates of the particle at $\tau = 0$. Thus, choosing (u_0, v_0) as the closest point of the orbit to the origin (see Figure 4.4), we have

$$p_u u_0 + p_v v_0 = 0$$

and from (4.73) we obtain

$$\mathrm{d}t = \left(\frac{{p_u}^2 + {p_v}^2}{m^2}\tau^2 + {u_0}^2 + {v_0}^2\right)\mathrm{d}\tau = \left(\frac{4k}{m}\tau^2 + 2r_{\min}\right)\mathrm{d}\tau, \tag{4.74}$$

hence, we can take

$$t = \frac{4k}{3m}\tau^3 + 2r_{\min}\tau \tag{4.75}$$

(so that $t = 0$ and $\tau = 0$ when $r = r_{\min}$).

Exercise 4.30. Show that the time is related to the distance to the origin by

$$\pm t = \frac{2}{3}\sqrt{\frac{m}{2k}}\,(r + 2r_{\min})\,\sqrt{r - r_{\min}}.$$

(Cf. Exercise 2.1.)

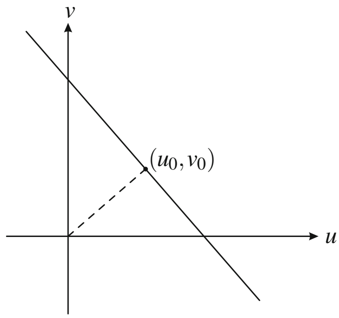


Fig. 4.4 When $E = 0$, the orbit in the uv-plane is a straight line with direction vector (p_u, p_v) [see Equations (4.67)]. The point (u_0, v_0) is chosen as the point of the straight line closest to the origin

Exercise 4.31. Show that the vector **A**, defined in (4.70), can also be expressed as

$$\mathbf{A} = \mathbf{p} \times \mathbf{L} - \frac{mk}{r}\mathbf{r}, \tag{4.76}$$

identifying a vector (a_1, a_2) in the plane with the vector $(a_1, a_2, 0)$, and $\mathbf{L} = (0, 0, L_3)$. The vector **A** is the *Laplace–Runge–Lenz vector* or *Runge–Lenz vector*.

Exercise 4.32. Making use of the solution of the equations of motion in terms of the coordinates u, v, p_u, p_v, given by (4.67), with the aid of (4.51) one can find the trajectory followed by the linear momentum of the particle, (p_x, p_y), as a function of τ. Substituting (4.67) into (4.51) and eliminating the parameter τ, show that the curve described by the momentum (called the *hodograph*) is an arc of the circle passing through the origin given by

$$\left(p_x + \frac{A_2}{L_3}\right)^2 + \left(p_y - \frac{A_1}{L_3}\right)^2 = \frac{{A_1}^2 + {A_2}^2}{{L_3}^2} = \left(\frac{mk}{L_3}\right)^2 \tag{4.77}$$

(provided that $L_3 \neq 0$; when $L_3 = 0$, the hodograph is part of a straight line passing through the origin).

The center of the circle (4.77) is located at the point $(-A_2/L_3, A_1/L_3)$, which can be expressed in the form

$$\frac{\mathbf{L} \times \mathbf{A}}{\mathbf{L}^2}. \tag{4.78}$$

This constant vector is known as *Hamilton's vector*. (Here, again, $\mathbf{L} = (0, 0, L_3)$, and a vector (a_1, a_2) in the plane is identified with the vector $(a_1, a_2, 0)$.)

In the case where $E < 0$, h_E has the form of the Hamiltonian of a *two-dimensional isotropic harmonic oscillator* with angular frequency

$$\omega \equiv \sqrt{-\frac{2E}{m}} \tag{4.79}$$

[compare Equations (4.16) and (4.45)]. Hence, in the uv-plane, the orbit is an ellipse centered at the origin (see Section 2.1) and, by rotating the axes u and v, if necessary, we can assume that the solution of the equations of motion is given by

$$u = \beta \cos \omega\tau, \qquad v = \alpha \sin \omega\tau, \tag{4.80}$$

where α and β are two real numbers with $|\alpha| \geqslant |\beta|$ ($|\alpha|$ and $|\beta|$ are the semiaxes of the ellipse, see Figure 4.5). (It is convenient to allow for negative values of α and β; the sign of $\alpha\beta$ determines the sense in which the orbit is traversed.) Then, from (4.48) and (4.49),

$$p_u = -m\omega\beta \sin \omega\tau, \qquad p_v = m\omega\alpha \cos \omega\tau \tag{4.81}$$

and [see Equations (4.40)]

$$x = \tfrac{1}{4}\big[\beta^2 - \alpha^2 + (\alpha^2 + \beta^2) \cos 2\omega\tau\big], \qquad y = \tfrac{1}{2}\alpha\beta \sin 2\omega\tau, \tag{4.82}$$

which are parametric equations of an ellipse in the xy-plane, with semiaxes a and $|b|$, where

$$a \equiv \tfrac{1}{4}(\alpha^2 + \beta^2), \qquad b \equiv \tfrac{1}{2}\alpha\beta, \tag{4.83}$$

centered at the point $\big((\beta^2 - \alpha^2)/4, 0\big)$. Hence, the distance from the center of the ellipse to the foci is $c = \sqrt{a^2 - b^2} = (\alpha^2 - \beta^2)/4$, which means that one of the foci of the ellipse coincides with the origin (see Figure 4.5). Equations (4.82) show that the sign of b determines the sense in which the ellipse is traversed. See also Equation (4.85).

Since h_E has the value $2k$, substituting (4.80) and (4.81) into (4.45), we find that

$$\frac{\alpha^2 + \beta^2}{4} = \frac{k}{m\omega^2} = -\frac{k}{2E},$$

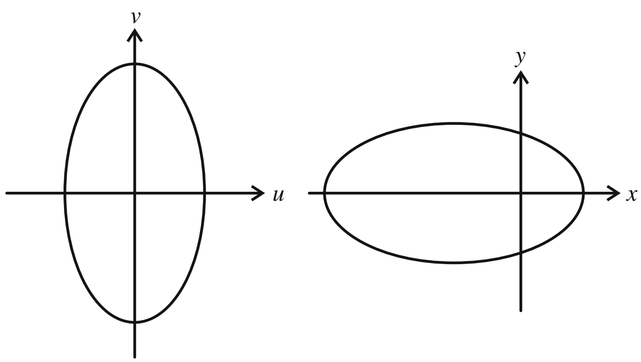


Fig. 4.5 Equations (4.80) are parametric equations of an ellipse centered at the origin in the uv-plane, with semiaxes $|\alpha|$ and $|\beta|$. Equations (4.82) correspond to an ellipse with one of its foci at the origin in the xy-plane

that is,

$$E = -\frac{k}{2a}, \tag{4.84}$$

which shows that the energy E is a function of the major semiaxis of the orbit only [cf. (2.21)]. Similarly, from (4.68) we obtain $2L_3 = m\omega\alpha\beta$, which amounts to

$$b = \frac{L_3}{m\omega} = \frac{L_3}{\sqrt{-2mE}}. \tag{4.85}$$

According to Equations (4.47), (4.80), and (4.83), the time is related to the parameter τ by

$$\begin{aligned} \mathrm{d}t &= (u^2 + v^2)\,\mathrm{d}\tau \\ &= (\beta^2 \cos^2 \omega\tau + \alpha^2 \sin^2 \omega\tau)\,\mathrm{d}\tau = 2(a - c\,\cos 2\omega\tau)\,\mathrm{d}\tau \\ &= 2a(1 - e\,\cos 2\omega\tau)\,\mathrm{d}\tau, \end{aligned}$$

where $e \equiv c/a$ is the eccentricity of the ellipse. Hence, we can take

$$t = 2a\tau - \frac{ea}{\omega}\sin 2\omega\tau \tag{4.86}$$

(so that $t = 0$ when $\tau = 0$). From (4.82) we see that, in a complete period, τ is increased by π/ω and from (4.86) it follows that the period of the motion, T, is

$$T = \frac{2\pi a}{\omega}.$$

With the aid of this expression [and Equations (4.79) and (4.84)] we can readily obtain Kepler's third law,

$$T^2 = \frac{4\pi^2 a^2}{\omega^2} = \frac{4\pi^2 a^2}{-2E/m} = -\frac{4\pi^2 a^2 m}{-k/a} = \frac{4\pi^2 m}{k} a^3.$$

In place of the parameter τ, it is convenient to use the dimensionless variable

$$\psi \equiv 2\omega\tau, \tag{4.87}$$

then

$$x = a(\cos\psi - e), \qquad y = b\sin\psi, \qquad \frac{2\pi}{T}t = \psi - e\sin\psi. \tag{4.88}$$

The last equation in (4.88), which relates the time with the auxiliary parameter ψ, is known as *Kepler's equation* [cf. Equation (2.24)]. In order to find the coordinates

x and y in terms of t, one would have to solve Kepler's equation for ψ as a function of t, which is not a simple task when $e \neq 0$ (the interested reader may consult Fasano and Marmi [8, Sects. 5.4–5.5], Bowman [1, §109] and, for a more detailed discussion, Colwell [3]). The geometrical meaning of ψ can be seen in Figure 4.6.

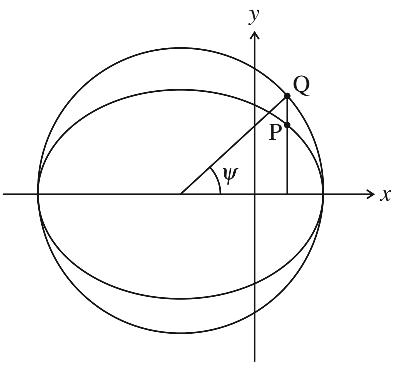


Fig. 4.6 The vertical line passing through the point P $= (a\cos\psi - ae, b\sin\psi)$ of the orbit intersects the circle of radius a and center $(-ae, 0)$ at the point Q $= (a\cos\psi - ae, a\sin\psi)$. Hence, the eccentric anomaly, ψ, is the angle between the x-axis and the radius of the circle passing through Q

A similar treatment of the Kepler problem with negative energy, making use of the Lagrangian formalism, can be found in Pars [11, Sect. 26.10].

Exercise 4.33. Consider the Hamiltonian (4.45) with $E > 0$. Show that the corresponding equations of motion have solutions of the form $u = \beta\cosh\omega\tau$, $v = \alpha\sinh\omega\tau$, where α and β are real constants, with $|\alpha| > |\beta|$ and $\omega \equiv \sqrt{2E/m}$. Show that, in the xy-plane, the orbit is a branch of a hyperbola with semiaxes a and $|b|$, with $a = \frac{1}{4}(\alpha^2 - \beta^2)$, $b = \frac{1}{2}\alpha\beta$. Show that also in this case the energy is a function of the major semiaxis only and obtain the analog of the Kepler equation.

Exercise 4.34. Making use of vector calculus show that in the case of the Kepler problem, characterized by the equations of motion

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\frac{k\,\mathbf{r}}{r^3},$$

where k is a constant and $r = |\mathbf{r}|$, the Laplace–Runge–Lenz vector (4.76) is conserved, for any value of the energy, E. Show that, for any value of E, the hodograph is a circle, or part of a circle, centered at $\mathbf{L} \times \mathbf{A}/\mathbf{L}^2$, of radius

$$R = \frac{mk}{|\mathbf{L}|}.$$

(Cf. Equations (4.78) and (4.77).)

Chapter 5
Canonical Transformations

One of the main reasons why the Hamiltonian formalism is more powerful than the Lagrangian formalism is that the set of coordinate transformations that leave invariant the form of the Hamilton equations is much broader than the set of coordinate transformations that leave invariant the form of the Lagrange equations. As we have seen, *any* transformation of coordinates, $q'_i = q'_i(q_j, t)$, preserves the form of the Lagrange equations (in fact, as shown in Section 2.4, the form of the Lagrange equations is invariant also under the coordinate transformations $q'_i = q'_i(q_j, t)$, $t' = t'(q_j, t)$, if the Lagrangian is suitably transformed).

As we shall see, not every coordinate transformation in the extended phase space preserves the form of the Hamilton equations; however, there exists a wide class of coordinate transformations, called canonical transformations, which maintain the form of the Hamilton equations for any Hamiltonian function. Furthermore, any canonical transformation can be obtained from a single real-valued function of $2n + 1$ variables, where n is the number of degrees of freedom of the system, which is therefore a generating function of the transformation.

In Sections 5.1 and 5.2 the canonical transformations are defined and the existence of generating functions is proved. The proof of the existence of a generating function for each canonical transformation given below is direct and elementary, and, by contrast with the standard approach, it is not based on the calculus of variations.

In order to present the ideas and results in a simple way, we consider first the case where there is only one degree of freedom. In Section 5.3 we study the one-parameter families of canonical transformations, and in Section 5.4, we show that *any* constant of motion (that may depend explicitly on the time) is associated with a one-parameter group of canonical transformations that leave invariant the Hamiltonian.

In Section 5.5 we discuss the coordinate transformations that preserve the form of the Hamilton equations for a given Hamiltonian, but are not canonical. We show that each of these transformations leads to a set of constants of motion.

G. F. Torres del Castillo, *An Introduction to Hamiltonian Mechanics*, Birkhäuser Advanced Texts Basler Lehrbücher, https://doi.org/10.1007/978-3-319-95225-3_5

5.1 Systems with One Degree of Freedom

We consider a system with one degree of freedom, described by a Hamiltonian function $H(q, p, t)$. This means that the equations of motion of the system are given by the Hamilton equations

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial H}{\partial p}, \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial q}. \tag{5.1}$$

We want to find other coordinate systems, (Q, P, t), of the extended phase space such that the equations of motion also have the *form* of the Hamilton equations; that is, we look for the coordinate transformations, $Q = Q(q, p, t)$, $P = P(q, p, t)$, such that Equations (5.1) are equivalent to

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \frac{\partial K}{\partial P}, \qquad \frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{\partial K}{\partial Q}, \tag{5.2}$$

where K is some function.

Assuming that the transformation $Q = Q(q, p, t)$, $P = P(q, p, t)$ is differentiable (as well as invertible), making use repeatedly of the chain rule, and of Equations (5.1) and (5.2) we find that the function K would have to satisfy

$$\begin{aligned}
\frac{\partial K}{\partial q} &= \frac{\partial K}{\partial Q}\frac{\partial Q}{\partial q} + \frac{\partial K}{\partial P}\frac{\partial P}{\partial q} \\
&= -\frac{\mathrm{d}P}{\mathrm{d}t}\frac{\partial Q}{\partial q} + \frac{\mathrm{d}Q}{\mathrm{d}t}\frac{\partial P}{\partial q} \\
&= -\left(\frac{\partial P}{\partial t} + \frac{\partial P}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{\partial P}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t}\right)\frac{\partial Q}{\partial q} + \left(\frac{\partial Q}{\partial t} + \frac{\partial Q}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{\partial Q}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t}\right)\frac{\partial P}{\partial q} \\
&= -\left(\frac{\partial P}{\partial t} + \frac{\partial P}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial P}{\partial p}\frac{\partial H}{\partial q}\right)\frac{\partial Q}{\partial q} + \left(\frac{\partial Q}{\partial t} + \frac{\partial Q}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial H}{\partial q}\right)\frac{\partial P}{\partial q} \\
&= \{Q, P\}\frac{\partial H}{\partial q} + \frac{\partial Q}{\partial t}\frac{\partial P}{\partial q} - \frac{\partial P}{\partial t}\frac{\partial Q}{\partial q},
\end{aligned} \tag{5.3}$$

where we have made use of the definition of the Poisson bracket (4.21)

$$\{f, g\} \equiv \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial q}. \tag{5.4}$$

In a similar way, we obtain

$$\frac{\partial K}{\partial p} = \{Q, P\}\frac{\partial H}{\partial p} + \frac{\partial Q}{\partial t}\frac{\partial P}{\partial p} - \frac{\partial P}{\partial t}\frac{\partial Q}{\partial p}. \tag{5.5}$$

Thus, given H, Q, and P in terms of q, p, and t, Equations (5.3) and (5.5) determine the partial derivatives of K with respect to q and p, but these two expressions might be incompatible; the (local) existence of a function K satisfying (5.3) and (5.5) is equivalent to the fulfillment of the integrability condition

$$0 = \frac{\partial}{\partial q}\frac{\partial K}{\partial p} - \frac{\partial}{\partial p}\frac{\partial K}{\partial q},$$

that is, assuming that the partial derivatives of P, Q, and H commute,

$$\begin{aligned} 0 &= \frac{\partial}{\partial q}\left[\{Q,P\}\frac{\partial H}{\partial p} + \frac{\partial Q}{\partial t}\frac{\partial P}{\partial p} - \frac{\partial P}{\partial t}\frac{\partial Q}{\partial p}\right] - \frac{\partial}{\partial p}\left[\{Q,P\}\frac{\partial H}{\partial q} + \frac{\partial Q}{\partial t}\frac{\partial P}{\partial q} - \frac{\partial P}{\partial t}\frac{\partial Q}{\partial q}\right] \\ &= \{\{Q,P\},H\} + \frac{\partial\{Q,P\}}{\partial t}, \end{aligned}$$

which means that $\{Q, P\}$ is a constant of motion [cf. Equation (4.23)]. Thus we have proved the following result.

Proposition 5.1. *The coordinate transformation $Q = Q(q, p, t)$, $P = P(q, p, t)$ preserves the form of the Hamilton equations if and only if $\{Q, P\}$ is a constant of motion.*

Example 5.2 (One-dimensional harmonic oscillator). In the case of the transformation

$$Q = \arctan\frac{m\omega q}{p}, \qquad P = \sqrt{p^2 + m^2\omega^2 q^2}, \tag{5.6}$$

the Poisson bracket $\{Q, P\}$ is equal to $m\omega(p^2+m^2\omega^2q^2)^{-1/2}$, which is a constant of motion if the Hamiltonian is, for instance, $H = \frac{1}{2m}(p^2 + m^2\omega^2q^2)$ (corresponding to a one-dimensional harmonic oscillator). The Hamilton equations (5.1) yield $\mathrm{d}q/\mathrm{d}t = p/m$, and $\mathrm{d}p/\mathrm{d}t = -m\omega^2 q$ and, therefore,

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \omega, \qquad \frac{\mathrm{d}P}{\mathrm{d}t} = 0, \tag{5.7}$$

which can be expressed as the Hamilton equations (5.2) with, e.g., $K = \omega P$. (Alternatively, from Equations (5.3) and (5.5) we have

$$\frac{\partial K}{\partial q} = m^2\omega^3 q(p^2 + m^2\omega^2q^2)^{-1/2}, \qquad \frac{\partial K}{\partial p} = \omega p(p^2 + m^2\omega^2q^2)^{-1/2},$$

which lead to $K = \omega(p^2+m^2\omega^2q^2)^{1/2} = \omega P$, up to an additive function of t only.)

The solution of the equations of motion (5.7) is $P = P_0$ and $Q = \omega t + Q_0$, where P_0 and Q_0 are constants. Then, inverting the relations (5.6), we obtain the solution of the equations of motion in terms of the original coordinates: $m\omega q = P_0 \sin(\omega t + Q_0)$, $p = P_0 \cos(\omega t + Q_0)$.

Of course, if we already have the equations of motion (5.7), which are easily integrated, it might seem of little interest to see if they can be expressed in the form of the Hamilton equations or not. However, as we have confirmed in the preceding chapter, and as we shall continue verifying in what follows, there are many advantages in expressing a system of equations in the Hamiltonian form. On the other hand, by contrast with (5.7), in most cases, we will not find directly the equations of motion in terms of the new variables. It is simpler to find the new Hamiltonian, which determines the entire new system of equations.

Exercise 5.3. Show that if

$$Q = \tfrac{1}{2}mq^2 - qpt, \qquad P = p,$$

and the Hamiltonian is that of a free particle,

$$H = \frac{p^2}{2m},$$

then $\{Q, P\}$ is a (nontrivial) constant of motion. Find a Hamiltonian K for the new coordinates.

Now, a function of (q, p, t) is a constant of motion or not, depending on the equations of motion, that is, depending on H. Thus, if we look for a coordinate transformation that preserves the form of the Hamilton equations regardless of the form of H, the Poisson bracket $\{Q, P\}$ must be a constant of motion regardless of the form of H, and this means that $\{Q, P\}$ has to be a *trivial constant* (i.e., a function that does not depend on q, p, or t). Since a nonzero constant factor can be absorbed in the definition of the coordinates, it is enough to consider transformations such that

$$\{Q, P\} = 1. \tag{5.8}$$

(If $\{Q, P\} = 0$, then (Q, P, t) is not functionally independent and cannot be used as a coordinate system.) The coordinate transformations satisfying Equation (5.8) are called *canonical transformations*, and the coordinate transformations such that $\{Q, P\}$ is a nontrivial constant of motion are sometimes called *canonoid transformations*. In what follows we shall consider almost exclusively canonical transformations. (The canonoid transformations will be treated in some detail in Section 5.5.)

There are several additional reasons to consider canonical transformations only; one of them is that if $(q_i(t), p_i(t))$ is the solution of the Hamilton equations, for a given Hamiltonian, then the relation between $(q_i(t_0), p_i(t_0))$ and $(q_i(t), p_i(t))$, for any t_0 and t, is a canonical transformation (see pp. 186–190, below). Another reason is that the Poisson brackets (5.4) are invariant under these transformations, in the sense that

$$\frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial q} = \frac{\partial f}{\partial Q}\frac{\partial g}{\partial P} - \frac{\partial f}{\partial P}\frac{\partial g}{\partial Q},$$

for any pair of functions f, g, defined on the extended phase space, if and only if Equation (5.8) holds. In fact, making use of the chain rule, one can readily show that

$$\frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial q} = \{Q, P\}\left(\frac{\partial f}{\partial Q}\frac{\partial g}{\partial P} - \frac{\partial f}{\partial P}\frac{\partial g}{\partial Q}\right). \tag{5.9}$$

It may be noticed that this equation is equivalent to the relation between Jacobians

$$\frac{\partial(f, g)}{\partial(q, p)} = \frac{\partial(Q, P)}{\partial(q, p)}\frac{\partial(f, g)}{\partial(Q, P)}, \tag{5.10}$$

and that Equation (5.8) amounts to

$$\frac{\partial(Q, P)}{\partial(q, p)} = 1. \tag{5.11}$$

The above calculation shows that, conversely, the coordinate transformations that leave the Poisson bracket invariant are the canonical transformations (see also Proposition 5.44, below).

The canonical transformations also arise in a natural way when one performs an *arbitrary* coordinate transformation $Q = Q(q, t)$ in the extended configuration space. According to (4.17), a change of the coordinate in the extended configuration space, $Q = Q(q, t)$, leads to a new canonical momentum, P, given by

$$P = p\frac{\partial q}{\partial Q}.$$

Then, using the definition of the Poisson bracket and that $\partial Q/\partial p = 0$, we have

$$\{Q, P\} = \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} = \frac{\partial Q}{\partial q}\frac{\partial q}{\partial Q} = 1.$$

Another reason to restrict the discussion to canonical coordinates is that *any* constant of motion can be associated with a one-parameter group of canonical transformations that leaves the Hamiltonian invariant (see Section 5.4, below).

Thus, restricting ourselves to canonical transformations, Equations (5.3) and (5.5) reduce to

$$\frac{\partial(K - H)}{\partial q} = \frac{\partial Q}{\partial t}\frac{\partial P}{\partial q} - \frac{\partial P}{\partial t}\frac{\partial Q}{\partial q}, \qquad \frac{\partial(K - H)}{\partial p} = \frac{\partial Q}{\partial t}\frac{\partial P}{\partial p} - \frac{\partial P}{\partial t}\frac{\partial Q}{\partial p}. \tag{5.12}$$

These equations show that, given a canonical transformation (so that the right-hand sides of Equations (5.12) can be calculated), the difference $K - H$ is determined *up* *to an additive function of t only*. In particular, if the new coordinates do not depend on t, we can take $K = H$. It may be noticed that the expressions on the right-hand sides of Equations (5.12) are the Jacobians $\frac{\partial(Q,P)}{\partial(t,q)}$ and $\frac{\partial(Q,P)}{\partial(t,p)}$.

Example 5.4. The coordinate transformation

$$Q = tq^2, \qquad P = \frac{p}{2tq} \tag{5.13}$$

is canonical since

$$\{Q, P\} = \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} = 2tq\frac{1}{2tq} - 0\left(-\frac{p}{2tq^2}\right) = 1.$$

(Except at $tq = 0$, where the transformation is not defined.) According to Equations (5.12),

$$\frac{\partial(K-H)}{\partial q} = \frac{p}{2t}, \qquad \frac{\partial(K-H)}{\partial p} = \frac{q}{2t},$$

hence, $K = H + pq/2t + f(t) = H + PQ/t + f(t)$, where $f(t)$ is a function of t only, which can be taken equal to zero without affecting Equations (5.2).

Exercise 5.5. Show that the coordinate transformation

$$Q = q + \tfrac{1}{2}gt^2, \qquad P = p + mgt,$$

where m and g are constants, is canonical and find the expression for K in terms of H.

Exercise 5.6. Show that the coordinate transformation

$$\begin{aligned} Q &= q(\cos\omega t + \omega t \sin\omega t) + \frac{p}{m\omega}(\omega t\cos\omega t - \sin\omega t), \\ P &= m\omega q \sin\omega t + p\cos\omega t, \end{aligned} \tag{5.14}$$

where m and ω are constants, is canonical. Show that if

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2,$$

then, up to an additive function of t only,

$$K = \frac{P^2}{2m}.$$

Therefore, under the coordinate transformation (5.14) the equations of motion of a one-dimensional harmonic oscillator are converted into those of a free particle.

(It may be noticed that in the limit as ω goes to zero, the coordinate transformation (5.14) reduces to the identity and H becomes the standard Hamiltonian of a free particle.)

It should be remarked that even though the canonical transformations are defined taking into account the form of the equations of motion, a canonical transformation, such as (5.14), is a coordinate transformation in the extended phase space and, among other things, the coordinates, q, p, and t are *independent variables* (this point is essential in what follows).

One can convince oneself (for example, looking for explicit examples) that it is not easy to find canonical transformations, guided solely by the definition (5.8), and that it is even more difficult to find a canonical transformation useful to simplify the Hamilton equations for a given Hamiltonian. As we shall see in the next paragraphs, at least the first of these difficulties can be easily surmounted. We will be able to construct plenty of canonical transformations very easily with the aid of the generating functions defined below.

Generating Functions

Another reason for centering the attention to the canonical transformations is that any of these transformations is determined by a single real-valued function.

Equations (5.8) and (5.12) are necessary and sufficient conditions for the local existence of a function, F_1, defined on some open region of the extended phase space, such that

$$p\mathrm{d}q - H\mathrm{d}t - (P\mathrm{d}Q - K\mathrm{d}t) = \mathrm{d}F_1 \tag{5.15}$$

(that is, the function F_1 may not be defined in all the extended phase space, we can only assure its existence in some neighborhood of each point of the extended phase space where the left-hand side of (5.15) is defined). In fact, writing the left-hand side of (5.15) in terms of the single set of coordinates (q, p, t) of the extended phase space,

$$\left(p - P\frac{\partial Q}{\partial q}\right)\mathrm{d}q - P\frac{\partial Q}{\partial p}\mathrm{d}p + \left(K - H - P\frac{\partial Q}{\partial t}\right)\mathrm{d}t,$$

and applying the standard criterion for a linear (or Pfaffian) differential form to be exact, one finds that the left-hand side of (5.15) is locally exact if and only if Equations (5.8) and (5.12) are satisfied. For instance, by considering the crossed partial derivatives of the coefficients of dq and dt, we obtain

$$\frac{\partial}{\partial q}\left(K - H - P\frac{\partial Q}{\partial t}\right) - \frac{\partial}{\partial t}\left(p - P\frac{\partial Q}{\partial q}\right) = \frac{\partial(K - H)}{\partial q} - \frac{\partial P}{\partial q}\frac{\partial Q}{\partial t} + \frac{\partial P}{\partial t}\frac{\partial Q}{\partial q},$$

which is equal to zero if and only if (5.12) holds.

If (q, Q, t) is functionally independent (which amounts to say that p can be expressed as a function of q, Q, t), then the function F_1 appearing in

Equation (5.15) can be expressed in terms of q, Q, and t (in a *unique* way), and from Equation (5.15) it follows that

$$p = \left(\frac{\partial F_1}{\partial q}\right)_{Q,t}, \qquad P = -\left(\frac{\partial F_1}{\partial Q}\right)_{q,t}, \qquad K - H = \left(\frac{\partial F_1}{\partial t}\right)_{q,Q}, \tag{5.16}$$

and, necessarily, $\partial^2 F_1/\partial q \partial Q \neq 0$ (otherwise q and p would not be independent). Conversely, given a function $F_1(q, Q, t)$ such that $\partial^2 F_1/\partial q \partial Q \neq 0$, Equation (5.16) can be locally inverted to find Q and P in terms of q, p, and t. In this way, F_1 is a *generating function* of a canonical transformation. Thus, we can easily construct canonical transformations; we only have to choose a function $F_1(q, Q, t)$ such that $\partial^2 F_1/\partial q \partial Q \neq 0$. What is desirable is to find a canonical transformation that simplifies a given problem or that relates two problems of interest.

Example 5.7. From Equations (5.14) we see that, except at the points where $\tan \omega t = \omega t$, (q, Q, t) is functionally independent (for instance, Q cannot be written as a function of q and t only), and that p and P are given by

$$p = \frac{m\omega[Q - q(\cos \omega t + \omega t \sin \omega t)]}{\omega t \cos \omega t - \sin \omega t}, \qquad P = \frac{m\omega(Q \cos \omega t - q)}{\omega t \cos \omega t - \sin \omega t}$$

in terms of (q, Q, t). Substituting these expressions into the left-hand sides of the first two equations (5.16) we find that

$$F_1 = \frac{m\omega}{\omega t \cos \omega t - \sin \omega t}\left[qQ - \tfrac{1}{2}Q^2 \cos \omega t - \tfrac{1}{2}q^2(\cos \omega t + \omega t \sin \omega t)\right] + f(t),$$

where $f(t)$ is a function of t only. Then, the last equation in (5.16) gives

$$K - H = -\frac{m\omega^2}{2}\left(\frac{Q \sin \omega t - q\omega t}{\omega t \cos \omega t - \sin \omega t}\right)^2 + f'(t)$$

and with the aid of this expression we can find the new Hamiltonian, K, for any given Hamiltonian H. (The term $f'(t)$ has no effect on the Hamilton equations and can be dropped.)

Example 5.8. The function

$$F_1 = -Q \cot \frac{q}{2}$$

satisfies the condition $\partial^2 F_1/\partial q \partial Q \neq 0$. According to Equations (5.16), the canonical transformation generated by F_1 is such that $p = \frac{1}{2}Q \csc^2 \frac{1}{2}q$ and $P = \cot \frac{1}{2}q$, hence, $Q = 2p \sin^2 \frac{1}{2}q$, $P = \cot \frac{1}{2}q$. Furthermore, $K = H$. (Note that in this case we *assume* a priori that (q, Q, t) is functionally independent.)

On the other hand, if (q, Q, t) is functionally dependent (that is, Q can be expressed as a function of q and t only, as in Example 5.4), the function F_1 appearing in Equation (5.15) can be written in infinitely many ways in terms of q, Q, and t, and the first two equations in (5.16) make no sense (since, e.g., keeping q and t constant in the partial differentiation with respect to Q, would make Q also constant). In such a case, the set of variables (q, P, t) [as well as (p, Q, t)] is necessarily functionally independent (otherwise (Q, P, t) or (q, p, t) would be functionally dependent). Using the fact that $P\mathrm{d}Q = \mathrm{d}(PQ) - Q\mathrm{d}P$, we can write Equation (5.15) in the equivalent form

$$p\mathrm{d}q - H\mathrm{d}t + Q\mathrm{d}P + K\mathrm{d}t = \mathrm{d}F_2, \tag{5.17}$$

where $F_2 \equiv F_1 + PQ$. Then, the function F_2 can be expressed in a *unique* way as a function of q, P, and t, and from (5.17) it follows that the canonical transformation is determined by

$$p = \left(\frac{\partial F_2}{\partial q}\right)_{P,t}, \qquad Q = \left(\frac{\partial F_2}{\partial P}\right)_{q,t}, \qquad K - H = \left(\frac{\partial F_2}{\partial t}\right)_{q,P}, \tag{5.18}$$

with $\partial^2 F_2/\partial q \partial P \neq 0$. F_2 is then a generating function of the canonical transformation. It should be noted that Equations (5.18) are applicable if (q, P, t) is functionally independent, regardless of whether (q, Q, t) is functionally independent or not (see Example 5.9, below). Conversely, any function $F_2(q, P, t)$ such that $\partial^2 F_2/\partial q \partial P \neq 0$ defines a canonical transformation by means of the first two equations in (5.18). The relation between F_1 and F_2 is another example of a Legendre transformation (compare the relation $F_2 = -Q(\partial F_1/\partial Q) + F_1$ with the definition of the Hamiltonian, $H = \dot{q}_i(\partial L/\partial \dot{q}_i) - L$).

Example 5.9. In the case of the coordinate transformation (5.14), except for some values of t, the sets (q, Q, t) and (q, P, t) can be used as coordinates of the extended phase space. For instance, from (5.14) we have

$$p = \frac{P - m\omega q \sin\omega t}{\cos\omega t}, \qquad Q = \frac{q}{\cos\omega t} + \frac{P(\omega t \cos\omega t - \sin\omega t)}{m\omega \cos\omega t},$$

which explicitly show that when $\cos\omega t \neq 0$, (q, P, t) can be used as coordinates of the extended phase space (since (q, p, t) can be expressed in terms of (q, P, t)). Substituting these expressions on the left-hand sides of the first two equations (5.18) one finds that

$$F_2 = \frac{1}{\cos\omega t}\left[qP - \frac{m\omega q^2}{2}\sin\omega t + \frac{P^2}{2m\omega}(\omega t \cos\omega t - \sin\omega t)\right] + f(t),$$

where $f(t)$ is a function of t only.

In a similar way, if (p, Q, t) is functionally independent, one can make use of a generating function $F_3 \equiv F_1 - pq$. In that case, F_3 can be expressed in a unique way in terms of (p, Q, t), and from (5.15) we have

$$-q\mathrm{d}p - H\mathrm{d}t - (P\mathrm{d}Q - K\mathrm{d}t) = \mathrm{d}F_3,$$

which implies that

$$q = -\left(\frac{\partial F_3}{\partial p}\right)_{Q,t}, \qquad P = -\left(\frac{\partial F_3}{\partial Q}\right)_{p,t}, \qquad K - H = \left(\frac{\partial F_3}{\partial t}\right)_{p,Q}$$

and, when (p, P, t) is functionally independent, defining $F_4 \equiv F_1 - pq + PQ$, one obtains

$$-q\mathrm{d}p - H\mathrm{d}t + Q\mathrm{d}P + K\mathrm{d}t = \mathrm{d}F_4$$

therefore

$$q = -\left(\frac{\partial F_4}{\partial p}\right)_{P,t}, \qquad Q = \left(\frac{\partial F_4}{\partial P}\right)_{p,t}, \qquad K - H = \left(\frac{\partial F_4}{\partial t}\right)_{p,P}.$$

Example 5.10. We shall consider the coordinate transformation (5.13) again,

$$Q = tq^2, \qquad P = \frac{p}{2tq}.$$

Substituting these expressions into the left-hand side of (5.15) we obtain

$$\begin{aligned} p\mathrm{d}q - H\mathrm{d}t - (P\mathrm{d}Q - K\mathrm{d}t) &= p\mathrm{d}q - H\mathrm{d}t - \frac{p}{2tq}\,\mathrm{d}\big(tq^2\big) + K\mathrm{d}t \\ &= \left(K - H - \frac{pq}{2t}\right)\mathrm{d}t, \end{aligned}$$

which is an exact differential if and only if $K - H - pq/2t$ is a function of t only. (Applying again the criterion of the equality of the crossed partial derivatives, it follows that the differential form $A\,\mathrm{d}t + 0\,\mathrm{d}q + 0\,\mathrm{d}p$ is exact if and only if A is a function of t only.) Choosing $K = H + pq/2t$, we obtain $\mathrm{d}F_1 = 0$ (that is, F_1 is a trivial constant).

In the present case (q, Q, t) *is not* functionally independent but since q and p can be written in terms of (q, P, t), these last variables can be used as coordinates in the extended phase space, and the canonical transformation can be expressed in terms of a generating function of the type F_2. Choosing $F_1 = 0$, we obtain $F_2 = PQ = Ptq^2$ and it can be readily verified that Equations (5.18) reproduce the formulas (5.13).

Alternatively, expressing p and Q in terms of (q, P, t), we have

$$p = 2tPq, \qquad Q = tq^2,$$

and comparing with the first two equations in (5.18) we find that $F_2 = Ptq^2 + \phi(t)$, where $\phi(t)$ is a function of t only. Then, from the last equation in (5.18) we obtain $K = H + Pq^2 + \phi'(t) = H + pq/2t + \phi'(t)$, which is equivalent to the previous result.

Note that, in the present example, the sets (q, P, t), (p, Q, t), and (p, P, t) are functionally independent and, therefore, the transformation can be obtained by means of generating functions of type F_2, F_3, and F_4.

Exercise 5.11. Find a generating function of type F_2 for the canonical transformation considered in Exercise 5.5. Is it possible to find generating functions of types F_1, F_3, or F_4 for this transformation?

Exercise 5.12. Does the function

$$F_1 = q(k \ln t - \ln Q),$$

where k is a constant, define a canonical transformation? If it does, find Q and P in terms of (q, p, t) and K in terms of H.

Example 5.13 (Damped harmonic oscillator). The Hamiltonian

$$H = \mathrm{e}^{-2\gamma t} \frac{p^2}{2m} + \mathrm{e}^{2\gamma t} \frac{m\omega^2}{2} q^2,$$

where γ is a positive constant, corresponds to a damped harmonic oscillator (see Example 2.6). With the aid of (5.15) one verifies that the coordinate transformation

$$Q = \mathrm{e}^{\gamma t} q, \qquad P = \mathrm{e}^{-\gamma t} p + m\gamma\, \mathrm{e}^{\gamma t} q$$

is canonical and that the new Hamiltonian can be chosen as $K = H + \gamma pq$, with $F_1 = -m\gamma\, \mathrm{e}^{2\gamma t} q^2/2$, thus

$$K = \mathrm{e}^{-2\gamma t} \frac{p^2}{2m} + \mathrm{e}^{2\gamma t} \frac{m\omega^2}{2} q^2 + \gamma pq = \frac{P^2}{2m} + \frac{m(\omega^2 - \gamma^2)}{2} Q^2,$$

which has the form of the standard Hamiltonian of a harmonic oscillator of frequency $\sqrt{\omega^2 - \gamma^2}$. Hence, if $\omega^2 - \gamma^2 > 0$, the solution of the Hamilton equations is given by $Q = A\cos(\sqrt{\omega^2 - \gamma^2}\, t + \delta)$, where A and δ are some constants and, therefore, $q = A\mathrm{e}^{-\gamma t} \cos(\sqrt{\omega^2 - \gamma^2}\, t + \delta)$.

(Note that the explicit expression of the function F_1 given above was not employed. In order to verify that the coordinate transformation is canonical and to find $K - H$ it is only necessary to show the *existence* of F_1. In this example the set (q, Q, t) is not independent and therefore F_1 cannot be a generating function of the transformation. Note also that the new Hamiltonian, K, being independent of t, is conserved. This constant of motion was already found in Example 2.20.)

Example 5.14 (Galilean transformations in one dimension). The Cartesian coordinate q, of a particle of mass m, measured in an inertial reference frame S, is related to the Cartesian coordinate Q of the same particle, measured in a frame S′ moving with respect to S with (constant) velocity v, by means of

$$Q = q - vt \tag{5.19}$$

if the origins coincide at $t = 0$. Then, the usual linear momenta of the particle, p and P, measured in S and S′, respectively, are related by

$$P = p - mv. \tag{5.20}$$

Equations (5.19) and (5.20) define a canonical transformation since $\{Q, P\} = 1$ and with the aid of Equations (5.12) we can find $K - H$ (up to an additive function of t only). Alternatively, we compute the differential form on the left-hand side of (5.15)

$$\begin{aligned} p\mathrm{d}q - H\mathrm{d}t - (P\mathrm{d}Q - K\mathrm{d}t) &= p\mathrm{d}q - (p - mv)(\mathrm{d}q - v\mathrm{d}t) + (K - H)\mathrm{d}t \\ &= vp\mathrm{d}t + mv\mathrm{d}q - mv^2\mathrm{d}t + (K - H)\mathrm{d}t \\ &= \mathrm{d}(mvq - \tfrac{1}{2}mv^2t) + (K - H + vp - \tfrac{1}{2}mv^2)\mathrm{d}t. \end{aligned} \tag{5.21}$$

By comparing this last expression with Equation (5.15) we conclude that the coordinate transformation given by (5.19) and (5.20) is canonical, and also that the new Hamiltonian must be given by

$$K = H - vp + \tfrac{1}{2}mv^2 + f(t), \tag{5.22}$$

where $f(t)$ is an arbitrary function of t only. If we choose $f = 0$, then by comparing (5.21) and (5.15) we see that (up to an additive constant) $F_1 = mvq - \frac{1}{2}mv^2t$.

It may be noticed that the last line of (5.21) can also be written as $\mathrm{d}(mvq - mv^2t) + (K - H + vp)\mathrm{d}t$, in which case the constant term $\frac{1}{2}mv^2$ would be absent from (5.22).

In this example (q, Q, t) is not functionally independent (since $Q = q - vt$) and therefore F_1 *cannot* be used as a generating function; however, (q, P, t) [as well as (p, Q, t)] can be used as coordinates in the extended phase space and the function $F_2(q, P, t) = F_1 + PQ = mvq + P(q - vt)$ generates the transformation (5.19)–(5.20), by means of (5.18). (Note that the function F_1 can be expressed in terms of q, Q, and t in infinitely many ways, some of them are $mvq - \frac{1}{2}mv^2t$, $mV(Q + vt) - \frac{1}{2}mv^2t$, and $\frac{1}{2}mv(q + Q + vt) - \frac{1}{2}mv^2t$.)

It may be remarked that it is not entirely a matter of choice which variables are used as coordinates in the extended phase space. It is the coordinate transformation what determines which combinations of old and new coordinates are functionally independent. Given a canonical transformation, the functions F_1, F_2, F_3, and F_4

always exist, at least locally, but they can be called generating functions *only if* the appropriate sets of variables are functionally independent. Note also that, in some cases, one of these functions can be equal to zero (see, e.g., Examples 5.10, 5.17, and 5.18).

The following example may help to understand what is happening here. The sets $\{\mathbf{i}, \mathbf{j}\}$ and $\{2\mathbf{i}, \mathbf{i}+\mathbf{j}\}$ are bases of $\mathbb{R}^2$; however, if we consider a set of two vectors formed by one vector from the first set and one vector from the second set, we may obtain a basis of $\mathbb{R}^2$ or not. In fact, $\{\mathbf{i}, \mathbf{i}+\mathbf{j}\}$, $\{\mathbf{j}, 2\mathbf{i}\}$, and $\{\mathbf{j}, \mathbf{i}+\mathbf{j}\}$ are bases of $\mathbb{R}^2$, but $\{\mathbf{i}, 2\mathbf{i}\}$ is not. In a similar way, in the case of a canonical transformation, we have two sets of coordinates of the phase space, (q, p) and (Q, P), but when we take one coordinate of the first set and one coordinate from the second set, not necessarily we will get a coordinate system of the phase space.

Exercise 5.15. Consider the canonical transformation found in Example 5.8. Is it possible to find generating functions of types F_2, F_3, or F_4 for this transformation?

Exercise 5.16. Is it possible to have a canonical transformation that can be obtained by means of a type F_1 generating function as well as by generating functions of types F_2, F_3, and F_4? (Note that, since the functions F_1, F_2, F_3, and F_4 always exist, at least locally, the question is equivalent to ask if is it possible to have a canonical transformation such that the sets (q, Q, t), (q, P, t), (p, Q, t), and (p, P, t) are all functionally independent.)

Example 5.17 (The Emden–Fowler equation). The Hamiltonian

$$H = \frac{p^2}{2t^2} + \frac{t^2 q^{k+1}}{k+1}$$

leads to the Emden–Fowler equation for q [see Equations (2.49) and (2.51)]. The coordinate transformation $Q = t^{1/2} q$, $P = t^{-1/2} p$ is canonical and the new Hamiltonian can be taken as $K = H + \frac{1}{2} t^{-1} pq$. In fact,

$$\begin{aligned} p\mathrm{d}q - H\mathrm{d}t - (P\mathrm{d}Q - K\mathrm{d}t) &= p\mathrm{d}q - t^{-1/2} p \left(t^{1/2}\mathrm{d}q + \tfrac{1}{2} t^{-1/2} q\,\mathrm{d}t\right) + (K-H)\mathrm{d}t \\ &= \left(K - H - \tfrac{1}{2} t^{-1} pq\right)\mathrm{d}t \\ &= 0 \end{aligned}$$

(so that Equation (5.15) holds, with $F_1 = 0$). Thus,

$$K = \frac{P^2}{2t} + \frac{Q^{k+1}}{(k+1)\, t^{(k-3)/2}} + \frac{PQ}{2t}.$$

In the special case where $k = 5$, the new Hamiltonian is given by

$$K = \frac{P^2}{2t} + \frac{Q^6}{6t} + \frac{PQ}{2t}, \tag{5.23}$$

which implies that the function

$$\Phi \equiv 3P^2 + Q^6 + 3PQ \tag{5.24}$$

is a constant of motion (see Exercise 4.8).

The constant of motion (5.24) can be employed to solve the Hamilton equations. Indeed, making use of one of the Hamilton equations and eliminating P with the aid of (5.24), we obtain

$$\frac{dQ}{dt} = \frac{\partial K}{\partial P} = \frac{P}{t} + \frac{Q}{2t} = \pm\frac{1}{2t}\sqrt{Q^2 - \frac{4}{3}Q^6 + \frac{4}{3}\Phi},$$

which is a separable equation, viz.,

$$\pm\frac{dt}{t} = \frac{2dQ}{\sqrt{Q^2 - \frac{4}{3}Q^6 + \frac{4}{3}\Phi}}.$$

In terms of the original coordinates, the constant of motion (5.24) is given by

$$\Phi = \frac{3p^2}{t} + t^3q^6 + 3pq.$$

Example 5.18 (The Poisson–Boltzmann equation). The Poisson–Boltzmann equation

$$\ddot{q} = -\frac{k}{t}\dot{q} - a e^q,$$

where k and a are constants such that $k \neq 0$ and $a \in \{-1, 1\}$ can be obtained from the Lagrangian $L = \frac{1}{2}t^k\dot{q}^2 - at^k e^q$ (see Exercise 2.9) and the corresponding Hamiltonian is given by

$$H = \frac{p^2}{2t^k} + at^k e^q.$$

The transformation $Q = t^2 e^q$, $P = t^{-2}e^{-q}p$ is canonical. In fact,

$$\begin{aligned} p dq - H dt - (P dQ - K dt) &= p dq - t^{-2}e^{-q}p\big(t^2 e^q dq + 2te^q dt\big) + (K - H)dt \\ &= (K - H - 2t^{-1}p)dt \end{aligned}$$

and, therefore, we can take $F_1 = 0$ and

$$K = H + \frac{2p}{t} = \frac{P^2Q^2}{2t^k} + \frac{aQ}{t^{2-k}} + \frac{2PQ}{t}.$$

Thus, in the case where $k = 1$,

$$K = \frac{P^2 Q^2}{2t} + \frac{aQ}{t} + \frac{2PQ}{t},$$

which implies that, in this particular case, tK (that is, $tH + 2p$) is a constant of motion (see Exercise 4.8). With the aid of this constant of motion we can find a separable first-order ODE for Q as a function of t.

Exercise 5.19. Consider the Hamiltonian

$$H = \frac{p^2}{2m} - \frac{k}{q^2},$$

where m and k are constants. Show that the coordinate transformation $Q = t^{-1/2}q$, $P = t^{1/2}p$ is canonical, and find the new Hamiltonian, K. By analyzing the form of K, find a constant of motion and use it, together with H (which is also conserved), to find the solution of the equations of motion without solving any differential equation.

Exercise 5.20. The possibility of finding by inspection a constant of motion, as in Examples 5.17 and 5.18, is present in some second-order ODEs invariant under scaling transformations. For instance, the nonlinear ODE

$$\ddot{q} = \frac{\dot{q}}{t} + \frac{q^2}{t^3}$$

is invariant under the substitution $q \mapsto \lambda q$, $t \mapsto \lambda t$, where λ is any nonzero real constant. Following the standard procedure one obtains the Lagrangian

$$L = \frac{\dot{q}^2}{2t} + \frac{q^3}{3t^4}$$

(see Section 2.3) and the Hamiltonian

$$H = \frac{tp^2}{2} - \frac{q^3}{3t^4}$$

for this equation. Show that for any value of the real parameter s, the transformation $Q = t^s q$, $P = t^{-s} p$ is canonical and find the new Hamiltonian. Determine a convenient value of s for which one can readily identify a constant of motion from K. (Thus, despite the fact that in a canonical transformation the coordinate t is not transformed, we can make use of symmetry transformations that also affect the time.) With the aid of this constant of motion, reduce the solution of the equations of motion to quadrature.

Example 5.21. A more elaborate example, involving the simple Hamiltonian

$$H = \frac{p^2}{2m} + mgq,$$

is given by the coordinate transformation

$$Q = t^{-1/2}q + \tfrac{1}{2}gt^{3/2}, \qquad P = t^{1/2}p + mgt^{3/2}.$$

Making use of these definitions we calculate the differential form

$$\begin{aligned} p\mathrm{d}q - H\mathrm{d}t - (P\mathrm{d}Q - K\mathrm{d}t) \\ &= p\mathrm{d}q - (t^{1/2}p + mgt^{3/2})\big(t^{-1/2}\mathrm{d}q - \tfrac{1}{2}t^{-3/2}q\mathrm{d}t + \tfrac{3}{4}gt^{1/2}\mathrm{d}t\big) + (K - H)\mathrm{d}t \\ &= \mathrm{d}\big(-mgtq\big) + \left(\frac{pq}{2t} - \frac{3}{4}gtp + \frac{3}{2}mgq - \frac{3}{4}mg^2t^2 + K - H\right)\mathrm{d}t, \end{aligned}$$

thus showing that the transformation is canonical and that the new Hamiltonian must be given by

$$K = H - \frac{pq}{2t} + \frac{3}{4}gtp - \frac{3}{2}mgq + \frac{3}{4}mg^2t^2 + f(t),$$

where $f(t)$ is a function of t only. Substituting the expression for H (which has not been employed so far) and expressing the result in terms of the new canonical coordinates one finds the simple expression

$$K = \frac{P^2}{2mt} - \frac{PQ}{2t} + \frac{1}{2}mg^2t^2 + f(t).$$

Hence, it is convenient to choose $f(t) = -\frac{1}{2}mg^2t^2$, so that

$$K = \frac{1}{t}\left(\frac{P^2}{2m} - \frac{PQ}{2}\right).$$

By contrast with the original Hamiltonian, which is a constant of motion, K depends explicitly on the time and is not conserved, but the product tK is a constant of motion (see Exercise 4.8). In terms of the original coordinates, the product tK is given by

$$\frac{tp^2}{2m} - \frac{1}{2}mgtq - \frac{1}{2}qp + \frac{3}{4}gt^2p + \frac{1}{4}mg^2t^3.$$

Constructing Canonical Transformations

We can construct a canonical transformation starting from a single expression of the form $P = P(q, p, t)$, assuming that this relation can be inverted to obtain p as a function of (q, P, t) (that is, $p = p(q, P, t)$). This condition implies that the desired canonical transformation can be defined by a type F_2 generating function. The first equation in (5.18) gives $p = (\partial F_2/\partial q)_{P,t}$ which, by virtue of the fundamental theorem of calculus, is satisfied if we take

$$F_2(q, P, t) = \int^q p(q', P, t)\, \mathrm{d}q', \tag{5.25}$$

where, in order to avoid confusion, we have used the symbol q' for the integration variable and q for the limit of the integral. (Of course, this equation defines F_2 up to an arbitrary additive function of P and t only and, therefore, the canonical transformation is not unique.) The missing part of the canonical transformation is then obtained from

$$Q = \left(\frac{\partial F_2}{\partial P}\right)_{q,t} = \int^q \frac{\partial p(q', P, t)}{\partial P}\, \mathrm{d}q'. \tag{5.26}$$

Example 5.22. If we take $P = m\omega q \sin\omega t + p\cos\omega t$, as in (5.14), from Equation (5.25) we get

$$\begin{aligned} F_2 &= \int^q \frac{1}{\cos\omega t}(P - m\omega q' \sin\omega t)\, \mathrm{d}q' \\ &= \frac{1}{\cos\omega t}(Pq - \tfrac{1}{2}m\omega q^2 \sin\omega t) + f(P, t), \end{aligned} \tag{5.27}$$

where $f(P, t)$ is an arbitrary function of two variables. Then, according to (5.26),

$$Q = \frac{q}{\cos\omega t} + \frac{\partial f(P, t)}{\partial P}. \tag{5.28}$$

(Note that with

$$f(P, t) = \frac{P^2}{2m}\left(t - \frac{\tan\omega t}{\omega}\right)$$

the first equation in (5.14) is recovered.)

It should be remarked that we can start with almost any function $P = P(q, p, t)$, which will be one of the new canonical coordinates; the only requisite is that p may be expressed as a function of q, P, and t.

Using a Constant of Motion to Find the Solution of the Hamilton Equations
In the examples presented above, the coordinate transformations were given without a motivation and in some cases we were able to see their usefulness only after finding the new Hamiltonian in terms of the new coordinates. As we shall see now, when we have a constant of motion, we can find a canonical transformation leading to a new set of coordinates, (Q, P), which are constants of motion.

We shall assume that $P(q, p, t)$ is a constant of motion such that p can be expressed as a function of (q, P, t), as above. Then

$$0 = \frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{\partial K}{\partial Q},$$

which means that the new Hamiltonian must be a function of P and t only. On the other hand, from (5.18) we have $K = H + \partial F_2/\partial t$ and, as we have seen, this last expression must be a function of P and t only, but we also know that the generating function F_2 given by (5.25) is determined up to an arbitrary additive function of P and t; by suitably choosing this additive function we can make $K = 0$ and then Q is also a constant of motion (see Example 5.23, below). This result is essentially the Liouville theorem, which will be presented in a more general setting in Section 6.2.

Example 5.23 (Harmonic oscillator). The function $P = m\omega q \sin \omega t + p \cos \omega t$, considered in the previous example, is a constant of motion if the Hamiltonian is given by

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2} q^2,$$

where m and ω are constants. From (5.18) and (5.27) we have

$$\begin{aligned} K &= H + \frac{\partial F_2}{\partial t} \\ &= H + \frac{\omega \sin \omega t}{\cos^2 \omega t}\left(Pq - \tfrac{1}{2} m\omega q^2 \sin \omega t\right) - \frac{m\omega^2}{2} q^2 + \frac{\partial f(P,t)}{\partial t} \\ &= \frac{p^2}{2m} + \frac{\omega P q \sin \omega t}{\cos^2 \omega t} - \frac{m\omega^2 q^2 \sin^2 \omega t}{2\cos^2 \omega t} + \frac{\partial f(P,t)}{\partial t} \\ &= \frac{P^2}{2m \cos^2 \omega t} + \frac{\partial f(P,t)}{\partial t} \end{aligned}$$

(which is, indeed, a function of P and t only). Choosing

$$f(P,t) = -\frac{P^2 \tan \omega t}{2m\omega},$$

we have $K = 0$, and, according to (5.28), $Q = q \sec \omega t - P \tan \omega t / m\omega$, which has to be a constant of motion. This last expression gives q in terms of the constants of motion P and Q.

It may be noticed that if we already have a constant of motion, the solution of the equations of motion can be obtained in an elementary manner by solving a first-order ODE (of course, this is true only when there is one degree of freedom). In the present case, from the Hamilton equations we have $p = m\dot{q}$ and, therefore, $P = m\omega q \sin \omega t + m(\mathrm{d}q/\mathrm{d}t) \cos \omega t$, which is a linear equation for q that can be readily solved. However, here we want to emphasize the applications of the canonical transformations.

Exercise 5.24. The function $p + mgt$ is a constant of motion if the Hamiltonian is given by

$$H = \frac{p^2}{2m} + mgq.$$

Making use of this constant of motion find a canonical transformation such that the new Hamiltonian, K, is equal to zero and with the aid of this transformation find the solution of the equations of motion.

Action-Angle Variables

The main limitation for the application of the procedure presented above is that we have to know a constant of motion in order to construct a canonical transformation such that $K = 0$. In the cases where there exist periodic solutions, a convenient constant of motion can be readily obtained.

We consider a system with a time-independent Hamiltonian such that, at least in some region of the phase space, the level curves $H(q, p) =$ const. are closed. This means that, in this region, the motion is periodic and that p cannot be given by a single-valued function of q and H (see Figure 5.1). We take a point (q_0, p_0) belonging to this region and we define a new coordinate, P, by

$$P(q_0, p_0) \equiv \frac{1}{2\pi} \oint_{C_0} p \, \mathrm{d}q, \tag{5.29}$$

where C_0 is the level curve of H passing through (q_0, p_0), traversed in the sense of the time evolution (see Figure 5.1). Then, the value of P is the same at all the points of C_0 and, therefore, C_0 is also a level curve of P, which implies that H is a function of P only (and that P is a function of H only). Making use of Green's theorem one can see that, apart from the factor $1/2\pi$, which is introduced for later convenience, the integral (5.29) is the area enclosed by the curve C_0 (taking $\mathrm{d}q \, \mathrm{d}p$ as the area element in the phase space).

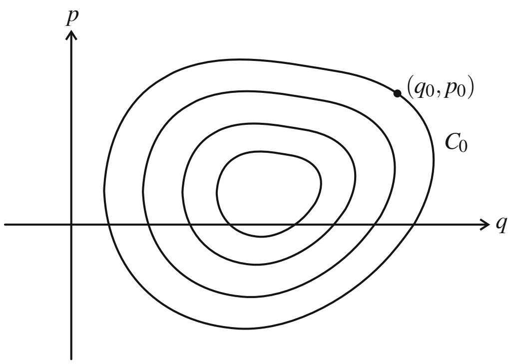


Fig. 5.1 If the level curves $H(q, p) =$ const. are closed, then the motion is periodic. Each point of the phase space, (q_0, p_0), belongs to one of these level curves, which we call C_0. The value of the so-called *action variable*, P, at (q_0, p_0) is the area enclosed by C_0 divided by 2π and, therefore, P has the same value at all the points of C_0

Since P is a function of q and p only [see (5.29)], we can choose F_2 in such a way that it is a function of q and P only [see (5.25)] and, in that way, $K = H$. Then, taking into account that K is a function of P only, the Hamilton equations give

$$\dot{Q} = \frac{\partial K}{\partial P}, \qquad \dot{P} = -\frac{\partial K}{\partial Q} = 0.$$

The derivative $\partial K/\partial P$ must be a function of P only and, defining

$$\omega(P) \equiv \frac{\partial K}{\partial P}, \tag{5.30}$$

the solution of the equations of motion in the coordinates (Q, P) is

$$P = P_0, \qquad Q = \omega(P_0)\, t + Q_0, \tag{5.31}$$

where P_0 and Q_0 are the values of P and Q at $t = 0$, respectively.

The coordinates P and Q defined in this way are called *action-angle variables*. Usually these variables are introduced starting from the Hamilton–Jacobi equation (which is presented in Chapter 6), and some special symbols are employed to denote them, such as I and J for the action variable (5.29), and θ, ϕ, and w, for the angle variable (5.26).

The curve C_0 (which is a closed curve in the qp-plane) corresponds to a straight line in the QP-plane (given by $P = P_0$, see Figure 5.2). Since the Jacobian of a canonical transformation is equal to 1 [see (5.11)], according to the formula for the change of variables in a multiple integral, the area of the region enclosed by the curve C_0 in the qp-plane (i.e., $2\pi P_0$) must be equal to the area of its image in the QP-plane under the canonical transformation $(q, p) \mapsto (Q, P)$, which is a rectangle bounded by the straight lines $P = P_0$, $P = 0$, $Q = Q_i$,and $Q = Q_f$,

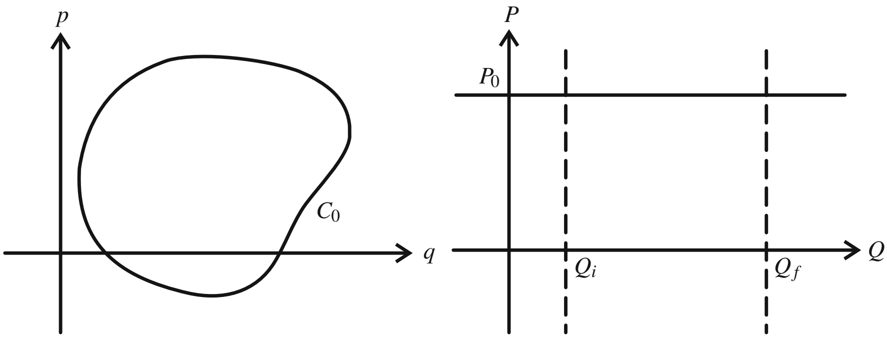


Fig. 5.2 Owing to the definition of the action variable, P, the closed curve C_0 in the qp-plane corresponds to a straight line, $P = P_0$, in the QP-plane. Since the Jacobian of a canonical transformation is equal to 1, the area under the line $P = P_0$, between the vertical lines $Q = Q_i$ and $Q = Q_f$, where Q_i and Q_f are the initial and final values of Q in the closed curve C_0, respectively, must be equal to the area enclosed by C_0, that is, $2\pi P_0$, which means that Q must increase by 2π in each cycle

where Q_i and Q_f are, respectively, the initial and final values of Q in a cycle. Hence, when the state of the system returns to the original point, the value of Q has to be increased by 2π (that is, $Q_f - Q_i = 2\pi$). This means that $\omega(P_0)$ is the angular frequency of the motion [see Equation (5.31)]. Note that the Hamiltonian need not have the simple form $p^2/2m + V$.

Example 5.25 (Action-angle variables for the one-dimensional harmonic oscillator). One of the simplest examples, which is commonly considered in the textbooks, is that of a one-dimensional harmonic oscillator. If

$$H = \frac{p^2}{2m} + \frac{m{\omega_0}^2}{2} q^2,$$

where m and ω_0 are constants, the level curves $H(q, p) = \text{const.}$ in the phase space are ellipses centered at the origin. The level curve of H passing through (q_0, p_0) is an ellipse with semiaxes

$$\sqrt{\frac{2H(q_0, p_0)}{m{\omega_0}^2}} \qquad \text{and} \qquad \sqrt{2mH(q_0, p_0)}$$

and, therefore, the area of this ellipse is equal to $2\pi H(q_0, p_0)/\omega_0$ (recalling that the area enclosed by an ellipse with semiaxes a and b is πab). From Equation (5.29) we obtain $P = H/\omega_0$ (so that, in effect, P is a function of H only), and Equation (5.30) gives $\omega(P) = \omega_0$ (a constant, as expected).

We can also obtain the solution of the equations of motion in terms of the original coordinates with the aid of (5.26). We have

$$Q = \int^q \frac{\partial}{\partial P} \sqrt{2m\omega_0 P - m^2{\omega_0}^2 q'^2}\, \mathrm{d}q' = \int^q \frac{m\omega_0 \mathrm{d}q'}{\sqrt{2m\omega_0 P - m^2{\omega_0}^2 q'^2}}.$$

Making use of the change of variable $m\omega_0 q' = \sqrt{2m\omega_0 P} \cos\alpha$ we find

$$Q = -\arccos\left(\sqrt{\frac{m\omega_0}{2P}}\, q\right) + \text{const.},$$

thus, using the solution of the equations of motion in terms of the action-angle variables, (5.31),

$$q = \sqrt{\frac{2P}{m\omega_0}} \cos(\omega_0 t + Q_0).$$

Example 5.26. Another commonly encountered example is given by the Hamiltonian

$$H = \frac{p^2}{2m} + V_0 \tan^2\left(\frac{\pi q}{2a}\right), \qquad -a < q < a,$$

where V_0 and a are positive constants. As in the case of the harmonic oscillator, all the level curves of this function $H(q, p)$ are closed. According to (5.29) we have

$$P(q_0, p_0) = \frac{2}{\pi}\int_0^{q_m} \sqrt{2m\big[H(q_0, p_0) - V_0 \tan^2(\pi q'/2a)\big]}\, \mathrm{d}q',$$

where $q_m \in [0, a)$ is such that

$$H(q_0, p_0) - V_0 \tan^2\left(\frac{\pi q_m}{2a}\right) = 0.$$

One can verify that

$$\int\sqrt{a - b\tan^2 x}\, \mathrm{d}x = \sqrt{a+b}\ \arctan\left(\frac{\sqrt{a+b}\ \tan x}{\sqrt{a-b\tan^2 x}}\right) - \sqrt{b}\ \arctan\left(\frac{\sqrt{b}\ \tan x}{\sqrt{a-b\tan^2 x}}\right),$$

hence,

$$P = \frac{2a}{\pi}\left[\sqrt{2m(H + V_0)} - \sqrt{2mV_0}\,\right]$$

and, therefore, inverting this last relation,

$$H = \frac{\pi^2 P^2}{8ma^2} + \frac{\pi P\sqrt{2mV_0}}{2ma},$$

which leads to the angular frequency

$$\omega = \frac{\pi^2 P}{4ma^2} + \frac{\pi\sqrt{2mV_0}}{2ma} = \frac{\pi}{a}\sqrt{\frac{H + V_0}{2m}}.$$

(The higher the value of H, the smallest the period of the motion.)

In order to find the solution of the equations of motion in terms of the original canonical coordinates, it is convenient to write Equation (5.26) in the form

$$Q = \int^q \frac{\partial p(q', H)}{\partial H}\frac{\mathrm{d}H}{\mathrm{d}P}\, \mathrm{d}q' = \int^q \frac{m\omega\, \mathrm{d}q'}{\sqrt{2mH - 2mV_0\tan^2(\pi q'/2a)}}.$$

Then, with the aid of the formula

$$\int \frac{\mathrm{d}x}{\sqrt{a - b\tan^2 x}} = \frac{1}{\sqrt{a+b}} \arcsin\left(\sqrt{\frac{a+b}{a}}\sin x\right), \tag{5.32}$$

one obtains

$$Q = \arcsin\left[\sqrt{\frac{H+V_0}{H}}\sin\left(\frac{\pi q}{2a}\right)\right] + \text{const.},$$

so that

$$q = \frac{2a}{\pi}\arcsin\left[\sqrt{\frac{H}{H+V_0}}\sin(\omega t + Q_0)\right].$$

Further examples can be found, e.g., in Percival and Richards [12, Chap. 7].

It may be pointed out that if we have a system with a time-independent Hamiltonian, then the fact that H is conserved allows us to reduce the equations of motion to a single first-order ODE, without having to perform a coordinate transformation. For instance, in the case of the Hamiltonian considered in Example 5.26, denoting by E the constant value of H, we have $E = \frac{1}{2}m\dot{q}^2 + V_0\tan^2(\pi q/2a)$, which leads to the first-order ODE

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \pm\sqrt{\frac{2}{m}\left[E - V_0\tan^2\left(\frac{\pi q}{2a}\right)\right]}$$

then, making use of (5.32), one obtains

$$q = \pm\frac{2a}{\pi}\arcsin\left\{\sqrt{\frac{E}{E+V_0}}\sin\left[\frac{\pi}{a}\sqrt{\frac{E+V_0}{2m}}\,(t-t_0)\right]\right\}$$

which, among other things, gives us the angular frequency of the motion.

However, the action-angle variables are interesting for two reasons at least. One of them is that the action variables are "adiabatic invariants" (see, e.g., Fasano and Marmi [8, Sect. 12.7] and Wells and Siklos [21]). Another reason is that the action-angle variables are useful in the study of perturbations (see, e.g., Fasano and Marmi [8, Sect. 12.7]).

Exercise 5.27. Consider a Hamiltonian of the form

$$H = \frac{p^2}{2m} + V(q),$$

where $V(q)$ is a function with a single minimum in a certain domain such that, for each value E of H, the motion is limited to an interval $q_-(E) \leqslant q \leqslant q_+(E)$ (see Figure 5.3). By adding a suitable constant term to $V(q)$ we can assume that the minimum value of $V(q)$ is zero (note that the Hamiltonians considered in Examples 5.25 and 5.26 satisfy these conditions). Show that the action variable, P, can be expressed as

$$P(E_0) = \frac{1}{\pi}\sqrt{\frac{m}{2}} \int_0^{E_0} \frac{[q_+(E) - q_-(E)]}{\sqrt{E_0 - E}} \mathrm{d}E, \tag{5.33}$$

where $q_+(E)$ and $q_-(E)$ are the turning points corresponding to the energy E, that is, the solutions of the equation $V(q) = E$, with $q_+(E) \geqslant q_-(E)$ (see Figure 5.3).

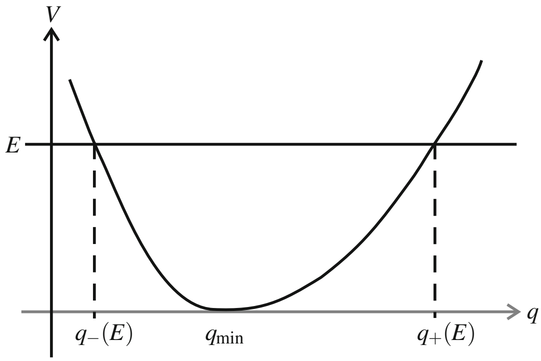


Fig. 5.3 Graph of the function $V(q)$. We assume that for each real number, $E > 0$, there exist two values of q, denoted as $q_+(E)$ and $q_-(E)$, where the value of $V(q)$ coincides with E. The potential has a minimum value equal to zero at some point $q_{\min}$. The function $w(E) \equiv q_+(E) - q_-(E)$ measures the width of the graph of $V(q)$ at the height E

The result of this exercise shows that, in the cases satisfying the imposed conditions, the dependence of the action variable (and that of the angular frequency) on the Hamiltonian is determined by the function $w(E) \equiv q_+(E) - q_-(E)$ (see Figure 5.3). The function $w(E)$ does not define the potential V uniquely, in fact, there are an infinite number of potentials with the same function $w(E)$. For instance, in the case of the harmonic oscillator (considered in Example 5.25), $w(E) = 2\sqrt{2E/m\omega_0{}^2}$, and this function leads to a constant angular frequency. Another well-known potential that also leads to a constant angular frequency is given by

$$V(q) = a^2q^2 + \frac{b^2}{q^2} - 2ab = \left(aq - \frac{b}{q}\right)^2,$$

where a, b are positive constants. One obtains $w(E) = \sqrt{E}/a$, which is also proportional to the square root of E and therefore also produces a constant angular frequency.

Equation (5.33) amounts to say that $P(E)$ is, up to a constant factor, the Abel transform of $w(E)$. This relation can be inverted and one can show that

$$w(E) = \sqrt{\frac{2}{m}} \int_0^E \frac{\mathrm{d}P(E_0)}{\mathrm{d}E_0} \frac{\mathrm{d}E_0}{\sqrt{E - E_0}},$$

which means that the dependence of the action variable on H only depends on the function $w(E)$ and conversely.

5.2 Systems with an Arbitrary Number of Degrees of Freedom

When the number of degrees of freedom is greater than 1, the canonical transformations are defined in a similar manner to that given in the case where there is one degree of freedom, and the existence of a generating function for any canonical transformation can be demonstrated following essentially the same steps as in the preceding section. We start assuming that the set of Hamilton equations

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q_i} \tag{5.34}$$

$(i = 1, 2, \ldots, n)$, is equivalent to the set

$$\frac{\mathrm{d}Q_i}{\mathrm{d}t} = \frac{\partial K}{\partial P_i}, \qquad \frac{\mathrm{d}P_i}{\mathrm{d}t} = -\frac{\partial K}{\partial Q_i}, \tag{5.35}$$

where the new coordinates Q_i and P_i can be expressed as functions of q_i, p_i, and possibly also of the time, and K is some function. Then, by virtue of the chain rule and Equations (5.34) and (5.35) we find that the partial derivatives of K with respect to the original coordinates must be given by

$$\begin{aligned}
\frac{\partial K}{\partial q_i} &= \frac{\partial K}{\partial Q_j}\frac{\partial Q_j}{\partial q_i} + \frac{\partial K}{\partial P_j}\frac{\partial P_j}{\partial q_i} \\
&= -\frac{\mathrm{d}P_j}{\mathrm{d}t}\frac{\partial Q_j}{\partial q_i} + \frac{\mathrm{d}Q_j}{\mathrm{d}t}\frac{\partial P_j}{\partial q_i} \\
&= -\left(\frac{\partial P_j}{\partial t} + \frac{\partial P_j}{\partial q_k}\frac{\partial H}{\partial p_k} - \frac{\partial P_j}{\partial p_k}\frac{\partial H}{\partial q_k}\right)\frac{\partial Q_j}{\partial q_i}
\end{aligned}$$

$$+ \left(\frac{\partial Q_j}{\partial t} + \frac{\partial Q_j}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial Q_j}{\partial p_k} \frac{\partial H}{\partial q_k} \right) \frac{\partial P_j}{\partial q_i}$$
$$= \left(\frac{\partial Q_j}{\partial q_i} \frac{\partial P_j}{\partial p_k} - \frac{\partial P_j}{\partial q_i} \frac{\partial Q_j}{\partial p_k} \right) \frac{\partial H}{\partial q_k} - \left(\frac{\partial Q_j}{\partial q_i} \frac{\partial P_j}{\partial q_k} - \frac{\partial P_j}{\partial q_i} \frac{\partial Q_j}{\partial q_k} \right) \frac{\partial H}{\partial p_k}$$
$$+ \frac{\partial Q_j}{\partial t} \frac{\partial P_j}{\partial q_i} - \frac{\partial P_j}{\partial t} \frac{\partial Q_j}{\partial q_i}. \tag{5.36}$$

In order to write the last expression in a compact form, it is convenient to introduce the *Lagrange bracket*

$$[u, v] \equiv \frac{\partial Q_j}{\partial u} \frac{\partial P_j}{\partial v} - \frac{\partial P_j}{\partial u} \frac{\partial Q_j}{\partial v}, \tag{5.37}$$

where u and v are any two variables belonging to the set q_i, p_i, t. It may be noticed that, in the special case where $n = 1$, the Lagrange bracket $[q, p]$, defined above, is just the Poisson bracket $\{Q, P\}$. (Sometimes, in order to specify which variables are being differentiated, the expression (5.37) is denoted as $[u, v]_{Q,P}$.) Note also that the Lagrange bracket is antisymmetric, $[u, v] = -[v, u]$.

Then, Equation (5.36) can be written as

$$\frac{\partial K}{\partial q_i} = [q_i, p_k] \frac{\partial H}{\partial q_k} - [q_i, q_k] \frac{\partial H}{\partial p_k} + [t, q_i]. \tag{5.38}$$

In a similar way one finds that

$$\frac{\partial K}{\partial p_i} = [p_i, p_k] \frac{\partial P_j}{\partial p_k} - [p_i, q_k] \frac{\partial H}{\partial q_k} + [t, p_i] \tag{5.39}$$

[cf. Equations (5.3) and (5.5)].

Exercise 5.28. Show that

$$\frac{\partial}{\partial u}[v, w]_{Q,P} + \frac{\partial}{\partial v}[w, u]_{Q,P} + \frac{\partial}{\partial w}[u, v]_{Q,P} = 0, \tag{5.40}$$

where u, v, w are any three variables belonging to the set q_i, p_i, t.

The equality of the $(2n)(2n-1)/2$ mixed partial derivatives of K obtainable from (5.38) and (5.39) gives necessary and sufficient conditions for the (local) existence of K. (These conditions are analyzed in Section 5.5, at the end of this chapter.) As in the case of systems with one degree of freedom, we will be mainly interested in canonical transformations, which are *defined* here by the conditions

$$[q_i, q_k] = 0, \qquad [q_i, p_k] = \delta_{ik}, \qquad [p_i, p_k] = 0 \tag{5.41}$$

[cf. Equation (5.8)]. Note that these equations do not contain the function H or K, which means that in order to verify that a given coordinate transformation is canonical it is not necessary to specify the original or the new Hamiltonian. Then, Equations (5.38) and (5.39) reduce to

$$\frac{\partial(K-H)}{\partial q_i} = [t, q_i], \qquad \frac{\partial(K-H)}{\partial p_i} = [t, p_i] \tag{5.42}$$

[cf. Equations (5.12)]. (In Section 5.5 it is shown that Equations (5.41) are necessary conditions for the existence of a new Hamiltonian, K, for *any* Hamiltonian, H.)

As we shall show now, for any canonical transformation, there exists (at least locally) a function K that satisfies Equations (5.42). For this purpose it is convenient to note that the $2n$ equations in (5.42) can be expressed in the common form

$$\frac{\partial(K-H)}{\partial x_\alpha} = [t, x_\alpha]$$

($\alpha = 1, 2, \ldots, 2n$), where

$$(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{2n}) \equiv (q_1, \ldots, q_n, p_1, \ldots, p_n). \tag{5.43}$$

Then, making use of (5.40),

$$\frac{\partial^2(K-H)}{\partial x_\alpha \partial x_\beta} - \frac{\partial^2(K-H)}{\partial x_\beta \partial x_\alpha} = \frac{\partial}{\partial x_\alpha}[t, x_\beta] - \frac{\partial}{\partial x_\beta}[t, x_\alpha] = \frac{\partial}{\partial t}[x_\alpha, x_\beta],$$

which is equal to zero by virtue of (5.41). Thus, for a given canonical transformation, Equations (5.42) determine $K - H$ up to an additive function of t only, which can be arbitrarily chosen because its presence does not affect the right-hand sides of (5.35).

In the case of a canonical transformation that does not involve the time, that is, $Q_i = Q_i(q_j, p_j)$, $P_i = P_i(q_j, p_j)$, according to (5.42), K can be taken equal to H.

Example 5.29 (Uniformly rotating frame in the three-dimensional Euclidean space). The coordinate transformation

$$Q_1 = q_1 \cos\omega_r t + q_2 \sin\omega_r t, \quad Q_2 = -q_1 \sin\omega_r t + q_2 \cos\omega_r t, \quad Q_3 = q_3,$$
$$P_1 = p_1 \cos\omega_r t + p_2 \sin\omega_r t, \quad P_2 = -p_1 \sin\omega_r t + p_2 \cos\omega_r t, \quad P_3 = p_3, \tag{5.44}$$

where ω_r is a constant, gives the relation between the Cartesian coordinates and momenta of a particle viewed from two reference frames; the frame with coordinates (Q_1, Q_2, Q_3) rotates with respect to the reference frame with coordinates (q_1, q_2, q_3) in the positive sense about the third axis with angular velocity ω_r.

Note that, as in all the previous examples, the transformation (5.44) is *passive*, that is, Equations (5.44) give the relation between the coordinates of a point with respect to *two different frames*. This is the standard point of view when we talk about coordinate transformations. For instance, in the formulas $x = r\cos\theta$, $y = r\sin\theta$, relating the polar and the Cartesian coordinates of the Euclidean plane, the variables x, y, r and θ correspond to a single point of the plane. The Galilean and the Lorentz transformations are usually viewed in this way. Equations (5.44) can also be interpreted as defining an *active* transformation; in that case, the q_i are the coordinates of a point before the rotation and the Q_i are the coordinates of the rotated point, all referred to a single frame.

We can readily see that the coordinate transformation (5.44) is canonical. Since $\partial P_j/\partial q_k = 0 = \partial Q_j/\partial p_k$, the first and third set of equations of (5.41) are automatically satisfied, and by means of a straightforward computation one can verify that the second set of equations of (5.41) also hold. For instance, we have

$$\begin{aligned}\frac{\partial Q_j}{\partial q_1}\frac{\partial P_j}{\partial p_1} - \frac{\partial P_j}{\partial q_1}\frac{\partial Q_j}{\partial p_1} &= \frac{\partial Q_j}{\partial q_1}\frac{\partial P_j}{\partial p_1} \\ &= \cos\omega_r t\cos\omega_r t + (-\sin\omega_r t)(-\sin\omega_r t) \\ &= 1.\end{aligned}$$

On the other hand, from Equations (5.42), we have, for instance,

$$\frac{\partial (K-H)}{\partial q_1} = \frac{\partial Q_j}{\partial t}\frac{\partial P_j}{\partial q_1} - \frac{\partial P_j}{\partial t}\frac{\partial Q_j}{\partial q_1} = -\frac{\partial P_j}{\partial t}\frac{\partial Q_j}{\partial q_1} = -\omega_r p_2.$$

Continuing in this manner one finds that

$$K - H = -\omega_r(q_1 p_2 - q_2 p_1) + f(t), \tag{5.45}$$

where $f(t)$ is a function of t only.

Equations (5.41) and (5.42) are necessary and sufficient conditions for the local existence of a function F_1 defined on the extended phase space such that

$$p_i \mathrm{d}q_i - H\mathrm{d}t - (P_i\mathrm{d}Q_i - K\mathrm{d}t) = \mathrm{d}F_1, \tag{5.46}$$

as can be readily verified writing the left-hand side of the last equation in terms of the original variables

$$\left(p_i - P_j\frac{\partial Q_j}{\partial q_i}\right)\mathrm{d}q_i - P_j\frac{\partial Q_j}{\partial p_i}\mathrm{d}p_i + \left(K - H - P_j\frac{\partial Q_j}{\partial t}\right)\mathrm{d}t$$

and applying again the standard criterion for the local exactness of a linear differential form. Note that the function F_1 can be equal to zero (see, e.g., Example 5.36).

Example 5.30 (Charged particle in a uniform magnetic field). Consider the (linear) coordinate transformation

$$\begin{aligned}Q_1 &= \frac{1}{2}q_1 - \frac{1}{\mu}p_2, \ Q_2 = \frac{1}{2}q_1 + \frac{1}{\mu}p_2, \ Q_3 = q_3, \\ P_1 &= p_1 + \frac{\mu}{2}q_2, \quad P_2 = p_1 - \frac{\mu}{2}q_2, \quad P_3 = p_3,\end{aligned} \tag{5.47}$$

where μ is a constant different from zero. It can be readily verified that this transformation is canonical by computing the Lagrange brackets (5.41) (15 in total); however, it is more convenient to construct the differential form of the left-hand side of (5.46). Since the transformation (5.47) does not involve the time, we can take $K = H$, and, therefore

$$\begin{aligned}p_i \mathrm{d}q_i - H\mathrm{d}t - (P_i \mathrm{d}Q_i - K\mathrm{d}t) &= p_1\mathrm{d}q_1 + p_2\mathrm{d}q_2 - \left(p_1 + \frac{\mu}{2}q_2\right)\left(\frac{1}{2}\mathrm{d}q_1 - \frac{1}{\mu}\mathrm{d}p_2\right) \\ &\quad - \left(p_1 - \frac{\mu}{2}q_2\right)\left(\frac{1}{2}\mathrm{d}q_1 + \frac{1}{\mu}\mathrm{d}p_2\right) \\ &= \mathrm{d}(p_2 q_2),\end{aligned}$$

which shows that the transformation (5.47) is canonical.

Note that a coordinate transformation, such as (5.47), is canonical regardless of the meaning of the coordinates and of the range of values they can take.

The canonical transformation (5.47) is useful in connection with the Hamiltonian

$$H = \frac{1}{2m}\left[\left(p_x + \frac{eB_0}{2c}y\right)^2 + \left(p_y - \frac{eB_0}{2c}x\right)^2 + {p_z}^2\right], \tag{5.48}$$

which corresponds to a charged particle of mass m and electric charge e in a uniform magnetic field $\mathbf{B} = B_0\mathbf{k}$ (in cgs, or Gaussian, units), where c is the speed of light in vacuum, if the vector potential is chosen as $\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}$ [see (4.14)]. If we set $\mu = eB_0/c$, identifying the Cartesian coordinates x, y, z, p_x, p_y, p_z with $q_1, q_2, q_3, p_1, p_2, p_3$, respectively, we find that

$$K = \frac{1}{2m}\left[{P_1}^2 + \left(\frac{eB_0}{c}\right)^2 {Q_1}^2 + {P_3}^2\right], \tag{5.49}$$

which is the sum of the standard Hamiltonian of a one-dimensional harmonic oscillator of angular frequency eB_0/mc, and that of a free particle in one dimension. Among other things, we can take advantage of the fact that P_2, Q_2, and Q_3 do not appear in the expression for K and, correspondingly, Q_2, P_2, and P_3, are constants of motion, which, in terms of the original coordinates, are

$$Q_2 = \frac{c}{eB_0}\left(p_y + \frac{eB_0}{2c}x\right), \qquad P_2 = p_x - \frac{eB_0}{2c}y, \qquad P_3 = p_z. \tag{5.50}$$

The solution of the Hamilton equations corresponding to the Hamiltonian (5.49) is readily found to be

$$Q_1 = (Q_1)_0 \cos\omega_c t + \frac{c}{eB_0}(P_1)_0 \sin\omega_c t,\ Q_2 = (Q_2)_0,\ Q_3 = \frac{1}{m}(P_3)_0\, t + (Q_3)_0,$$
$$P_1 = (P_1)_0 \cos\omega_c t - \frac{eB_0}{c}(Q_1)_0 \sin\omega_c t,\ \ P_2 = (P_2)_0,\ \ P_3 = (P_3)_0,$$

where the $(Q_i)_0$ and $(P_i)_0$ are constants that denote the values of Q_i and P_i at $t = 0$, respectively, and $\omega_c \equiv eB_0/mc$ (the cyclotron frequency). Thus, from (5.47) (with $\mu = eB_0/c$) we obtain, e.g.,

$$\begin{aligned} x &= (Q_2)_0 + (Q_1)_0 \cos\omega_c t + \frac{c}{eB_0}(P_1)_0 \sin\omega_c t, \\ y &= -\frac{c}{eB_0}(P_2)_0 - (Q_1)_0 \sin\omega_c t + \frac{c}{eB_0}(P_1)_0 \cos\omega_c t, \\ z &= (Q_3)_0 + \frac{1}{m}(P_3)_0\, t, \end{aligned} \tag{5.51}$$

which corresponds to the well-known result that the orbit in the configuration space is a circle (if $(P_3)_0 = 0$) or a helix (if $(P_3)_0 \neq 0$). The constants of motion Q_2 and P_2 are related to the coordinates of the center of the circle described on the xy-plane (cf. Example 1.19).

The momenta p_x, p_y, p_z, as functions of the time, can be obtained from (5.47), in the same manner as we obtained Equations (5.51), or by differentiating Equations (5.51) with respect to the time and comparing the result with the Hamilton equations, e.g., $\dot{x} = \partial H/\partial p_x = \frac{1}{m}(p_x + eB_0 y/2c)$.

Example 5.31 (Charged particle in a uniform magnetic field). Another useful canonical transformation for the problem of a charged particle in a uniform magnetic field is the one given in Example 5.29. Indeed, from Equations (5.45) and (5.48) we find that the new Hamiltonian is given by

$$\begin{aligned} K = \frac{1}{2m}\left[{p_x}^2 + {p_y}^2 + {p_z}^2 - \frac{eB_0}{c}(xp_y - yp_x) + \left(\frac{eB_0}{2c}\right)^2 (x^2 + y^2)\right] \\ - \omega_r(xp_y - yp_x) + f(t), \end{aligned}$$

hence, choosing $\omega_r = -eB_0/2mc$ and $f(t) = 0$, we obtain

$$\begin{aligned} K &= \frac{{p_x}^2 + {p_y}^2 + {p_z}^2}{2m} + \frac{m}{2}\left(\frac{eB_0}{2mc}\right)^2 (x^2 + y^2) \\ &= \frac{{P_1}^2 + {P_2}^2 + {P_3}^2}{2m} + \frac{m}{2}\left(\frac{eB_0}{2mc}\right)^2 ({Q_1}^2 + {Q_2}^2), \end{aligned}$$

which is the sum of the standard Hamiltonian for a two-dimensional isotropic harmonic oscillator (of frequency $eB_0/2mc$) and that of a free particle in one dimension [cf. (5.49)].

Making use of the solution of the equations of motion of a two-dimensional isotropic harmonic oscillator, namely

$$\begin{aligned}
Q_1 &= (Q_1)_0 \cos\frac{\omega_c t}{2} + \frac{2}{m\omega_c}(P_1)_0 \sin\frac{\omega_c t}{2}, \\
P_1 &= (P_1)_0 \cos\frac{\omega_c t}{2} - \frac{m\omega_c}{2}(Q_1)_0 \sin\frac{\omega_c t}{2}, \\
Q_2 &= (Q_2)_0 \cos\frac{\omega_c t}{2} + \frac{2}{m\omega_c}(P_2)_0 \sin\frac{\omega_c t}{2}, \\
P_2 &= (P_2)_0 \cos\frac{\omega_c t}{2} - \frac{m\omega_c}{2}(Q_2)_0 \sin\frac{\omega_c t}{2},
\end{aligned} \tag{5.52}$$

where the constants $(Q_1)_0, (Q_2)_0, (P_1)_0, (P_2)_0$ are the values of Q_1, Q_2, P_1, P_2, respectively, at $t = 0$, and $\omega_c \equiv eB_0/mc$ (see Section 2.1), from Equations (5.44), with $\omega_r = -\omega_c/2$ (the *Larmor frequency*), we obtain

$$x = \frac{(Q_1)_0}{2} + \frac{(P_2)_0}{m\omega_c} + \left[\frac{(Q_1)_0}{2} - \frac{(P_2)_0}{m\omega_c}\right]\cos\omega_c t + \left[\frac{(Q_2)_0}{2} + \frac{(P_1)_0}{m\omega_c}\right]\sin\omega_c t,$$

$$y = \frac{(Q_2)_0}{2} - \frac{(P_1)_0}{m\omega_c} - \left[\frac{(Q_1)_0}{2} - \frac{(P_2)_0}{m\omega_c}\right]\sin\omega_c t + \left[\frac{(Q_2)_0}{2} + \frac{(P_1)_0}{m\omega_c}\right]\cos\omega_c t,$$

corresponding to a circular motion with constant angular velocity ω_c. It is interesting to note that with respect to the rotating axes, (Q_1, Q_2), the particle describes an ellipse centered at the origin [see (5.52)], while in the xy-plane the particle describes a circle whose center may be any point of this plane.

Generating Functions

If (q_i, Q_i, t) is functionally independent, then the function F_1, appearing on the right-hand side of Equation (5.46), can be expressed as a function of q_i, Q_i, and t, in a unique way, and Equation (5.46) implies that

$$p_i = \left(\frac{\partial F_1}{\partial q_i}\right)_{q_j, Q_j, t}, \qquad P_i = -\left(\frac{\partial F_1}{\partial Q_i}\right)_{q_j, Q_j, t}, \qquad K - H = \left(\frac{\partial F_1}{\partial t}\right)_{q_j, Q_j}. \tag{5.53}$$

The independence of the $2n$ variables q_i, p_i requires that $\det(\partial^2 F_1/\partial q_i \partial Q_j) \neq 0$. Conversely, given a function $F_1(q_i, Q_i, t)$ satisfying this condition, according to the implicit function theorem, Equations (5.53) define a local canonical transformation.

The coordinates q_i, Q_i are not always functionally independent; for instance, in Example 5.30 we have the relation $q_1 = Q_1 + Q_2$ [see Equations (5.47)] and therefore, in this case, (q_i, Q_i, t) *cannot* be used as coordinates in the extended phase space [see also Equations (5.44)].

For the canonical transformations such that the set (q_i, Q_i, t) is functionally dependent, one can employ generating functions that depend on other combinations of n old and n new variables, and t; some or all of the q_i can be replaced by their conjugate momenta p_i and, similarly, some or all of the Q_i can be replaced by their conjugate momenta P_i, giving a total of 2^{2n} possibilities (see, e.g., Exercise 5.32, below). For instance, expressing (5.46) in the form

$$p_i \mathrm{d}q_i - H\mathrm{d}t + Q_i \mathrm{d}P_i + K\mathrm{d}t = \mathrm{d}F_2, \tag{5.54}$$

where $F_2 \equiv F_1 + Q_i P_i$, it follows that if the set (q_i, P_i, t) is functionally independent, then

$$p_i = \left(\frac{\partial F_2}{\partial q_i}\right)_{q_j, P_j, t}, \qquad Q_i = \left(\frac{\partial F_2}{\partial P_i}\right)_{q_j, P_j, t}, \qquad K - H = \left(\frac{\partial F_2}{\partial t}\right)_{q_j, P_j}. \tag{5.55}$$

Similarly, if (p_i, Q_i, t) is functionally independent, writing (5.46) in the equivalent form

$$-q_i \mathrm{d}p_i - H\mathrm{d}t - (P_i \mathrm{d}Q_i - K\mathrm{d}t) = \mathrm{d}F_3,$$

where $F_3 \equiv F_1 - p_i q_i$, we see that the canonical transformation is determined by

$$q_i = -\left(\frac{\partial F_3}{\partial p_i}\right)_{p_j, Q_j, t}, \quad P_i = -\left(\frac{\partial F_3}{\partial Q_i}\right)_{p_j, Q_j, t}, \quad K - H = \left(\frac{\partial F_3}{\partial t}\right)_{p_j, Q_j}. \tag{5.56}$$

When (p_i, P_i, t) is functionally independent, we can make use of the generating function $F_4 \equiv F_1 - p_i q_i + P_i Q_i$. Then

$$-q_i \mathrm{d}p_i - H\mathrm{d}t + Q_i \mathrm{d}P_i + K\mathrm{d}t = \mathrm{d}F_4,$$

which implies that

$$q_i = -\left(\frac{\partial F_4}{\partial p_i}\right)_{p_j, P_j, t}, \qquad Q_i = \left(\frac{\partial F_4}{\partial P_i}\right)_{p_j, P_j, t}, \qquad K - H = \left(\frac{\partial F_4}{\partial t}\right)_{p_j, P_j}. \tag{5.57}$$

Exercise 5.32. Show that, in the case of the canonical transformation given by (5.47), the set $(q_1, q_2, q_3, P_1, Q_2, P_3)$ is functionally independent and find a generating function $F(q_1, q_2, q_3, P_1, Q_2, P_3)$ for this transformation. Is it possible to find generating functions of types F_1, F_2, F_3, or F_4 for this canonical transformation?

Example 5.33. The function

$$F_2 = a(t)\, q_i P_i, \tag{5.58}$$

where $a(t)$ is an arbitrary nonvanishing function of t only, satisfies the condition $\det(\partial^2 F_2/\partial q_i \partial P_j) \neq 0$ and according to (5.55) generates the canonical transformation given by $p_i = a(t)\, P_i$, $Q_i = a(t)\, q_i$, that is

$$Q_i = a(t)\, q_i, \qquad P_i = \frac{1}{a(t)} p_i$$

(a possibly time-dependent change of scale) and $K = H + a'(t)\, Q_i P_i/a(t)$. Similarly,

$$F_1 = a(t)\, q_i Q_i$$

generates the canonical transformation given by $p_i = a(t)\, Q_i$, $P_i = -a(t)\, q_i$ [see (5.53)], that is

$$Q_i = \frac{1}{a(t)} p_i, \qquad P_i = -a(t)\, q_i$$

(an exchange of the coordinates and momenta with a possible change of scale) and $K = H - a'(t)\, Q_i P_i/a(t)$.

Exercise 5.34. Find a generating function for the canonical transformations considered in Example 5.29.

Example 5.35 (Particle in a uniform gravitational field). The function

$$F_3(p_1, p_2, Q_1, Q_2, t) = -p_1 Q_1 - p_2 Q_2 + \frac{{p_1}^2 p_2}{2m^2 g}, \tag{5.59}$$

where m and g are constants, satisfies the condition $\det(\partial^2 F_3/\partial p_i \partial Q_j) \neq 0$ everywhere (note that here it is *assumed* that the p_i and Q_i are functionally independent). Indeed

$$\frac{\partial F_3}{\partial p_1} = -Q_1 + \frac{p_1 p_2}{m^2 g}, \qquad \frac{\partial F_3}{\partial p_2} = -Q_2 + \frac{{p_1}^2}{2m^2 g},$$

and

$$\frac{\partial F_3}{\partial Q_1} = -p_1, \qquad \frac{\partial F_3}{\partial Q_2} = -p_2,$$

and therefore $\det(\partial^2 F_3/\partial p_i \partial Q_j) = 1$. Comparing with (5.56) we conclude that (5.59) generates the coordinate transformation

$$\begin{aligned} Q_1 &= q_1 + \frac{p_1 p_2}{m^2 g}, \quad & Q_2 &= q_2 + \frac{{p_1}^2}{2m^2 g}, \\ P_1 &= p_1, & P_2 &= p_2, \end{aligned} \tag{5.60}$$

which, by construction, is canonical.

This transformation can be applied to the Hamiltonian

$$H = \frac{{p_1}^2 + {p_2}^2}{2m} + mgq_2,$$

which corresponds to a particle of mass m in a uniform gravitational field. Since the coordinate transformation does not involve the time, we can take $K = H$, hence

$$K = \frac{{P_2}^2}{2m} + mgQ_2,$$

which has the form of the standard Hamiltonian for a particle in *one dimension* in a uniform gravitational field. As a consequence of the fact that Q_1 and P_1 do not appear in K, P_1 and Q_1, respectively, are constants of motion, that is [see Equations (5.60)]

$$p_1, \qquad q_1 + \frac{p_1 p_2}{m^2 g}$$

are constants of motion (as well as H). The conservation of p_1 also follows from the fact that q_1 is an ignorable coordinate in H. It should be clear that the equations of motion (and their solution) given the Hamiltonian H are equivalent to those given by K.

As shown above, for any canonical transformation, the function F_1 always exists (at least locally), but in order for F_1 to be useful as a generating function it is necessary that (q_i, Q_i, t) be functionally independent. Note that the function F_1 (as well as F_2, F_3, and F_4) can *always* be expressed in terms of q_i, p_i, and t, or of Q_i, P_i, and t (since any of these sets are coordinates on the extended phase space), and can be equal to zero (see, e.g., Example 5.36, below). The canonical transformations with $F_1 = 0$ are sometimes called *homogeneous* canonical transformations.

Example 5.36. In some cases we start with one half of the expressions that define a coordinate transformation in the extended phase space and we want to find the other half, in such a way that the transformation is canonical. For instance, we can have an *arbitrary* coordinate transformation in the configuration space

$$Q_i = Q_i(q_j, t) \tag{5.61}$$

(as usual, with the only condition that the transformation be invertible), and we want to find the expressions for the new momenta, P_i, in such a way that the resulting coordinate transformation in the extended phase space is canonical. Substituting (5.61) into the left-hand side of (5.46) we find that the new momenta P_i must be such that

$$\left(p_i - P_j \frac{\partial Q_j}{\partial q_i}\right) \mathrm{d}q_i + \left(K - H - P_i \frac{\partial Q_i}{\partial t}\right) \mathrm{d}t = \mathrm{d}F_1, \tag{5.62}$$

for some function F_1.

The left-hand side of (5.62) is trivially an exact differential if the coefficients of the $\mathrm{d}q_i$ and $\mathrm{d}t$ are all identically equal to zero; then the P_i are implicitly given by $p_i = P_j\, \partial Q_j/\partial q_i$, which amounts to

$$P_k = p_i \frac{\partial q_i}{\partial Q_k} \tag{5.63}$$

(making use of the fact that $(\partial Q_j/\partial q_i)(\partial q_i/\partial Q_k) = \delta_{jk}$), and

$$K = H + P_i \left(\frac{\partial Q_i}{\partial t}\right)_{q_j}. \tag{5.64}$$

For instance, the parabolic coordinates, u, v, can be defined in terms of the Cartesian coordinates of the plane by means of

$$x = \tfrac{1}{2}(u^2 - v^2), \qquad y = uv$$

[note that the relation is two-to-one; (u, v) and $(-u, -v)$ correspond to the same point (x, y)]. Equations (5.63) yield

$$p_u = p_x \frac{\partial x}{\partial u} + p_y \frac{\partial y}{\partial u} = p_x u + p_y v, \qquad p_v = p_x \frac{\partial x}{\partial v} + p_y \frac{\partial y}{\partial v} = -p_x v + p_y u$$

[cf. Equation (4.51)] and, from (5.64), $K = H$.

Going back to (5.62), we find that the most general expression for the P_i accompanying (5.61) is implicitly given by

$$p_i = P_j \frac{\partial Q_j}{\partial q_i} + \frac{\partial F_1}{\partial q_i}, \tag{5.65}$$

where F_1 is an *arbitrary* function of q_i and t only, and

$$K = H + P_i \frac{\partial Q_i}{\partial t} + \frac{\partial F_1}{\partial t} \tag{5.66}$$

(cf. Equation (5.53), note that in this case F_1 is not a generating function since, owing to (5.61), the set (q_i, Q_i, t) is not independent). In conclusion, given an arbitrary coordinate transformation in the configuration space (5.61), the new conjugate momenta are not defined in a unique way; they are given by (5.65) with F_1 being an arbitrary function of q_i and t.

In particular, if $Q_i = q_i$, from the last two equations we obtain

$$P_i = p_i - \frac{\partial F_1}{\partial q_i}, \qquad K = H + \frac{\partial F_1}{\partial t}, \tag{5.67}$$

where, as stated above, F_1 is an arbitrary function of q_i and t only. For instance, in the case of the Hamiltonian

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy, \tag{5.68}$$

where m and g are constants, we can eliminate the last term choosing $F_1 = -mgyt$. Then, from Equations (5.67) we obtain $P_1 = p_x$, $P_2 = p_y + mgt$, and the new Hamiltonian is

$$K = \frac{{P_1}^2 + (P_2 - mgt)^2}{2m}, \tag{5.69}$$

which does not contain Q_1 and Q_2 and, therefore, P_1 and P_2 are constants of motion. (Note that, by contrast with H, K depends on the time and, therefore, it is not conserved.)

Exercise 5.37. Assuming that

$$Q_1 = q_1 \cos q_2, \qquad Q_2 = q_1 \sin q_2,$$

find the most general form of P_1 and P_2 as functions of (q_1, q_2, p_1, p_2, t) in order to have a canonical transformation.

Interaction with an Electromagnetic Field. Gauge Transformations

An important example of canonical transformations that do not change the coordinates in the configuration space (i.e., $Q_i = q_i$) are those associated with the gauge transformations.

The standard Lagrangian for a particle in a conservative field of force with potential $V(\mathbf{r})$ is given by

$$L = \tfrac{1}{2}m\mathbf{v}^2 - V(\mathbf{r}),$$

where m is the mass of the particle and $\mathbf{v}$ is its velocity (with respect to some inertial frame). The addition of the interaction with an electromagnetic field, $\mathbf{E}$, $\mathbf{B}$, is taken into account (in cgs units) by adding the terms $(e/c)\mathbf{A} \cdot \mathbf{v} - e\varphi$ to the Lagrangian L, where e is the electric charge of the particle and φ, $\mathbf{A}$ are potentials for the fields $\mathbf{E}$ and $\mathbf{B}$ [see Equation (1.68)]. Thus, the complete Lagrangian is

$$\tilde{L} = L + \frac{e}{c}\mathbf{A} \cdot \mathbf{v} - e\varphi. \tag{5.70}$$

If we denote by π_i the original canonical momenta (without the interaction with the electromagnetic field),

$$\pi_i \equiv \frac{\partial L}{\partial \dot{q}_i},$$

then, in the presence of the electromagnetic field, the canonical momenta are

$$p_i = \frac{\partial \tilde{L}}{\partial \dot{q}_i} = \pi_i + \frac{e}{c}A_i, \tag{5.71}$$

where [see Equation (1.42)]

$$A_i \equiv \mathbf{A} \cdot \frac{\partial \mathbf{v}}{\partial \dot{q}_i} = \mathbf{A} \cdot \frac{\partial \mathbf{r}}{\partial q_i}. \tag{5.72}$$

(Note that the q_i are arbitrary coordinates, not necessarily Cartesian or orthogonal.)

As pointed out in Section 1.2, the electromagnetic potentials are not unique: if $\mathbf{A}$ and φ are potentials for $\mathbf{E}$ and $\mathbf{B}$, then, for any well-behaved function $\xi(\mathbf{r}, t)$, $\mathbf{A} + \nabla\xi$, and $\varphi - (1/c)\partial\xi/\partial t$ are also potentials for $\mathbf{E}$ and $\mathbf{B}$. The transformation

$$\mathbf{A} \mapsto \mathbf{A} + \nabla\xi, \qquad \varphi \mapsto \varphi - \frac{1}{c}\frac{\partial\xi}{\partial t} \tag{5.73}$$

is called a gauge transformation and from Equations (5.70) and (5.73) we see that, under a gauge transformation, the Lagrangian $\tilde{L}$ transforms according to

$$\tilde{L} \mapsto \tilde{L} + \frac{\partial(e\xi/c)}{\partial q_i}\dot{q}_i + \frac{\partial(e\xi/c)}{\partial t}$$

[cf. (1.103)]. Thus, under the gauge transformation (5.73), the canonical momenta transform according to

$$p_i \mapsto p_i + \frac{\partial(e\xi/c)}{\partial q_i},$$

which is of the form (5.67), and, if $\tilde{H}$ denotes the Hamiltonian corresponding to the complete Lagrangian $\tilde{L}$, making use of the last two equations we have

$$\tilde{H} \mapsto \tilde{H} - \frac{\partial(e\xi/c)}{\partial t},$$

which is of the form (5.67).

In conclusion, in the presence of an electromagnetic field, the canonical momenta are gauge-dependent, but a gauge transformation corresponds to a canonical transformation that leaves the coordinates q_i unchanged. As we shall prove below, the Poisson bracket is invariant under canonical transformations [see Equation (5.112)], therefore, even though the canonical momenta are gauge-dependent, the Poisson bracket *does not* depend on the choice of the electromagnetic potentials.

Example 5.38 (Kepler problem in two dimensions with energy equal to zero). Apart from the canonical transformations arising from coordinate transformations in the configuration space, considered in Example 5.36, in some cases we may start with relations between the old canonical momenta, p_i, and the new coordinates, Q_i. For instance, considering again the Kepler problem in two dimensions with energy equal to zero (cf. Section 4.3), the Cartesian components of the linear momentum (p_x, p_y), appearing in (4.44), will be expressed in terms of new coordinates (Q_1, Q_2) by means of the inversion in a circle of unit radius

$$(Q_1, Q_2) \equiv \frac{(p_x, p_y)}{{p_x}^2 + {p_y}^2}. \tag{5.74}$$

Thus,

$$Q_1{}^2 + Q_2{}^2 = \frac{1}{{p_x}^2 + {p_y}^2} \tag{5.75}$$

and, therefore,

$$(p_x, p_y) = \frac{(Q_1, Q_2)}{Q_1{}^2 + Q_2{}^2} \tag{5.76}$$

[cf. Equation (5.74)].

According to (5.76), the set (x, y, Q_1, Q_2, t) is functionally independent and, therefore, the desired canonical transformation can be obtained by means of a type F_1 generating function. Since $p_i = (\partial F_1/\partial q_i)_{q_j, Q_j, t}$ [see (5.53)], we can take

$$F_1 = \frac{Q_1 x + Q_2 y}{Q_1{}^2 + Q_2{}^2}$$

and from $P_i = -(\partial F_1/\partial Q_i)_{q_j,Q_j,t}$ we obtain

$$P_1 = \frac{x(Q_1{}^2 - Q_2{}^2) + 2yQ_1Q_2}{(Q_1{}^2 + Q_2{}^2)^2}, \qquad P_2 = \frac{y(Q_2{}^2 - Q_1{}^2) + 2xQ_1Q_2}{(Q_1{}^2 + Q_2{}^2)^2}. \tag{5.77}$$

Furthermore, we can take $K = H$.

From Equations (5.77) we obtain

$$P_1{}^2 + P_2{}^2 = \frac{x^2 + y^2}{(Q_1{}^2 + Q_2{}^2)^2}$$

and, therefore, the new Hamiltonian is given by [see Equation (4.44)]

$$K = \frac{1}{2m(Q_1{}^2 + Q_2{}^2)}\left(1 - \frac{2mk}{\sqrt{P_1{}^2 + P_2{}^2}}\right).$$

Introducing

$$h \equiv \frac{P_1{}^2 + P_2{}^2}{2m}, \tag{5.78}$$

which has the form of the standard Hamiltonian of a free particle in Cartesian coordinates, we see that

$$K = \frac{1}{2m(Q_1{}^2 + Q_2{}^2)}\left(1 - \frac{2mk}{\sqrt{2mh}}\right), \tag{5.79}$$

and that the hypersurface $K = 0$ is also defined by $h = 2mk^2$.

Equation (5.79) is of the form (4.53), with

$$f(Q_i, P_i) = \frac{1}{2m(Q_1{}^2 + Q_2{}^2)}, \qquad g(x) = 1 - \frac{2mk}{\sqrt{2mx + 4m^2k^2}}$$

and, therefore, making use of the parameter τ defined by

$$\mathrm{d}\tau = \frac{\mathrm{d}t}{8m^2k^2(Q_1{}^2 + Q_2{}^2)} \tag{5.80}$$

[see Equation (4.55)], on the hypersurface $K = 0$, the equations of motion can be written in the form of the Hamilton equations

$$\frac{\mathrm{d}Q_i}{\mathrm{d}\tau} = \frac{\partial h}{\partial P_i}, \qquad \frac{\mathrm{d}P_i}{\mathrm{d}\tau} = -\frac{\partial h}{\partial Q_i}.$$

By virtue of the form of the Hamiltonian (5.78), one concludes that the solution of these last equations is given by

$$P_i = (P_i)_0, \qquad Q_i = \frac{(P_i)_0}{m}\tau + (Q_i)_0, \tag{5.81}$$

where the $(P_i)_0$ and $(Q_i)_0$ are constants. Thus, the orbit in the Q_1Q_2-plane is a straight line. (As we shall show now, owing to (5.76), this straight line corresponds to a circle in the $p_x p_y$-plane.)

Making use of Equations (5.76) and (5.77), we can find the solution of the equations of motion in the original Cartesian coordinates, albeit parameterized by τ. Instead of following this direct approach, it is convenient to take advantage of the fact that (as corresponds to a free particle) P_1 and P_2 are conserved as well as $Q_1P_2 - Q_2P_1$ [as can be seen from (5.81)] and with the aid of (5.77) and (5.76) we find that its value is

$$\begin{aligned} Q_1P_2 - Q_2P_1 &= \frac{Q_1\big[y({Q_2}^2 - {Q_1}^2) + 2xQ_1Q_2\big] - Q_2\big[x({Q_1}^2 - {Q_2}^2) + 2yQ_1Q_2\big]}{({Q_1}^2 + {Q_2}^2)^2} \\ &= \frac{xQ_2 - yQ_1}{{Q_1}^2 + {Q_2}^2} \\ &= xp_y - yp_x \\ &= L_3 \end{aligned}$$

(the angular momentum of the particle about the origin). On the other hand, from (5.74), we also have

$$Q_1P_2 - Q_2P_1 = \frac{p_xP_2 - p_yP_1}{{p_x}^2 + {p_y}^2}.$$

Hence, combining the two previous expressions, if $L_3 \neq 0$, we obtain

$${p_x}^2 + {p_y}^2 = \frac{p_xP_2 - p_yP_1}{L_3}$$

(which relates p_x, p_y with the constants P_1, P_2, and L_3) or, equivalently,

$$\left(p_x - \frac{P_2}{2L_3}\right)^2 + \left(p_y + \frac{P_1}{2L_3}\right)^2 = \frac{{P_1}^2 + {P_2}^2}{(2L_3)^2} = \left(\frac{mk}{L_3}\right)^2, \tag{5.82}$$

which is the equation of a circle in the $p_x p_y$-plane of radius $mk/|L_3|$ passing through the origin (though actually the point (p_x, p_y) never reaches the origin, as can be seen taking into account that in the present case the total energy is equal to zero and therefore ${p_x}^2 + {p_y}^2 = 2mk/r$; a closed curve would correspond to a periodic motion) [cf. Equation (4.77)]. Thus, we obtain, in the first place, not the orbit in the configuration space, but the hodograph.

By combining (5.77), (5.74), and (4.44), we find that on the hypersurface $K = 0$,

$$\begin{aligned} P_1 &= x({p_x}^2 - {p_y}^2) + 2yp_x p_y \\ &= x({p_x}^2 + {p_y}^2 - 2{p_y}^2) + 2yp_x p_y \\ &= ({p_x}^2 + {p_y}^2)\,x - 2p_y(xp_y - yp_x) \\ &= \frac{2mk}{r}x - 2p_y L_3 \end{aligned}$$

and in a similar way we find that

$$P_2 = \frac{2mk}{r}y + 2p_x L_3.$$

Thus, the constants of motion $-\frac{1}{2}P_1$ and $-\frac{1}{2}P_2$ are the x- and y-components, respectively, of the Laplace–Runge–Lenz vector

$$\mathbf{A} = \mathbf{p} \times \mathbf{L} - \frac{mk}{r}\mathbf{r},$$

if, as in Exercise 4.31, we identify a vector (a_1, a_2) in the plane with the vector $(a_1, a_2, 0)$, and $\mathbf{L} = (0, 0, L_3)$.

Substituting Equation (5.75) into (5.80) we see that, on the hypersurface $K = 0$,

$$\mathrm{d}\tau = \frac{\mathrm{d}t}{8m^2k^2({Q_1}^2 + {Q_2}^2)} = \frac{({p_x}^2 + {p_y}^2)\,\mathrm{d}t}{8m^2k^2} = \frac{\mathrm{d}t}{4mk\sqrt{x^2 + y^2}},$$

which shows that the parameter τ employed here is, apart from the constant factor $1/2mk$, the auxiliary parameter introduced in Section 4.3 [see Equations (4.47) and (4.41)].

In the following example we show that a similar treatment can be given in the case where $E < 0$, with the aid of an appropriate mapping. (The case with $E > 0$ can be found in, e.g., Torres del Castillo [16].)

Example 5.39 (Kepler problem in two dimensions with negative energy). In the case where $E < 0$, it is convenient to parameterize the Cartesian components of the momentum of the particle in terms of the standard coordinates, θ, ϕ, of a sphere of radius

$$p_0 \equiv \sqrt{-2mE} \tag{5.83}$$

according to (see Figure 5.4)

$$(p_x, p_y) = \frac{p_0}{1-\cos\theta}(\sin\theta\cos\phi, \sin\theta\sin\phi) = p_0 \cot\frac{\theta}{2}(\cos\phi, \sin\phi), \qquad (5.84)$$

where $0 \leqslant \theta \leqslant \pi$, $0 \leqslant \phi \leqslant 2\pi$. Then,

$$p_x{}^2 + p_y{}^2 = p_0{}^2 \cot^2\frac{\theta}{2}, \qquad (5.85)$$

and

$$\begin{aligned} p_x \mathrm{d}x + p_y \mathrm{d}y &= p_0 \cot\frac{\theta}{2}(\cos\phi\, \mathrm{d}x + \sin\phi\, \mathrm{d}y) \\ &= \mathrm{d}\left[p_0 \cot\frac{\theta}{2}(x\cos\phi + y\sin\phi)\right] \\ &\quad + \frac{1}{2}p_0 \csc^2\frac{\theta}{2}\big(x\cos\phi + y\sin\phi\big)\mathrm{d}\theta + p_0\cot\frac{\theta}{2}\big(x\sin\phi - y\cos\phi\big)\mathrm{d}\phi. \end{aligned}$$

The last expression shows that θ, ϕ, and

$$p_\theta \equiv \frac{1}{2}p_0 \csc^2\frac{\theta}{2}\big(x\cos\phi + y\sin\phi\big), \qquad p_\phi \equiv p_0 \cot\frac{\theta}{2}\big(x\sin\phi - y\cos\phi\big) \qquad (5.86)$$

are canonical coordinates and from Equations (5.86) it follows that

$$x^2 + y^2 = \frac{4}{p_0{}^2}\sin^4\frac{\theta}{2}\left(p_\theta{}^2 + \frac{p_\phi{}^2}{\sin^2\theta}\right). \qquad (5.87)$$

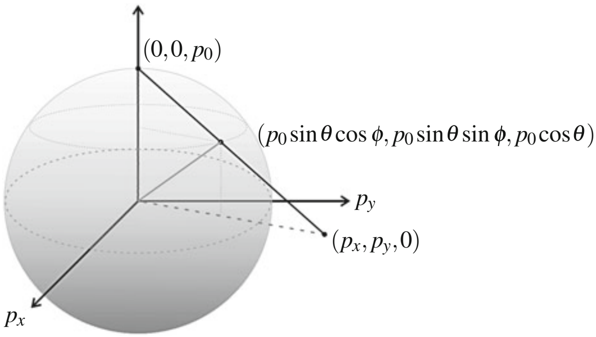


Fig. 5.4 By means of the stereographic projection, the Cartesian components of the momentum, p_x, p_y, are parameterized by the spherical coordinates θ, ϕ of a sphere of radius $p_0 = \sqrt{-2mE}$. The point $(p_0 \sin\theta\cos\phi,\ p_0\sin\theta\sin\phi,\ p_0\cos\theta)$ is the intersection of this sphere with the straight line joining the points $(0, 0, p_0)$ and $(p_x, p_y, 0)$

(Alternatively, we can see that (5.84) can be obtained by means of the generating function $F_1 = p_0 \cot \frac{\theta}{2}(x \cos\phi + y \sin\phi)$.)

Since the canonical transformation given by (5.84) and (5.86) does not involve the time explicitly, we can choose $K = H$, hence, making use of (5.85), (5.87), and (5.83),

$$\begin{aligned} K &= \frac{{p_0}^2}{2m} \cot^2 \frac{\theta}{2} - \frac{k p_0}{2 \sin^2(\theta/2)\sqrt{{p_\theta}^2 + {p_\phi}^2/\sin^2\theta}} \\ &= E + \frac{k p_0}{2} \csc^2 \frac{\theta}{2} \left(\frac{p_0}{mk} - \frac{1}{\sqrt{2m h_E}} \right), \end{aligned} \tag{5.88}$$

where we have introduced the auxiliary function

$$h_E \equiv \frac{1}{2m} \left({p_\theta}^2 + \frac{{p_\phi}^2}{\sin^2\theta} \right), \tag{5.89}$$

which has the form of the standard Hamiltonian of a free particle on a sphere of radius 1 (a spherical pendulum without the gravitational field), in spherical coordinates [see Equation (2.31)].

Equation (5.88) is of the form (4.53) with

$$f(\theta, \phi, p_\theta, p_\phi) = \frac{k p_0}{2} \csc^2 \frac{\theta}{2}, \qquad g(x) = -\frac{p_0}{\sqrt{2m{p_0}^2 x + m^2 k^2}}.$$

Hence, on the hypersurface $K = E$ (or, equivalently, $h_E = mk^2/2{p_0}^2$), the equations of motion are also given by the equivalent Hamiltonian h_E with the parameter τ being related to the time by [see Equation (4.55)]

$$\mathrm{d}\tau = \frac{{p_0}^4}{2m^2k^2} \csc^2 \frac{\theta}{2}\, \mathrm{d}t.$$

Making use of (5.85) and (5.83) we find that, on the hypersurface $H = E$, we have the equivalent expressions

$$\mathrm{d}\tau = \frac{{p_0}^4}{2m^2k^2} \left(1 + \cot^2 \frac{\theta}{2} \right) \mathrm{d}t = \frac{{p_0}^2}{2m^2k^2} \left({p_0}^2 + {p_x}^2 + {p_y}^2 \right) \mathrm{d}t = -\frac{2E\, \mathrm{d}t}{k\sqrt{x^2 + y^2}},$$

which show that the parameter τ is, apart from the constant factor $-4E/k$, the auxiliary parameter introduced in Section 4.3 [Equation (4.47)].

Since h_E has the form of the Hamiltonian of a free particle in a sphere, the three Cartesian components of the angular momentum,

$$l_1 = -\sin\phi\, p_\theta - \cot\theta \cos\phi\, p_\phi, \quad l_2 = \cos\phi\, p_\theta - \cot\theta \sin\phi\, p_\phi, \quad l_3 = p_\phi \tag{5.90}$$

are constants of motion, but they are not independent. In fact, making use of (5.90) and (5.89) one readily verifies that

$$l_1{}^2 + l_2{}^2 + l_3{}^2 = p_\theta{}^2 + \frac{p_\phi{}^2}{\sin^2\theta} = 2mh_E = \left(\frac{mk}{p_0}\right)^2. \tag{5.91}$$

Making use of the explicit expressions (5.90) one obtains

$$l_1 \sin\theta \cos\phi + l_2 \sin\theta \sin\phi + l_3 \cos\theta = 0,$$

which is the equation of a plane passing through the origin with normal vector (l_1, l_2, l_3) and shows that the orbit in the sphere is a great circle (that is, the intersection of the sphere with a plane passing through its center). Making use of (5.84) we can eliminate $\cos\phi$ and $\sin\phi$ in favor of p_x and p_y, which yields

$$l_1 \sin\theta \frac{p_x}{p_0 \cot(\theta/2)} + l_2 \sin\theta \frac{p_y}{p_0 \cot(\theta/2)} + l_3 \cos\theta = 0,$$

and with the aid of (5.85) we obtain

$$2p_0(l_1 p_x + l_2 p_y) + l_3(p_x{}^2 + p_y{}^2 - p_0{}^2) = 0,$$

so that if $l_3 \neq 0$,

$$\left(p_x + \frac{p_0 l_1}{l_3}\right)^2 + \left(p_y + \frac{p_0 l_2}{l_3}\right)^2 = \frac{p_0{}^2}{l_3{}^2}(l_1{}^2 + l_2{}^2 + l_3{}^2) = \left(\frac{mk}{l_3}\right)^2$$

[see (5.91)]. This last equation shows that the hodograph is a circle enclosing the origin.

Time Evolution

As pointed out in Section 5.1, if $q_i(t)$, $p_i(t)$ represent the solution of the Hamilton equations, then the relation between $q_i(t_0)$, $p_i(t_0)$ and $q_i(t)$, $p_i(t)$ is a canonical transformation, for any values of t_0 and t. More precisely, if

$$q_i = q_i(Q_j, P_j, t), \qquad p_i = p_i(Q_j, P_j, t), \tag{5.92}$$

is the solution of the Hamilton equations, where Q_i and P_i are the values of q_i and p_i at $t = t_0$, respectively, that is

$$q_i(Q_j, P_j, t_0) = Q_i, \qquad p_i(Q_j, P_j, t_0) = P_i, \tag{5.93}$$

then (5.92) is a canonical transformation.

Example 5.40. The solution of the Hamilton equations corresponding to the Hamiltonian

$$H = \frac{p^2}{2m} - ktq,$$

where k is some constant, is given by

$$q = Q + \frac{Pt}{m} + \frac{kt^3}{6m}, \qquad p = P + \frac{kt^2}{2}, \tag{5.94}$$

where Q and P are the values of q and p, respectively, at $t = 0$. With the aid of the Poisson bracket, one can readily verify that (5.94) is a canonical transformation (that is, $\partial(q, p)/\partial(Q, P) = 1$).

In order to prove that (5.92) is always a canonical transformation it suffices to show that the Lagrange brackets $[Q_i, Q_j]_{q,p}$, $[Q_i, P_j]_{q,p}$, and $[P_i, P_j]_{q,p}$ do not depend on the time (see Equations (5.37); note that here we are taking Q_i and P_i as the "old" coordinates and q_i and p_i as the new ones). Making use of (5.40), the definition of the Lagrange brackets, the fact that, by hypothesis,

$$\frac{\partial q_i(Q_j, P_j, t)}{\partial t} = \frac{\partial H}{\partial p_i}, \qquad \frac{\partial p_i(Q_j, P_j, t)}{\partial t} = -\frac{\partial H}{\partial q_i}, \tag{5.95}$$

and the chain rule, we have, for instance,

$$\begin{aligned}
\frac{\partial}{\partial t}[Q_i, Q_j]_{q,p} &= -\frac{\partial}{\partial Q_i}[Q_j, t]_{q,p} - \frac{\partial}{\partial Q_j}[t, Q_i]_{q,p} \\
&= \frac{\partial}{\partial Q_i}[t, Q_j]_{q,p} - \frac{\partial}{\partial Q_j}[t, Q_i]_{q,p} \\
&= \frac{\partial}{\partial Q_i}\left(\frac{\partial q_k}{\partial t}\frac{\partial p_k}{\partial Q_j} - \frac{\partial p_k}{\partial t}\frac{\partial q_k}{\partial Q_j}\right) - \frac{\partial}{\partial Q_j}\left(\frac{\partial q_k}{\partial t}\frac{\partial p_k}{\partial Q_i} - \frac{\partial p_k}{\partial t}\frac{\partial q_k}{\partial Q_i}\right) \\
&= \frac{\partial}{\partial Q_i}\left(\frac{\partial H}{\partial p_k}\frac{\partial p_k}{\partial Q_j} + \frac{\partial H}{\partial q_k}\frac{\partial q_k}{\partial Q_j}\right) - \frac{\partial}{\partial Q_j}\left(\frac{\partial H}{\partial p_k}\frac{\partial p_k}{\partial Q_i} + \frac{\partial H}{\partial q_k}\frac{\partial q_k}{\partial Q_i}\right) \\
&= \frac{\partial^2 H}{\partial Q_i \partial Q_j} - \frac{\partial^2 H}{\partial Q_j \partial Q_i} \\
&= 0,
\end{aligned} \tag{5.96}$$

and in a similar way one finds that $[Q_i, P_j]_{q,p}$, and $[P_i, P_j]_{q,p}$ do not depend on the time. According to (5.93), at $t = t_0$, the Lagrange brackets $[Q_i, Q_j]_{q,p}$, $[Q_i, P_j]_{q,p}$, and $[P_i, P_j]_{q,p}$ have the values 0, δ_{ij}, and 0, respectively, and, therefore, these values are maintained all the time, thus showing that the transformation (5.92) is canonical.

Of course, if Equations (5.92) do correspond to a canonical transformation, a relation of the form (5.46) must be satisfied. In fact, we have the following simple and very useful result.

Proposition 5.41. *If the equations $q_i = q_i(Q_j, P_j, t)$, $p_i = p_i(Q_j, P_j, t)$ express the solution of the Hamilton equations, where Q_i and P_i are the values of q_i and p_i, respectively, at some time $t = t_0$, then*

$$p_i \mathrm{d}q_i - H\mathrm{d}t - P_i \mathrm{d}Q_i = \mathrm{d}F, \tag{5.97}$$

for some locally defined function F.

Proof. Writing the left-hand side of (5.97) in the form

$$\left(p_j \frac{\partial q_j}{\partial Q_i} - P_i\right) \mathrm{d}Q_i + p_j \frac{\partial q_j}{\partial P_i} \mathrm{d}P_i + \left(p_j \frac{\partial q_j}{\partial t} - H\right) \mathrm{d}t$$

and using the fact that (5.92) is a canonical transformation (i.e., Equations (5.41) hold), we only have to verify that

$$\frac{\partial}{\partial t}\left(p_j \frac{\partial q_j}{\partial Q_i} - P_i\right) = \frac{\partial}{\partial Q_i}\left(p_j \frac{\partial q_j}{\partial t} - H\right), \quad \frac{\partial}{\partial t}\left(p_j \frac{\partial q_j}{\partial P_i}\right) = \frac{\partial}{\partial P_i}\left(p_j \frac{\partial q_j}{\partial t} - H\right),$$

using Q_i, P_i, and t as the independent variables. We have, for instance, by virtue of (5.95) and the chain rule,

$$\begin{aligned}
\frac{\partial}{\partial t}&\left(p_j \frac{\partial q_j}{\partial Q_i} - P_i\right) - \frac{\partial}{\partial Q_i}\left(p_j \frac{\partial q_j}{\partial t} - H\right) \\
&= \frac{\partial p_j}{\partial t}\frac{\partial q_j}{\partial Q_i} + p_j \frac{\partial^2 q_j}{\partial t \partial Q_i} - 0 - \frac{\partial p_j}{\partial Q_i}\frac{\partial q_j}{\partial t} - p_j \frac{\partial^2 q_j}{\partial Q_i \partial t} + \frac{\partial H}{\partial Q_i} \\
&= -\frac{\partial H}{\partial q_j}\frac{\partial q_j}{\partial Q_i} - \frac{\partial p_j}{\partial Q_i}\frac{\partial H}{\partial p_j} + \frac{\partial H}{\partial Q_i} \\
&= 0,
\end{aligned}$$

and in a similar manner one readily finds that the remaining conditions are satisfied, showing that the left-hand side of (5.97) is locally exact. □

Comparing (5.97) with (5.46) we see that the Hamiltonian corresponding to the coordinates Q_i, P_i can be taken as $K = 0$, which is consistent with the fact that Q_i and P_i, being initial values, are constants of motion. The function F appearing on the right-hand side of Equation (5.97) is called *Hamilton's principal function*.

For instance, substituting the relation (5.94) into the left-hand side of (5.97) one finds that

$$p\mathrm{d}q - H\mathrm{d}t - P\mathrm{d}Q = \mathrm{d}\left[\frac{kt^2q}{3} + \frac{kt^2Q}{6} + \frac{m}{2t}(q-Q)^2 - \frac{k^2t^5}{90m}\right], \tag{5.98}$$

which is indeed of the form (5.97). For $t \neq 0$, the function inside the brackets on the right-hand side of the last equation is a generating function of the canonical transformation (5.94), which gives the time evolution of the system considered. (However, q, P, and t are functionally independent everywhere and, therefore, a generating function of type F_2 will reproduce the transformation (5.94) without restrictions.)

If $\tilde{Q}_i$, $\tilde{P}_i$ are canonical coordinates and, at the same time, constants of motion, then they must be functions of the initial conditions, Q_i, P_i, only, and there must exist a time-independent canonical transformation relating these sets of canonical coordinates, that is,

$$P_i\mathrm{d}Q_i - \tilde{P}_i\mathrm{d}\tilde{Q}_i = \mathrm{d}Z, \tag{5.99}$$

for some function Z. Thus, from (5.97) and (5.99) it follows that

$$p_i\mathrm{d}q_i - H\mathrm{d}t - \tilde{P}_i\mathrm{d}\tilde{Q}_i = \mathrm{d}(F+Z)$$

and, therefore, dropping the tildes and renaming $F+Z$ as F, we have the following result.

Proposition 5.42. *If the equations $q_i = q_i(Q_j, P_j, t)$, $p_i = p_i(Q_j, P_j, t)$ express the solution of the Hamilton equations in terms of a set of canonical coordinates, Q_i, P_i, which are constants of motion, then*

$$p_i\mathrm{d}q_i - H\mathrm{d}t - P_i\mathrm{d}Q_i = \mathrm{d}F, \tag{5.100}$$

for some locally defined function F.

By abuse, the function F appearing on the right-hand side of Equation (5.100) is also called Hamilton's principal function.

Example 5.43. The functions

$$Q = \frac{p^2}{2m} + mgq, \qquad P = \frac{p}{mg} + t, \tag{5.101}$$

where g is a constant, are canonical coordinates (that is, $\{Q, P\} = 1$) and constants of motion if the Hamiltonian is given by

$$H = \frac{p^2}{2m} + mgq.$$

(Note that by simply setting $t = 0$ in Equations (5.101) one finds the relation between Q and P and the initial values of q and p, and that Q and P are not the initial values of q and p, respectively.) By inverting the expressions for Q and P one finds that

$$q = \frac{Q}{mg} - \frac{g}{2}(P - t)^2, \qquad p = mg(P - t), \tag{5.102}$$

which give the solution of the Hamilton equations, and, making use of these explicit expressions, one obtains

$$p\,\mathrm{d}q - H\,\mathrm{d}t - P\,\mathrm{d}Q = \mathrm{d}\left[-Qt \mp \frac{1}{3m^2 g}\left(2mQ - 2m^2 gq\right)^{3/2}\right]. \tag{5.103}$$

Since (q, Q, t) is functionally independent, the function inside the brackets in the last equation is a generating function of the canonical transformation (5.101), which represents the time evolution of the system.

Propositions 5.41 and 5.42 contain a remarkable result: for any Hamiltonian, with any number of degrees of freedom, the entire solution of the corresponding Hamilton equations is given by some generating function. In the two preceding examples, that generating function has been obtained from the explicit expression of the solution of the Hamilton equations. However, as we shall see in the next chapter, the generating function of the time evolution can be obtained directly if one is able to solve certain PDE satisfied by the generating function.

Characterization of the Canonical Transformations by Means of the Poisson Brackets

The definition of the canonical transformations given above is expressed in terms of Lagrange brackets [see Equations (5.41)] and, as we have shown, leads directly to the existence of the generating functions. As we shall show now, the canonical transformations can also be defined making use of the Poisson bracket. (Recall that when the number of degrees of freedom is equal to 1, the Lagrange bracket $[q, p]$ coincides with the Poisson bracket $\{Q, P\}$.)

Proposition 5.44. *Equation (5.41) is equivalent to*

$$\{Q_i, Q_k\} = 0, \qquad \{P_i, P_k\} = 0, \qquad \{Q_i, P_k\} = \delta_{ik}. \tag{5.104}$$

In other words, a transformation is canonical if and only if Equations (5.104) hold.

Proof. In order to prove this Proposition it is convenient to employ the notation

$$(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{2n}) \equiv (q_1, \ldots, q_n, p_1, \ldots, p_n)$$

[see Equation (4.25)] and, similarly,

$$(y_1, \ldots, y_n, y_{n+1}, \ldots, y_{2n}) \equiv (Q_1, \ldots, Q_n, P_1, \ldots, P_n).$$

Then, making use of the $2n \times 2n$ block matrix

$$(\epsilon_{\alpha\beta}) \equiv \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{5.105}$$

($\alpha, \beta = 1, 2, \ldots, 2n$), where I is the $n \times n$ unit matrix, we have

$$\{f, g\} = \varepsilon_{\alpha\beta} \frac{\partial f}{\partial x_\alpha} \frac{\partial g}{\partial x_\beta}, \tag{5.106}$$

with sum over repeated indices [see Equation (4.28)]. In a similar manner, the Lagrange bracket (5.37) can be written as

$$[u, v]_{Q,P} = \varepsilon_{\alpha\beta} \frac{\partial y_\alpha}{\partial u} \frac{\partial y_\beta}{\partial v}. \tag{5.107}$$

Thus, according to the definition (5.41), the transformation $Q_i = Q_i(q_j, p_j, t)$, $P_i = P_i(q_j, p_j, t)$ is canonical if

$$\varepsilon_{\alpha\beta} \frac{\partial y_\alpha}{\partial x_\mu} \frac{\partial y_\beta}{\partial x_\nu} = \varepsilon_{\mu\nu}, \tag{5.108}$$

while Equations (5.104) are equivalent to

$$\varepsilon_{\mu\nu} \frac{\partial y_\alpha}{\partial x_\mu} \frac{\partial y_\beta}{\partial x_\nu} = \varepsilon_{\alpha\beta} \tag{5.109}$$

and we want to prove that (5.108) is equivalent to (5.109).

Multiplying both sides of Equation (5.108) by $\partial x_\nu / \partial y_\gamma$, using the fact that, as a consequence of the chain rule, $(\partial y_\beta / \partial x_\nu)(\partial x_\nu / \partial y_\gamma) = \delta_{\beta\gamma}$ we obtain

$$\varepsilon_{\alpha\gamma} \frac{\partial y_\alpha}{\partial x_\mu} = \varepsilon_{\mu\nu} \frac{\partial x_\nu}{\partial y_\gamma}.$$

Now, multiplying both sides of the last equation by $\varepsilon_{\beta\gamma}$, using the fact that $\varepsilon_{\alpha\gamma}\varepsilon_{\beta\gamma} = \delta_{\alpha\beta}$, which follows from (5.105), we have

$$\frac{\partial y_\beta}{\partial x_\mu} = \varepsilon_{\beta\gamma}\varepsilon_{\mu\nu}\frac{\partial x_\nu}{\partial y_\gamma}. \tag{5.110}$$

Making use of (5.110), the antisymmetry of the matrix $(\varepsilon_{\alpha\beta})$, the relations employed above, and the chain rule we have (changing the names of the indices as necessary, in order to avoid that a summation index appears more than twice in a term)

$$\begin{aligned}
\varepsilon_{\mu\nu}\frac{\partial y_\alpha}{\partial x_\mu}\frac{\partial y_\beta}{\partial x_\nu} &= \varepsilon_{\mu\nu}\frac{\partial y_\alpha}{\partial x_\mu}\left(\varepsilon_{\beta\gamma}\varepsilon_{\nu\rho}\frac{\partial x_\rho}{\partial y_\gamma}\right) \\
&= -\varepsilon_{\mu\nu}\varepsilon_{\beta\gamma}\varepsilon_{\rho\nu}\frac{\partial y_\alpha}{\partial x_\mu}\frac{\partial x_\rho}{\partial y_\gamma} \\
&= -\delta_{\mu\rho}\varepsilon_{\beta\gamma}\frac{\partial y_\alpha}{\partial x_\mu}\frac{\partial x_\rho}{\partial y_\gamma} \\
&= -\varepsilon_{\beta\gamma}\frac{\partial y_\alpha}{\partial x_\mu}\frac{\partial x_\mu}{\partial y_\gamma} \\
&= -\varepsilon_{\beta\gamma}\delta_{\alpha\gamma} \\
&= \varepsilon_{\alpha\beta},
\end{aligned}$$

thus showing that (5.109) follows from (5.108). In a similar way, starting from (5.109) one obtains (5.110), which, in turn, leads to (5.108). □

Note that Equations (5.110) amount to the more explicit expressions

$$\frac{\partial Q_i}{\partial q_j} = \frac{\partial p_j}{\partial P_i}, \qquad \frac{\partial Q_i}{\partial p_j} = -\frac{\partial q_j}{\partial P_i}, \qquad \frac{\partial P_i}{\partial q_j} = -\frac{\partial p_j}{\partial Q_i}, \qquad \frac{\partial P_i}{\partial p_j} = \frac{\partial q_j}{\partial Q_i}. \tag{5.111}$$

The properties of the Poisson bracket given in Section 4.2 make it more convenient than the Lagrange bracket to determine if a given transformation is canonical or not. For instance, in the case of the coordinate transformation (5.47), making use of the bilinearity of the Poisson bracket and the fact that the only nonzero Poisson brackets between the coordinates q_i, p_i are those containing one coordinate and its conjugate momentum [see (4.30)] one readily finds that the brackets $\{Q_i, Q_j\}$ and $\{P_i, P_j\}$ are all equal to zero and

$$\{Q_1, P_3\} = \{Q_2, P_3\} = \{Q_3, P_1\} = \{Q_3, P_2\} = 0.$$

On the other hand

$$\{Q_1, P_1\} = \{\frac{1}{2}q_1 - \frac{1}{\mu}p_2, p_1 + \frac{\mu}{2}q_2\} = \frac{1}{2}\{q_1, p_1\} - \frac{1}{2}\{p_2, q_2\} = 1,$$

$$\{Q_1, P_2\} = \{\frac{1}{2}q_1 - \frac{1}{\mu}p_2, p_1 - \frac{\mu}{2}q_2\} = \frac{1}{2}\{q_1, p_1\} + \frac{1}{2}\{p_2, q_2\} = 0,$$

$$\{Q_2, P_1\} = \{\frac{1}{2}q_1 + \frac{1}{\mu}p_2, p_1 + \frac{\mu}{2}q_2\} = \frac{1}{2}\{q_1, p_1\} + \frac{1}{2}\{p_2, q_2\} = 0,$$

$$\{Q_2, P_2\} = \{\frac{1}{2}q_1 + \frac{1}{\mu}p_2, p_1 - \frac{\mu}{2}q_2\} = \frac{1}{2}\{q_1, p_1\} - \frac{1}{2}\{p_2, q_2\} = 1,$$

thus showing that the transformation is canonical.

Example 5.45. Making use of Proposition 5.44 and the properties of the Poisson bracket established in Section 4.2, we can readily verify that the coordinate transformation

$$Q_1 = \frac{c}{eB_0}\left[\left(p_x + \frac{eB_0}{2c}y\right)\cos\omega_c t - \left(p_y - \frac{eB_0}{2c}x\right)\sin\omega_c t\right],$$

$$Q_2 = \frac{c}{eB_0}\left(p_y + \frac{eB_0}{2c}x\right),$$

$$P_1 = \left(p_y - \frac{eB_0}{2c}x\right)\cos\omega_c t + \left(p_x + \frac{eB_0}{2c}y\right)\sin\omega_c t,$$

$$P_2 = p_x - \frac{eB_0}{2c}y,$$

where $\omega_c \equiv eB_0/mc$, is canonical. Indeed, letting

$$\pi_1 \equiv p_x + \frac{eB_0}{2c}y, \qquad \pi_2 \equiv p_y - \frac{eB_0}{2c}x, \qquad \sigma_1 \equiv p_x - \frac{eB_0}{2c}y, \qquad \sigma_2 \equiv p_y + \frac{eB_0}{2c}x,$$

we find that (see Exercises 4.10 and 4.14)

$$\{\pi_1, \pi_2\} = \frac{eB_0}{c}, \qquad \{\sigma_1, \sigma_2\} = -\frac{eB_0}{c}, \qquad \{\pi_i, \sigma_j\} = 0.$$

Then, making use again of the bilinearity and skewsymmetry of the Poisson bracket, we have, for instance,

$$\{Q_1, P_1\} = \frac{c}{eB_0}\{\pi_1\cos\omega_c t - \pi_2\sin\omega_c t, \pi_2\cos\omega_c t + \pi_1\sin\omega_c t\} = \frac{c}{eB_0}\{\pi_1, \pi_2\} = 1,$$

and

$$\{Q_2, P_2\} = \frac{c}{eB_0}\{\sigma_2, \sigma_1\} = 1,$$

and so on, thus proving that the coordinate transformation is canonical.

Since Q_1, Q_2, P_1, P_2 are canonical coordinates, the equations of motion must be expressible in the form (5.35), for some function K. If we take

$$H = \frac{1}{2m}\left[\left(p_x + \frac{eB_0}{2c}y\right)^2 + \left(p_y - \frac{eB_0}{2c}x\right)^2\right],$$

which corresponds to a charged particle of mass m and electric charge e in a uniform magnetic field B_0, then Q_1, Q_2, P_1, and P_2 are constants of motion. In fact,

$$H = \frac{1}{2m}({\pi_1}^2 + {\pi_2}^2)$$

and, for instance,

$$\begin{aligned}
&\frac{\partial Q_1}{\partial t} + \{Q_1, H\} \\
&= \frac{c}{eB_0}\left[\frac{\partial(\pi_1 \cos\omega_c t - \pi_2 \sin\omega_c t)}{\partial t} + \frac{1}{2m}\{\pi_1 \cos\omega_c t - \pi_2 \sin\omega_c t, {\pi_1}^2 + {\pi_2}^2\}\right] \\
&= \frac{c}{eB_0}\left[-\omega_c(\pi_1 \sin\omega_c t + \pi_2 \cos\omega_c t) + \frac{1}{m}\{\pi_1, \pi_2\}\pi_2 \cos\omega_c t\right. \\
&\qquad\left. - \frac{1}{m}\{\pi_2, \pi_1\}\pi_1 \sin\omega_c t\right] \\
&= 0.
\end{aligned}$$

In a similar manner one can show that Q_2, P_1, and P_2 are constants of motion and, therefore, we can choose $K = 0$.

Exercise 5.46. Show that for any coordinate transformation $Q_i = Q_i(q_j, p_j, t)$, $P_i = P_i(q_j, p_j, t)$ (that is, not necessarily canonical),

$$[x_\alpha, x_\beta]_{Q,P}\ \{x_\gamma, x_\beta\}_{Q,P} = \delta_{\alpha\gamma}.$$

Invariance of the Poisson Bracket Under Canonical Transformations

As in the case where the number of degrees of freedom is 1, for an arbitrary number of degrees of freedom, the canonical transformations leave invariant the Poisson bracket in the sense that, for an arbitrary pair of differentiable functions, f, g, defined on the extended phase space,

$$\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} = \frac{\partial f}{\partial Q_i}\frac{\partial g}{\partial P_i} - \frac{\partial f}{\partial P_i}\frac{\partial g}{\partial Q_i}. \tag{5.112}$$

In order to demonstrate the validity of (5.112), making use of the definition of the Poisson bracket and the chain rule, we obtain the *identity*

$$\begin{aligned}\{f, g\} &= \frac{\partial f}{\partial q_i}\left(\frac{\partial g}{\partial Q_j}\frac{\partial Q_j}{\partial p_i} + \frac{\partial g}{\partial P_j}\frac{\partial P_j}{\partial p_i}\right) - \frac{\partial f}{\partial p_i}\left(\frac{\partial g}{\partial Q_j}\frac{\partial Q_j}{\partial q_i} + \frac{\partial g}{\partial P_j}\frac{\partial P_j}{\partial q_i}\right)\\ &= \{f, Q_j\}\frac{\partial g}{\partial Q_j} + \{f, P_j\}\frac{\partial g}{\partial P_j},\end{aligned} \tag{5.113}$$

hence, making use of the antisymmetry of the Poisson bracket and (5.113) again,

$$\begin{aligned}\{f, g\} &= -\{Q_j, f\}\frac{\partial g}{\partial Q_j} - \{P_j, f\}\frac{\partial g}{\partial P_j}\\ &= -\left(\{Q_j, Q_k\}\frac{\partial f}{\partial Q_k} + \{Q_j, P_k\}\frac{\partial f}{\partial P_k}\right)\frac{\partial g}{\partial Q_j}\\ &\quad - \left(\{P_j, Q_k\}\frac{\partial f}{\partial Q_k} + \{P_j, P_k\}\frac{\partial f}{\partial P_k}\right)\frac{\partial g}{\partial P_j}\\ &= \{Q_k, Q_j\}\frac{\partial f}{\partial Q_k}\frac{\partial g}{\partial Q_j} + \{P_k, Q_j\}\frac{\partial f}{\partial P_k}\frac{\partial g}{\partial Q_j}\\ &\quad + \{Q_k, P_j\}\frac{\partial f}{\partial Q_k}\frac{\partial g}{\partial P_j} + \{P_k, P_j\}\frac{\partial f}{\partial P_k}\frac{\partial g}{\partial P_j}.\end{aligned} \tag{5.114}$$

Therefore, if Q_i, P_i are related to q_i, p_i, and t by means of a canonical transformation (that is, Equations (5.104) hold), from (5.114) we see that

$$\{f, g\} = \frac{\partial f}{\partial Q_j}\frac{\partial g}{\partial P_j} - \frac{\partial f}{\partial P_j}\frac{\partial g}{\partial Q_j},$$

as was to be shown.

Conversely, if (5.112) holds for any pair of differentiable functions, f, g, then, in particular,

$$\{Q_j, Q_k\} = \frac{\partial Q_j}{\partial Q_i}\frac{\partial Q_k}{\partial P_i} - \frac{\partial Q_j}{\partial P_i}\frac{\partial Q_k}{\partial Q_i} = 0,$$

and, in a similar way, one obtains $\{P_j, P_k\} = 0$, $\{Q_j, P_k\} = \delta_{jk}$, which, according to Proposition 5.44, means that Q_i and P_i are given in terms of q_i, p_i, and possibly of t, by a canonical transformation.

Hence, the canonical transformations can be defined as the coordinate transformations in the extended phase space that preserve the Poisson bracket.

The Liouville Theorem on the Invariance of the Volume Element of the Phase Space

The *volume element* of the phase space, $\mathrm{d}q_1 \cdots \mathrm{d}q_n \mathrm{d}p_1 \cdots \mathrm{d}p_n$, is invariant under canonical transformations, which amounts to say that the Jacobian of a canonical transformation is equal to 1

$$\frac{\partial(Q_i, P_i)}{\partial(q_i, p_i)} = 1$$

(this fact was already established in (5.11), when the number of degrees of freedom is 1). This result is usually known as the *Liouville theorem*. (There is another, lesser-known, Liouville theorem related to the Hamilton–Jacobi equation, which will be considered in Section 6.2.)

In order to demonstrate this theorem for an arbitrary value of n, we shall assume that the canonical transformation is such that q_i, P_i, and t are functionally independent (other cases are treated in a similar way) and, therefore, the transformation can be expressed in terms of a generating function $F_2(q_i, P_i, t)$ by Equations (5.55), and the three sets (q_i, p_i, t), (Q_i, P_i, t), and (q_i, P_i, t), can be used as coordinates of the extended phase space. Then, by virtue of the chain rule, we have

$$\frac{\partial(Q_i, P_i)}{\partial(q_i, p_i)} = \frac{\partial(Q_i, P_i)}{\partial(q_i, P_i)} \frac{\partial(q_i, P_i)}{\partial(q_i, p_i)}$$

(note that this is a relation between $2n \times 2n$ determinants, the subscripts in the coordinates only indicate that there are n variables of each type) and, similarly,

$$\frac{\partial(q_i, P_i)}{\partial(q_i, p_i)} \frac{\partial(q_i, p_i)}{\partial(q_i, P_i)} = 1,$$

hence,

$$\frac{\partial(Q_i, P_i)}{\partial(q_i, p_i)} = \frac{\partial(Q_i, P_i)}{\partial(q_i, P_i)} \left[\frac{\partial(q_i, p_i)}{\partial(q_i, P_i)} \right]^{-1}. \tag{5.115}$$

Then, making use of the definition of the Jacobian and Equations (5.55), we have

$$\frac{\partial(Q_i, P_i)}{\partial(q_i, P_i)} = \frac{\partial(Q_1, \ldots, Q_n)}{\partial(q_1, \ldots, q_n)} = \det\left(\frac{\partial^2 F_2}{\partial q_i \partial P_j}\right)$$

and

$$\frac{\partial(q_i, p_i)}{\partial(q_i, P_i)} = \frac{\partial(p_1, \ldots, p_n)}{\partial(P_1, \ldots, P_n)} = \det\left(\frac{\partial^2 F_2}{\partial P_i \partial q_j}\right).$$

Substituting these last two expressions into (5.115) one concludes that the Jacobian of a canonical transformation is equal to 1.

It may be remarked that when the number of degrees of freedom is equal to 1, a coordinate transformation $Q = Q(q, p, t)$, $P = P(q, p, t)$ is canonical if and only if the Jacobian $\partial(Q, P)/\partial(q, p)$ is equal to 1. However, when the number of degrees of freedom is greater than 1, there is no such equivalence; there exist coordinate transformations with Jacobian equal to 1 that are not canonical.

Exercise 5.47. Find an example of a noncanonical coordinate transformation with Jacobian equal to 1.

Comparison with the Standard Approach
As shown in Section 4.2.1, the curve $C(t) = (q_i(t), p_i(t), t)$ in the extended phase space is a solution of the Hamilton equations if and only if the integral

$$I(C) = \int_C (p_i \mathrm{d}q_i - H\mathrm{d}t) = \int_{t_0}^{t_1} \left[p_i(t)\frac{\mathrm{d}q_i(t)}{\mathrm{d}t} - H(q_i(t), p_i(t), t) \right] \mathrm{d}t$$

has a stationary value on this curve, in the space of curves in the extended phase space that share the same endpoints with C. If Q_i, P_i, t is another coordinate system of the extended phase space such that the Hamilton equations (5.35) are equivalent to (5.34), then the functional

$$\tilde{I}(C) = \int_C (P_i \mathrm{d}Q_i - K\mathrm{d}t) = \int_{t_0}^{t_1} \left[P_i(t)\frac{\mathrm{d}Q_i(t)}{\mathrm{d}t} - K(Q_i(t), P_i(t), t) \right] \mathrm{d}t$$

must possess the same stationary points as $I(C)$. An error present in many textbooks on analytical mechanics is the conclusion that the integrands of $I(C)$ and $\tilde{I}(C)$ must differ at most by a constant factor and an exact differential, that is,

$$p_i \mathrm{d}q_i - H\mathrm{d}t = \lambda(P_i \mathrm{d}Q_i - K\mathrm{d}t) + \mathrm{d}F, \tag{5.116}$$

where λ is a constant and F is a real-valued function defined on the extended phase space. In order to show that this conclusion is wrong, it is enough to consider Example 5.2. A straightforward computation gives

$$P\mathrm{d}Q - K\mathrm{d}t = \frac{m\omega(p\mathrm{d}q - q\mathrm{d}p - 2H\mathrm{d}t)}{\sqrt{p^2 + m^2\omega^2 q^2}} = \frac{2m\omega(p\mathrm{d}q - H\mathrm{d}t) - \mathrm{d}(m\omega pq)}{\sqrt{p^2 + m^2\omega^2 q^2}},$$

which is not of the form (5.116). In fact, if the claim was true, the canonoid transformations would not exist.

Another counterexample is given by a particle in a uniform gravitational field. The standard Hamiltonian in Cartesian coordinates is

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy$$

and, therefore,

$$p_i \mathrm{d}q_i - H\mathrm{d}t = p_x \mathrm{d}x + p_y \mathrm{d}y - \left(\frac{{p_x}^2 + {p_y}^2}{2m} + mgy \right) \mathrm{d}t.$$

On the other hand, in terms of the coordinates

$$Q_1 = x, \qquad Q_2 = y, \qquad P_1 = p_y, \qquad P_2 = p_x,$$

the equations of motion also have the form of the Hamilton equations with

$$K = \frac{P_1 P_2}{m} + mgQ_1$$

so that

$$P_i \mathrm{d}Q_i - K\mathrm{d}t = p_y \mathrm{d}x + p_x \mathrm{d}y - \left(\frac{p_x p_y}{m} + mgx \right) \mathrm{d}t$$

and we can see that $p_i \mathrm{d}q_i - H\mathrm{d}t$ and $P_i \mathrm{d}Q_i - K\mathrm{d}t$ are not related as in (5.116).

As shown above, the converse is true, that is, if Equation (5.116) holds then the Hamilton equations (5.35) are equivalent to (5.34).

This is analogous to the fact that the point $x = 0$ is a minimum of the functions $f(x) = x^4$ and $g(x) = \cosh x$, but these functions are not the same nor are linked by a relation of the form $f(x) = \lambda g(x) + \mu$, where λ and μ are constants. Of course, if $f(x) = \lambda g(x) + \mu$, with λ and μ constant, then the critical points of f and g coincide.

5.3 One-Parameter Groups of Canonical Transformations

In this section we shall consider one-parameter families of canonical transformations, that is, canonical transformations of the form

$$Q_i = Q_i(q_j, p_j, t, s), \qquad P_i = P_i(q_j, p_j, t, s), \tag{5.117}$$

where s is a real parameter that can take values in some open interval containing 0 (in many cases s can take all real values) and we shall assume that for $s = 0$ the transformation (5.117) reduces to the identity, that is,

$$Q_i(q_j, p_j, t, 0) = q_i, \qquad P_i(q_j, p_j, t, 0) = p_i. \tag{5.118}$$

As we shall see in the following section, the one-parameter families of canonical transformations are especially important, among other reasons, because of their relation with the constants of motion.

Example 5.48. Let us consider the one-parameter family of transformations given by

$$Q = q\mathrm{e}^{s} - \frac{tp}{m}(\mathrm{e}^{s} - \mathrm{e}^{-s}), \qquad P = p\mathrm{e}^{-s}, \tag{5.119}$$

where m is a constant and s is a parameter that can take any real value. We can readily verify that when $s = 0$, the right-hand sides of Equations (5.119) reduce to q and p, respectively, and that, for each *fixed* value of s, this transformation is canonical, in fact, one readily sees that $\{Q, P\} = 1$, or that

$$\begin{aligned} p\mathrm{d}q - H\mathrm{d}t - (P\mathrm{d}Q - K_s\mathrm{d}t) &= p\mathrm{d}q - p\mathrm{e}^{-s}\left[\mathrm{e}^{s}\mathrm{d}q - \frac{t}{m}(\mathrm{e}^{s} - \mathrm{e}^{-s})\mathrm{d}p - \frac{p}{m}(\mathrm{e}^{s} - \mathrm{e}^{-s})\mathrm{d}t\right] + (K_s - H)\mathrm{d}t \\ &= \mathrm{d}\left[\frac{tp^2}{2m}(1 - \mathrm{e}^{-2s})\right] + \left[K_s - H + \frac{p^2}{2m}(1 - \mathrm{e}^{-2s})\right]\mathrm{d}t, \end{aligned}$$

which shows that the transformation is canonical, and that the new Hamiltonian must be given by

$$K_s = H - \frac{p^2}{2m}(1 - \mathrm{e}^{-2s}) + f(t, s), \tag{5.120}$$

where $f(t, s)$ is an arbitrary function of t and s only, and we have written K_s instead of K in order to emphasize the dependence of the new Hamiltonian on the parameter s.

In most textbooks, the one-parameter families of canonical transformations are given in "infinitesimal form," by expressing the transformations (5.117) to first order in s. For that reason, such transformations are usually called *infinitesimal canonical transformations*. For example, the transformations (5.119) would be expressed as

$$Q = q + qs - \frac{2tps}{m}, \qquad P = p - ps.$$

(Note that, in this case, $\{Q, P\} = 1 - s^2$, which is equal to 1 "to first order in s.") A more precise expression would be

$$Q = q + qs - \frac{2tps}{m} + O(s^2), \qquad P = p - ps + O(s^2),$$

but it is seldom used. We shall avoid the use of infinitesimal canonical transformations, considering strict canonical transformations only.

As we shall show now, a one-parameter family of canonical transformations possesses a generating function.

Proposition 5.49. *If* $Q_i = Q_i(q_j, p_j, t, s)$, $P_i = P_i(q_j, p_j, t, s)$, *is a one-parameter family of canonical transformations then there exists locally a function* $G(q_i, p_i, t)$ *(defined up to an additive function of* t *only) such that*

$$\left.\frac{\partial Q_i}{\partial s}\right|_{s=0} = \frac{\partial G}{\partial p_i}, \qquad \left.\frac{\partial P_i}{\partial s}\right|_{s=0} = -\frac{\partial G}{\partial q_i}. \tag{5.121}$$

Proof. As we have seen in the preceding sections, if, for each value of s, $Q_i = Q_i(q_j, p_j, t, s)$, $P_i = P_i(q_j, p_j, t, s)$, is a canonical transformation, then there exists (locally) a function F_1, which may depend parametrically on s, such that [see Equation (5.46)]

$$p_i \mathrm{d}q_i - H\mathrm{d}t - (P_i \mathrm{d}Q_i - K_s \mathrm{d}t) = \mathrm{d}F_1.$$

Taking the partial derivative with respect to s (with q_i, p_i and t fixed), at $s = 0$, of both sides of the last equation we obtain

$$-\left.\frac{\partial P_i}{\partial s}\right|_{s=0} \mathrm{d}q_i - p_i\, \mathrm{d}\left(\left.\frac{\partial Q_i}{\partial s}\right|_{s=0}\right) + \left.\frac{\partial K_s}{\partial s}\right|_{s=0} \mathrm{d}t = \mathrm{d}\left(\left.\frac{\partial F_1}{\partial s}\right|_{s=0}\right)$$

or, equivalently,

$$-\left.\frac{\partial P_i}{\partial s}\right|_{s=0} \mathrm{d}q_i + \left.\frac{\partial Q_i}{\partial s}\right|_{s=0} \mathrm{d}p_i + \left.\frac{\partial K_s}{\partial s}\right|_{s=0} \mathrm{d}t = \mathrm{d}\left(\left.\frac{\partial F_1}{\partial s}\right|_{s=0} + p_i \left.\frac{\partial Q_i}{\partial s}\right|_{s=0}\right).$$

Letting

$$G \equiv \left.\frac{\partial F_1}{\partial s}\right|_{s=0} + p_i \left.\frac{\partial Q_i}{\partial s}\right|_{s=0} \tag{5.122}$$

we obtain Equations (5.121) and, in addition,

$$\left.\frac{\partial K_s}{\partial s}\right|_{s=0} = \frac{\partial G}{\partial t}. \tag{5.123}$$

□

As we know, for a given canonical transformation, the difference $K - H$ is determined up to an additive function of t (and s, in the case of a one-parameter family of canonical transformations); the choice of this function affects F_1 and G (see the examples below).

For instance, in the case of the one-parameter family of canonical transformations given in Example 5.48, directly from Equations (5.119) we obtain

$$\left.\frac{\partial Q}{\partial s}\right|_{s=0} = q - \frac{2tp}{m}, \qquad \left.\frac{\partial P}{\partial s}\right|_{s=0} = -p, \tag{5.124}$$

which can be written in the form (5.121) with

$$G = pq - \frac{tp^2}{m} + \chi(t), \tag{5.125}$$

where $\chi(t)$ is a function of t only, which can be determined by means of (5.123) [or from (5.122)] if the new Hamiltonian has been specified. In fact, substituting (5.120) and (5.125) into (5.123) one obtains

$$\left.\frac{\partial f}{\partial s}\right|_{s=0} = \chi'(t),$$

which shows that for a choice of $f(t,s)$ in (5.120), χ is determined up to an additive trivial constant, which can be taken equal to zero.

The family of canonical transformations (5.119) is a *one-parameter group* of transformations in the following sense: if, for each $s \in \mathbb{R}$, we define the map $\phi_s : \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\phi_s(q,p,t) = \left(q\mathrm{e}^s - \frac{tp}{m}(\mathrm{e}^s - \mathrm{e}^{-s}),\, p\mathrm{e}^{-s},\, t\right)$$

then, by means of a straightforward but somewhat lengthy computation, one verifies that

$$\phi_{s'} \circ \phi_s = \phi_{s'+s} \tag{5.126}$$

for all $s', s \in \mathbb{R}$, and ϕ_0 is the identity map of $\mathbb{R}^3$. The validity of (5.126) is a consequence of the fact that if the partial derivatives with respect to s of $Q(q,p,t,s)$ and $P(q,p,t,s)$ are written in terms of Q, P, t, and s, the resulting expressions *do not contain* s; in other words, $Q(q,p,t,s)$ and $P(q,p,t,s)$, considered as functions of s only, satisfy an *autonomous* system of first-order ODEs. Indeed, from Equations (5.119), we obtain

$$\begin{aligned}
\frac{\partial Q}{\partial s} &= q\,\mathrm{e}^s - \frac{tp}{m}(\mathrm{e}^s + \mathrm{e}^{-s}) = Q - \frac{2tP}{m},\\
\frac{\partial P}{\partial s} &= -p\,\mathrm{e}^{-s} = -P.
\end{aligned}$$

(Cf. Equations (5.124).)

As a converse of Proposition 5.49, *any* differentiable function, $G(q_i, p_i, t)$, defines a (possibly local, see below) one-parameter group of canonical transformations, in such a way that Equations (5.121) hold, in the following manner. Equation (5.121) can be written as

$$\left.\frac{\partial Q_i}{\partial s}\right|_{s=0} = \left.\frac{\partial G(Q_j, P_j, t)}{\partial P_i}\right|_{s=0}, \qquad \left.\frac{\partial P_i}{\partial s}\right|_{s=0} = -\left.\frac{\partial G(Q_j, P_j, t)}{\partial Q_i}\right|_{s=0} \tag{5.127}$$

and, in order to have a well-defined rule to construct a family of coordinate transformations, we demand that Equations (5.127) hold for all values of s (not only for $s = 0$), that is

$$\frac{\partial Q_i}{\partial s} = \frac{\partial G(Q_j, P_j, t)}{\partial P_i}, \qquad \frac{\partial P_i}{\partial s} = -\frac{\partial G(Q_j, P_j, t)}{\partial Q_i}. \tag{5.128}$$

In this way, Equations (5.128) constitute an *autonomous* system of $2n$ ODEs for Q_i and P_i as functions of s, with the initial conditions (5.118).

Denoting by $Q_i(q_j, p_j, t, s)$, $P_i(q_j, p_j, t, s)$, the solution of the system (5.128) such that

$$Q_i(q_j, p_j, t, 0) = q_i, \qquad P_i(q_j, p_j, t, 0) = p_i,$$

and following a procedure analogous to that employed in (5.96), one finds that the Lagrange brackets $[q_i, q_j]$, $[q_i, p_j]$, and $[p_i, p_j]$ do not depend on s and, since for $s = 0$ these brackets have the values 0, δ_{ij}, and 0, respectively, one concludes that the transformation is canonical for all values of s.

As a consequence of the uniqueness of the solutions of the system of equations (5.128), the solutions of the system (5.128) correspond to a, possibly local, *one-parameter group* of canonical transformations (see, e.g., Crampin and Pirani [5, Chap. 3]).

Example 5.50. In order to find the one-parameter group of canonical transformations generated by the function

$$G(q, p, t) = qp - \frac{3}{2}gt^2 p - \frac{tp^2}{m} + mgtq - \frac{1}{2}mg^2t^3, \tag{5.129}$$

where m and g are constants, we substitute this expression into Equations (5.128), which yields the autonomous system of ODEs

$$\frac{\mathrm{d}Q}{\mathrm{d}s} = Q - \frac{3}{2}gt^2 - \frac{2tP}{m}, \tag{5.130}$$

$$\frac{\mathrm{d}P}{\mathrm{d}s} = -P - mgt, \tag{5.131}$$

where we have written the derivatives with respect to s as ordinary derivatives, for convenience, taking into account that q, p, and t enter in (5.128) as parameters only, which specify the initial conditions. (Note that the last term in (5.129) is a function of t only, which does not appear in Equations (5.130) and (5.131).)

Equation (5.131) is a separable equation, and one readily obtains the solution

$$P = (p + mgt)\,\mathrm{e}^{-s} - mgt, \tag{5.132}$$

so that, for $s = 0$, $P = p$. Substituting this result into (5.130) one obtains the linear first-order ODE

$$\frac{\mathrm{d}Q}{\mathrm{d}s} - Q = \frac{1}{2}gt^2 - 2\left(\frac{tp}{m} + gt^2\right)\mathrm{e}^{-s},$$

whose solution is

$$Q = -\frac{1}{2}gt^2 + \left(\frac{tp}{m} + gt^2\right)\mathrm{e}^{-s} + \left(q - \frac{1}{2}gt^2 - \frac{tp}{m}\right)\mathrm{e}^{s}. \tag{5.133}$$

Alternatively, the expression for Q can be readily obtained from that for P using the fact that, in all cases,

$$G(q_i, p_i, t) = G\big(Q_i(q_j, p_j, t, s), P_i(q_j, p_j, t, s), t\big), \tag{5.134}$$

for all s, since, by virtue of Equations (5.128), the right-hand side of (5.134) does not depend on s.

As we have seen in the preceding sections, for a given canonical transformation the difference $K_s - H$ is determined up to an additive function of t only. In the present case, substituting (5.132) and (5.133) into Equations (5.12) we obtain

$$K_s - H = \frac{p^2}{2m}(\mathrm{e}^{-2s} - 1) + mgq(\mathrm{e}^{s} - 1) + gtp(\mathrm{e}^{-2s} - \mathrm{e}^{s}) + f(t, s), \tag{5.135}$$

where now $f(t, s)$ is a function of t and s only. Since in this example G is given from the start, the function $f(t, s)$ is not completely arbitrary; in order to satisfy Equation (5.123), $f(t, s)$ must be such that

$$\left.\frac{\partial f}{\partial s}\right|_{s=0} = -\frac{3}{2}mg^2t^2,$$

but, clearly, this condition does not determine the function $f(t, s)$. (See, however, Exercise 5.73.)

Exercise 5.51. Show the validity of (5.134).

Example 5.52 (A local one-parameter group of canonical transformations). In all the examples presented so far, the parameter s can take all real values, but this is not always the case. If we consider the generating function

$$G(q, p, t) = q^2 p,$$

the corresponding system of equations is [see Equations (5.128)]

$$\frac{\mathrm{d}Q}{\mathrm{d}s} = Q^2, \qquad \frac{\mathrm{d}P}{\mathrm{d}s} = -2QP,$$

and the family of canonical transformations generated by G is given by

$$Q = \frac{q}{1 - sq}, \qquad P = (1 - sq)^2 p,$$

which, if $q \neq 0$, are defined only for $s \neq 1/q$. Thus, we obtain a *local* one-parameter group of transformations, which means that $\phi_s(q, p, t) \equiv \big(\frac{q}{1-sq}, (1 - sq)^2 p, t\big)$ is not defined for all values of the parameter s, but satisfies Equation (5.126) whenever both sides of the equation are defined. Note that, in accordance with (5.134), $q^2 p = Q^2 P$.

It should be clear that in many cases the solution of the system of equations (5.128) may be quite difficult. However, for some purposes, the knowledge of the derivatives $\partial Q_i/\partial s$, $\partial P_i/\partial s$ is enough.

Example 5.53 (Passive rotations in the three-dimensional Euclidean space). Let us assume that the Cartesian coordinates and momenta of a particle moving in the three-dimensional Euclidean space transform according to

$$\begin{aligned} Q_1 &= q_1 \cos s + q_2 \sin s, \quad Q_2 = -q_1 \sin s + q_2 \cos s, \quad Q_3 = q_3, \\ P_1 &= p_1 \cos s + p_2 \sin s, \quad P_2 = -p_1 \sin s + p_2 \cos s, \quad P_3 = p_3, \end{aligned} \tag{5.136}$$

if the Cartesian axes (Q_1, Q_2, Q_3) are obtained from the Cartesian axes (q_1, q_2, q_3) by means of a rotation through an angle s, about the q_3-axis. Computing the partial derivatives

$$\left.\frac{\partial Q_1}{\partial s}\right|_{s=0} = q_2, \quad \left.\frac{\partial Q_2}{\partial s}\right|_{s=0} = -q_1, \qquad \left.\frac{\partial Q_3}{\partial s}\right|_{s=0} = 0,$$
$$\left.\frac{\partial P_1}{\partial s}\right|_{s=0} = p_2, \quad \left.\frac{\partial P_2}{\partial s}\right|_{s=0} = -p_1, \qquad \left.\frac{\partial P_3}{\partial s}\right|_{s=0} = 0,$$

and comparing with Equations (5.121) we find that the generating function of these rotations is

$$G = -q_1 p_2 + q_2 p_1 + \chi(t),$$

where $\chi(t)$ is some function of t only, which *is not* determined by the coordinate transformation (see also Example 5.56). The function $\chi(t)$ need not be equal to zero, and different expressions for $\chi(t)$ will produce different generating functions of rotations. (See Exercise 5.61.)

Equation (5.128) is somewhat similar to the Hamilton equations and, for that reason, it is usually stated that the Hamiltonian is the generating function of the time evolution (with the parameter s of the transformations being the time). There are some differences, however; in the derivation of (5.121) it was assumed that the parameter s is not related to the coordinates or the time and a generating function G does not depend on the parameter s, while a Hamiltonian may depend on the time.

We end this section remarking that a given one-parameter family of canonical transformations defines its generating function up to an additive function of t only, but a real-valued function, G, defined on the extended phase space completely determines a, possibly local, one-parameter group of canonical transformations.

5.4 Symmetries of the Hamiltonian and Constants of Motion

As we have seen in Sections 5.1 and 5.2, given a canonical transformation, the new Hamiltonian, K, is defined up to an additive function of t only [cf. also Equation (5.120)]. We shall say that a Hamiltonian $H(q_i, p_i, t)$ is *invariant* under a canonical transformation $Q_i = Q_i(q_j, p_j, t)$, $P_i = P_i(q_j, p_j, t)$ (which may also depend on some parameters) if the new Hamiltonian K can be *chosen* in such a way that

$$K(q_i, p_i, t) = H\big(Q_i(q_j, p_j, t), P_i(q_j, p_j, t), t\big). \tag{5.137}$$

This amounts to say that the Hamiltonian $H(q_i, p_i, t)$ is invariant under the canonical transformation $Q_i = Q_i(q_j, p_j, t)$, $P_i = P_i(q_j, p_j, t)$ if the equations of motion expressed in terms of Q_i and P_i have exactly the same form as the equations of motion written in terms of q_i and p_i. (Note that we restrict ourselves to canonical transformations.)

The invariance condition (5.137) can be expressed in a more precise manner making use of the mapping $\phi : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n+1}$, defined by

$$\phi(q_1, \ldots, q_n, p_1, \ldots, p_n, t) \equiv (Q_1, \ldots, Q_n, P_1, \ldots, P_n, t),$$

so that Equation (5.137) is equivalent to

$$K = H \circ \phi. \tag{5.138}$$

For instance, the Hamiltonian

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2} q^2$$

is invariant under the canonical transformation

$$Q = \frac{p}{m\omega}, \qquad P = -m\omega q, \tag{5.139}$$

since

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{p}{m}, \qquad \dot{p} = -\frac{\partial H}{\partial q} = -m\omega^2 q \tag{5.140}$$

and, therefore, combining (5.139) and (5.140),

$$\dot{Q} = \frac{\dot{p}}{m\omega} = -\omega q = \frac{P}{m}, \qquad \dot{P} = -m\omega \dot{q} = -\omega p = -m\omega^2 Q,$$

which are of the same *form* as Equations (5.140). Alternatively, since the transformation (5.139) does not involve the time, $K - H = f(t)$, where $f(t)$ is a function of t only. Hence,

$$\begin{aligned} K(q, p, t) &= \frac{p^2}{2m} + \frac{m\omega^2}{2} q^2 + f(t) \\ &= \frac{(m\omega Q)^2}{2m} + \frac{m\omega^2}{2} \left(-\frac{P}{m\omega} \right)^2 + f(t) \\ &= \frac{P^2}{2m} + \frac{m\omega^2}{2} Q^2 + f(t) \\ &= H(Q, P, t) + f(t). \end{aligned}$$

Choosing $f(t) = 0$, we see that condition (5.137) is indeed satisfied.

Example 5.54. Let us consider the standard Hamiltonian for a free particle

$$H = \frac{p^2}{2m}.$$

The corresponding Hamilton equations are

$$\frac{dq}{dt} = \frac{p}{m}, \qquad \frac{dp}{dt} = 0. \tag{5.141}$$

The form of the second of these equations does not change under the scaling transformation

$$P = \lambda p,$$

for any nonzero constant λ. Then, the first equation in (5.141) maintains its form if we take $Q = \lambda q$, but this is not a canonical transformation if $\lambda \neq \pm 1$ ($\{Q, P\} = \{\lambda q, \lambda p\} = \lambda^2$). However, we have a canonical transformation if

$$Q = \frac{q}{\lambda} + F(p, t),$$

where $F(p, t)$ is an arbitrary function of p and t only, which can be chosen in such a way that the form of the first equation in (5.141) is invariant, that is, $dQ/dt = P/m$. In fact, this last condition amounts to

$$\frac{1}{\lambda}\frac{dq}{dt} + \frac{\partial F}{\partial p}\frac{dp}{dt} + \frac{\partial F}{\partial t} = \frac{\lambda p}{m},$$

which, with the aid of the original equations of motion (5.141), is equivalent to

$$\frac{1}{\lambda}\frac{p}{m} + \frac{\partial F}{\partial t} = \frac{\lambda p}{m}.$$

Thus, choosing

$$F(p, t) = \left(\lambda - \frac{1}{\lambda}\right)\frac{tp}{m}$$

we conclude that the Hamiltonian $H(q, p, t) = p^2/2m$ is invariant under the canonical transformation

$$Q = \frac{q}{\lambda} + \left(\lambda - \frac{1}{\lambda}\right)\frac{tp}{m}, \qquad P = \lambda p,$$

for any nonzero value of λ. (These are precisely the transformations (5.119), if we express the parameter λ as e^{-s}.)

We shall say that a given Hamiltonian, H, is invariant under a one-parameter family of canonical transformations, ϕ_s, if H is invariant under ϕ_s for each value of s.

As we shall show, the invariance of a Hamiltonian under a one-parameter family of canonical transformations implies the existence of a constant of motion, and any constant of motion is the generating function of a, possibly local, one-parameter group of canonical transformations that leave the Hamiltonian invariant. As we shall see in Section 5.5 there are also constants of motion associated with discrete canonoid transformations, but these can be trivial.

Proposition 5.55. *If the Hamiltonian H is invariant under a one-parameter family of canonical transformations, then its generating function, G, is a constant of motion.*

Proof. Considering a one-parameter family of canonical transformations with generating function G, making use of Equations (5.123) and (5.121), and the chain rule we get

$$\begin{aligned}\frac{\partial G}{\partial t} + \frac{\partial G}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial G}{\partial p_i}\frac{\partial H}{\partial q_i} \\ &= \left.\frac{\partial K_s}{\partial s}\right|_{s=0} - \left.\frac{\partial P_i}{\partial s}\right|_{s=0}\frac{\partial H}{\partial p_i} - \left.\frac{\partial Q_i}{\partial s}\right|_{s=0}\frac{\partial H}{\partial q_i} \\ &= \frac{\partial}{\partial s}\big[K_s - H\big(Q_i(q_i, p_i, t, s), P_i(q_i, p_i, t, s), t\big)\big]\Big|_{s=0} \end{aligned} \tag{5.142}$$

which is equal to zero if H is invariant under these transformations [see (5.137)], thus showing that G is a constant of motion. □

Example 5.56 (A group of active translations). It can be readily seen that

$$Q = q + s, \qquad P = p, \tag{5.143}$$

is a one-parameter family of canonical transformations (in fact, a one-parameter *group* of canonical transformations). Making use of Equations (5.121) one finds that the generating function of this group is of the form

$$G = p + \chi(t),$$

where $\chi(t)$ is a function of t only. Since the transformation (5.143) does not involve the time,

$$K_s = H + f(t, s), \tag{5.144}$$

where $f(t, s)$ is a function of t and s only [see Equations (5.12)].

The Hamiltonian

$$H(q, p, t) = \frac{p^2}{2m} + mgq$$

is invariant under the group of canonical transformations (5.143) since, from (5.143) and (5.144), we have

$$
\begin{aligned}
K_s(q, p, t) &= \frac{p^2}{2m} + mgq + f(t, s) \\
&= \frac{P^2}{2m} + mgQ - mgs + f(t, s) \\
&= H(Q, P, t) - mgs + f(t, s)
\end{aligned}
$$

and, therefore, in order to satisfy the invariance condition (5.137), $f(t, s)$ must be chosen as mgs. Then, from (5.123) and the expressions derived above we have

$$
\frac{\partial G}{\partial t} = \left.\frac{\partial K_s}{\partial s}\right|_{s=0} = \left.\frac{\partial f}{\partial s}\right|_{s=0} = mg
$$

and therefore $G = p + mgt$, which is, indeed, a constant of motion.

It may be remarked that it is usually stated (in the context of classical mechanics and in the context of quantum mechanics) that the canonical momentum, p, is *the* generating function (or *the* "infinitesimal generator," as is usually called) of the active translations; however, as we have shown, there are an infinite number of different generating functions of the translations (5.143), depending on the choice of the function $\chi(t)$. (For the passive translations defined by $Q = q - s$, $P = p$, the generating function is of the form $G = -p + \chi(t)$.) (See also Exercise 5.58.)

Since in this example we are considering a particle in a *uniform* gravitational field, it is to be expected that its Hamiltonian be invariant under translations.

Exercise 5.57. Show that the Hamiltonian

$$
H = \frac{p^2}{2m} - F(t)q,
$$

where $F(t)$ is a function of t only, is invariant under the group of translations (5.143) and find the constant of motion associated with this invariance.

Exercise 5.58 (Passive translations in a uniform magnetic field). Show that the Hamiltonian (5.48), for a charged particle in a uniform magnetic field, is invariant under (passive) translations in an *arbitrary* direction (a_1, a_2, a_3), defined by

$$
\begin{aligned}
&Q_1 = x - sa_1, \qquad Q_2 = y - sa_2, \qquad Q_3 = z - sa_3, \\
&P_1 = p_x + \frac{eB_0}{2c} sa_2, \quad P_2 = p_y - \frac{eB_0}{2c} sa_1, \quad P_3 = p_z
\end{aligned}
\tag{5.145}
$$

(see Example 5.36), and find the *three* constants of motion associated with these symmetries. (Cf. Equations (5.50).) (The vector (sa_1, sa_2, sa_3) is the position vector of the origin of the Cartesian coordinates (Q_1, Q_2, Q_3) with respect to the origin of the Cartesian coordinates (x, y, z).) (Cf. also Example 1.19.)

(It should be kept in mind that when there is a magnetic field present, the canonical momenta, p_i, are gauge dependent [see (5.71)]. The Hamilton equations applied to the Hamiltonian (5.48) yield

$$m\dot{x} = p_x + \frac{eB_0}{2c}y, \qquad m\dot{y} = p_y - \frac{eB_0}{2c}x, \qquad m\dot{z} = p_z,$$

and one can verify that the right-hand sides of these last equations are invariant under the transformations (5.145), as one would expect in the case of a passive translation.)

Example 5.59 (Passive rotations in the field of a magnetic monopole). Under a passive rotation about the z-axis through an angle s, the Cartesian coordinates, $(q_1, q_2, q_3) = (x, y, z)$, of a particle transform as [cf. (5.136)]

$$Q_1 = q_1 \cos s + q_2 \sin s, \qquad Q_2 = -q_1 \sin s + q_2 \cos s, \qquad Q_3 = q_3 \tag{5.146}$$

and, in a similar way, if there is a magnetic field present, the Cartesian components of its *kinematic* momentum transform as

$$\begin{aligned} P_1 - \frac{e}{c}A_1(Q_1, Q_2, Q_3) &= \left(p_1 - \frac{e}{c}A_1\right)\cos s + \left(p_2 - \frac{e}{c}A_2\right)\sin s, \\ P_2 - \frac{e}{c}A_2(Q_1, Q_2, Q_3) &= -\left(p_1 - \frac{e}{c}A_1\right)\sin s + \left(p_2 - \frac{e}{c}A_2\right)\cos s, \\ P_3 - \frac{e}{c}A_3(Q_1, Q_2, Q_3) &= p_3 - \frac{e}{c}A_3, \end{aligned} \tag{5.147}$$

where e is the electric charge of the particle, and the A_i are the Cartesian components of a vector potential for the magnetic field [recall that canonical momenta are related to the kinematic momenta by means of (4.31)]. From Equations (5.146) it follows that

$$\left.\frac{\partial Q_1}{\partial s}\right|_{s=0} = q_2, \qquad \left.\frac{\partial Q_2}{\partial s}\right|_{s=0} = -q_1, \qquad \left.\frac{\partial Q_3}{\partial s}\right|_{s=0} = 0,$$

and using these expressions and the chain rule, from Equations (5.147) one finds that

$$\begin{aligned} \left.\frac{\partial P_1}{\partial s}\right|_{s=0} &= p_2 - \frac{e}{c}A_2 + \frac{e}{c}\left(q_2\frac{\partial A_1}{\partial q_1} - q_1\frac{\partial A_1}{\partial q_2}\right), \\ \left.\frac{\partial P_2}{\partial s}\right|_{s=0} &= -p_1 + \frac{e}{c}A_1 + \frac{e}{c}\left(q_2\frac{\partial A_2}{\partial q_1} - q_1\frac{\partial A_2}{\partial q_2}\right), \\ \left.\frac{\partial P_3}{\partial s}\right|_{s=0} &= \frac{e}{c}\left(q_2\frac{\partial A_3}{\partial q_1} - q_1\frac{\partial A_3}{\partial q_2}\right). \end{aligned}$$

As shown in Proposition 5.49, if, for each value of s, the coordinate transformation given by Equations (5.146) and (5.147) is canonical, there must exist a function, G, such that [see Equations (5.121)]

$$\frac{\partial G}{\partial p_1} = q_2, \qquad \frac{\partial G}{\partial p_2} = -q_1, \qquad \frac{\partial G}{\partial p_3} = 0, \tag{5.148}$$

and

$$\begin{aligned} \frac{\partial G}{\partial q_1} &= -p_2 + \frac{e}{c}A_2 - \frac{e}{c}\left(q_2\frac{\partial A_1}{\partial q_1} - q_1\frac{\partial A_1}{\partial q_2}\right), \\ \frac{\partial G}{\partial q_2} &= p_1 - \frac{e}{c}A_1 - \frac{e}{c}\left(q_2\frac{\partial A_2}{\partial q_1} - q_1\frac{\partial A_2}{\partial q_2}\right), \\ \frac{\partial G}{\partial q_3} &= -\frac{e}{c}\left(q_2\frac{\partial A_3}{\partial q_1} - q_1\frac{\partial A_3}{\partial q_2}\right). \end{aligned} \tag{5.149}$$

However, these equations for G are not always integrable. From the equality of the mixed second partial derivatives of G, one obtains the conditions

$$q_1\frac{\partial B_1}{\partial q_2} - q_2\frac{\partial B_1}{\partial q_1} = -B_2, \qquad q_1\frac{\partial B_2}{\partial q_2} - q_2\frac{\partial B_2}{\partial q_1} = B_1, \qquad q_1\frac{\partial B_3}{\partial q_2} - q_2\frac{\partial B_3}{\partial q_1} = 0,$$

which mean that the magnetic field has to be invariant under the rotations about the z-axis. (Cf. Example 1.21.)

Thus, the rotations defined by Equations (5.146) and (5.147) are canonical transformations if and only if the magnetic field is invariant under the rotations about the z-axis. (Examples of such a field are a uniform magnetic field along the z-axis, and the field of a magnetic dipole aligned with the z-axis.)

On the other hand, the standard Hamiltonian in Cartesian coordinates for a charged particle of mass m and electric charge e in a magnetic field is [see (4.14)]

$$H = \frac{1}{2m}\left[\left(p_1 - \frac{e}{c}A_1\right)^2 + \left(p_2 - \frac{e}{c}A_2\right)^2 + \left(p_3 - \frac{e}{c}A_3\right)^2\right],$$

where the A_i are the Cartesian components of a vector potential for the magnetic field present. Since the transformation (5.146)–(5.147) does not involve the time, $K_s = H + f(t,s)$, where $f(t,s)$ is a function of t and s only and one can readily see that, choosing $f(t,s) = 0$, H is invariant under the rotations (5.146)–(5.147) (even if the magnetic field is not invariant under these rotations!).

Hence, if the magnetic field is invariant under rotations about the z-axis, the one-parameter group of rotations (5.146)–(5.147) is formed by canonical transformations that leave H invariant and, therefore, its generating function, G, determined by Equations (5.148)–(5.149), and $\partial G/\partial t = 0$ [see Equation (5.123)] is a constant of motion. It should be clear that analogous results hold for rotations about the other coordinate axes.

The only magnetic field invariant under rotations about the three coordinate axes (and, therefore, about any axis passing through the origin) is of the form

$$\mathbf{B} = g\frac{\mathbf{r}}{|\mathbf{r}|^3}, \tag{5.150}$$

where g is a constant. This field would be produced by a magnetic monopole placed at the origin. Since $\nabla \cdot \mathbf{B}$ is different from zero at the origin, there does not exist a well-behaved vector potential for this field defined everywhere. However, one can readily verify that

$$\mathbf{A} = \frac{gz(y\,\mathbf{i} - x\,\mathbf{j})}{(x^2+y^2)\sqrt{x^2+y^2+z^2}}$$

is a vector potential for the field (5.150), which is not defined on the points of the z-axis. (This vector potential is invariant under rotations about the z-axis, but other equivalent vector potentials may not have this symmetry.)

Substituting this vector potential into the right-hand sides of Equations (5.149) one finds that the solution of Equations (5.148)–(5.149) and $\partial G/\partial t = 0$ is $G = -xp_y + yp_x$, which is a constant of motion, as one can directly verify. Expressing the canonical momentum in terms of the kinematic momentum, $p_i = m\dot{q}_i + (e/c)A_i$, one obtains the equivalent expression

$$G = -(xm\dot{y} - ym\dot{x}) + \frac{eg}{c}\frac{z}{r}. \tag{5.151}$$

In this case we have two additional constants of motion, associated with the rotations about the x- and y-axes. With all these constants of motion we can form the conserved vector

$$\mathbf{r} \times m\dot{\mathbf{r}} - \frac{eg}{c}\frac{\mathbf{r}}{r}, \tag{5.152}$$

which reduces to the elementary definition of the angular momentum when $g = 0$. (Cf. Example 5.53.)

Exercise 5.60. The most general vector potential corresponding to the field (5.150) is of the form

$$\mathbf{A} = \frac{gz(y\,\mathbf{i} - x\,\mathbf{j})}{(x^2+y^2)\sqrt{x^2+y^2+z^2}} + \nabla\xi,$$

where ξ is some function of the coordinates x, y, z, only. Show explicitly that the generating function G of the rotations (5.146)–(5.147), expressed in terms of the kinematic momentum, is given by (5.151) and, therefore, it does not depend on the function ξ. (This behavior is to be expected since a constant of motion must not depend on the choice of the vector potential. Recall that, by contrast with the canonical momentum, which is gauge-dependent, the kinematic momentum does not depend on the gauge.)

Exercise 5.61. Show that the function G defined by Equations (5.148)–(5.149) is given by

$$G = q_2\left(p_1 - \frac{e}{c}A_1\right) - q_1\left(p_2 - \frac{e}{c}A_2\right) + \frac{e}{c}\int (q_1B_1 + q_2B_2)\mathrm{d}q_3 - B_3(q_1\mathrm{d}q_1 + q_2\mathrm{d}q_2).$$

(The integrand is the z-component of the vector field $\mathbf{r} \times (\mathrm{d}\mathbf{r} \times \mathbf{B})$.) Note that the differences $p_i - (e/c)A_i$ are the components of the kinematic momentum and are, therefore, gauge-independent. In the case of the magnetic field (5.150), this expression reduces to (5.151).

Exercise 5.62 (Passive translations in a magnetic field). Consider a charged particle in a (not necessarily uniform) static magnetic field. Under a passive translation by a distance s along the z-axis, the Cartesian coordinates of the particle transform according to

$$Q_1 = q_1, \qquad Q_2 = q_2, \qquad Q_3 = q_3 - s.$$

Assuming that the canonical momenta transform in such a way that

$$P_i - \frac{e}{c}A_i(Q_1, Q_2, Q_3) = p_i - \frac{e}{c}A_i(q_1, q_2, q_3)$$

($i = 1, 2, 3$) (that is, the Cartesian components of the kinematic momentum do not vary), show that these transformations are canonical if and only if the magnetic field is invariant under the translations (that is, the Cartesian components of the magnetic field are functions of q_1 and q_2 only) and that the generating function of this group is

$$G = -\left(p_3 - \frac{e}{c}A_3\right) + \frac{e}{c}\int (B_2\mathrm{d}q_1 - B_1\mathrm{d}q_2)$$

(cf. Exercise 4.14).

Example 5.63 (Galilean transformations in one dimension). As shown in Example 5.14, the coordinate transformations

$$Q = q - vt, \qquad P = p - mv, \tag{5.153}$$

which depend on the parameter v, are canonical and the new Hamiltonian is given by

$$K_v = H - vp + \tfrac{1}{2}mv^2 + f(t, v),$$

where $f(t, v)$ is a function of t and v only. Hence, taking H as in Example 5.56,

$$H(q, p, t) = \frac{p^2}{2m} + mgq,$$

we have

$$\begin{aligned} K_v(q, p, t) &= \frac{p^2}{2m} + mgq - vp + \frac{1}{2}mv^2 + f(t, v) \\ &= \frac{P^2}{2m} + mgQ + mgvt + f(t, v) \\ &= H(Q, P, t) + mgvt + f(t, v). \end{aligned}$$

Thus, H is invariant under the one-parameter group of canonical transformations (5.153) if we take $f(t, v) = -mgvt$ and, with this choice, the right-hand side of (5.21) is the differential of the function

$$F_1 = mvq - \tfrac{1}{2}mv^2t - \tfrac{1}{2}mgvt^2$$

(which, again, *is not* a generating function because the set (q, Q, t) is not functionally independent). Finally, substituting this last expression into (5.122) (with the parameter v in place of s), we find that

$$G = mq - \tfrac{1}{2}mgt^2 - pt,$$

which, as one can readily verify, is a constant of motion. (G can also be obtained from Equations (5.121) and (5.123).)

Assuming that H has the usual form

$$H = \frac{p^2}{2m} + V(q, t),$$

where $V(q, t)$ is a function of q and t only, we can find the most general form of the potential $V(q, t)$ such that H is invariant under the Galilean transformations. The invariance condition (5.137) reduces to

$$V(Q + vt, t) = V(Q, t) - f(t, v), \tag{5.154}$$

which must be valid for all values of v. Hence, taking the derivative with respect to v, at $v = 0$, of both sides of the last equation we obtain

$$t\frac{\partial V}{\partial q} = -\left.\frac{\partial f(t, v)}{\partial v}\right|_{v=0} \equiv -F(t),$$

where $F(t)$ is a function of t only. Hence, up to an irrelevant additive function of t only,

$$V(q,t) = -\frac{F(t)}{t}q, \tag{5.155}$$

which corresponds to a possibly time-dependent uniform field, and, from (5.154), it follows that $f(t,v) = F(t)\,v$. Then, making use of Equations (5.121) and (5.123), we find that the associated constant of motion is

$$G = mq - pt + \int F(t)\,\mathrm{d}t.$$

An important conclusion from the examples above is that there is no unique conserved generating function associated with a one-parameter group of canonical transformations. A similar result holds in the framework of quantum mechanics (see, e.g., Torres del Castillo and Herrera Flores [18]).

Exercise 5.64. According to Equations (5.121), the generating function of the Galilean transformations (5.153) is of the form $G = mq - pt + \chi(t)$, where $\chi(t)$ is a function of t only. Assuming that the Hamiltonian has the form $H = p^2/2m + V(q,t)$, show that (5.155) is the most general potential for which G is conserved.

Exercise 5.65. Show that the Hamiltonian

$$H = \frac{{p_1}^2}{2m_1} + \frac{{p_2}^2}{2m_2} + V(|q_2 - q_1|),$$

corresponding to a system of two particles of masses m_1 and m_2 interacting through a potential, V, that only depends on their separation, is invariant under the Galilean transformations,

$$Q_1 = q_1 - vt, \qquad P_1 = p_1 - m_1 v,$$
$$Q_2 = q_2 - vt, \qquad P_2 = p_2 - m_2 v.$$

Find the generating function $G(q_i, p_i, t)$ of this group of transformations and identify its physical meaning.

Exercise 5.66 (A nonstandard Hamiltonian). Consider the Hamiltonian

$$H = \frac{p_1 p_2}{m} + m\omega^2 q_1 q_2,$$

where m and ω are constants. Find the equations of motion for q_1 and q_2. What is the mechanical system they correspond to? Now, consider the one-parameter family of coordinate transformations

$$Q_1 = \mathrm{e}^s q_1, \qquad Q_2 = \mathrm{e}^{-s} q_2, \qquad P_1 = \mathrm{e}^{-s} p_1, \qquad P_2 = \mathrm{e}^s p_2.$$

Show that this is a family of canonical transformations that leave invariant H. Find the associated constant of motion and its physical meaning.

It may be remarked that, apart from simple cases (e.g., Hamiltonians with ignorable coordinates), it is usually difficult to identify the symmetries of a given Hamiltonian (that is, the one-parameter groups of canonical transformations that leave H invariant) and, therefore, the constants of motion obtained with the aid of Proposition 5.55 are relatively few. More often, the constants of motion are found by the direct integration of the equations of motion, making use of some ansatz (see, e.g., Perelomov [13, Sect. 2.8]), the application of the Poisson Theorem, the method of equivalent Hamiltonians presented in Section 4.3, by inspection (see, e.g., Examples 5.17 and 5.18), or the solution of the Hamilton–Jacobi equation (to be considered in Chapter 6).

The Symmetries of the Hamiltonian Generated by a Constant of Motion
As pointed out above, the converse of Proposition 5.55 is also true: any constant of motion, G, is the generating function of a (possibly local) one-parameter group of canonical transformations that leave H invariant (see Proposition 5.74, below). Among other things, this result implies that any Hamiltonian possesses an infinite number of one-parameter symmetry groups, since there are an infinite number of constants of motion (but only $2n$ can be functionally independent). We start presenting some examples and exercises where this fact is illustrated, postponing the proof to the end of this section.

Example 5.67. One can readily verify that the function

$$G = q - \frac{tp}{m} + \frac{kt^3}{3m}$$

is a constant of motion if the Hamiltonian is

$$H = \frac{p^2}{2m} - ktq. \tag{5.156}$$

Making use of (5.128) one finds that the one-parameter group of canonical transformations generated by G is given by

$$Q = q - \frac{ts}{m}, \qquad P = p - s. \tag{5.157}$$

Then, from Equations (5.12) one finds that the new Hamiltonian is

$$\begin{aligned} K_s &= H - \frac{ps}{m} + f(t,s) \\ &= \frac{p^2}{2m} - ktq - \frac{ps}{m} + f(t,s), \end{aligned}$$

where $f(t,s)$ is a function of t and s only.

On the other hand, from (5.156) and (5.157),

$$H(Q(q,p,t,s),P(q,p,t,s),t) = \frac{(p-s)^2}{2m} - kt\left(q - \frac{ts}{m}\right),$$

which coincides with the expression of K_s obtained above if

$$f(t,s) = \frac{s^2}{2m} + \frac{kt^2 s}{m},$$

thus showing that H is invariant under the group of canonical transformations generated by the constant of motion G.

Exercise 5.68. Show by direct substitution that the form of the equations of motion in terms of the canonical coordinates Q and P given by (5.157) is the same as the form of the equations of motion in terms of q and p, if the Hamiltonian is given by (5.156).

Example 5.69. The function

$$G = p_x p_y + m^2 g x \tag{5.158}$$

is a constant of motion if the Hamiltonian is the standard one for a particle of mass m in a uniform gravitational field

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy \tag{5.159}$$

(see, e.g., Example 5.35). The one-parameter group of canonical transformations generated by G is determined by the system of equations [see Equations (5.128)]

$$\frac{\mathrm{d}Q_1}{\mathrm{d}s} = P_2, \qquad \frac{\mathrm{d}Q_2}{\mathrm{d}s} = P_1, \qquad \frac{\mathrm{d}P_1}{\mathrm{d}s} = -m^2 g, \qquad \frac{\mathrm{d}P_2}{\mathrm{d}s} = 0$$

with the initial conditions $Q_1(0) = x$, $Q_2(0) = y$, $P_1(0) = p_x$, $P_2(0) = p_y$. The last two equations are readily integrated, giving

$$P_1 = p_x - m^2 g s, \qquad P_2 = p_y. \tag{5.160}$$

Then, substituting these expressions into the first pair, we obtain

$$Q_1 = x + p_y s, \qquad Q_2 = y + p_x s - \tfrac{1}{2} m^2 g s^2. \tag{5.161}$$

Since the canonical transformation (5.160)–(5.161) does not involve the time, $K_s = H + f(t,s)$, where $f(t,s)$ is a function of t and s only. Hence,

$$\begin{aligned}
K_s &= \frac{{p_x}^2 + {p_y}^2}{2m} + mgy + f(t,s) \\
&= \frac{1}{2m}\big[(P_1 + m^2gs)^2 + {P_2}^2\big] + mg\big(Q_2 - P_1s - \tfrac{1}{2}m^2gs^2\big) + f(t,s) \\
&= \frac{{P_1}^2 + {P_2}^2}{2m} + mgsP_1 + \frac{1}{2}m^3g^2s^2 + mgQ_2 - mgP_1s - \frac{1}{2}m^3g^2s^2 + f(t,s) \\
&= \frac{{P_1}^2 + {P_2}^2}{2m} + mgQ_2 + f(t,s) \\
&= H(Q_i(q_j,p_j), P_i(q_j,p_j)) + f(t,s).
\end{aligned}$$

Hence, choosing $f(t,s) = 0$, we see that the invariance condition [Equation (5.137)] is satisfied.

Exercise 5.70. Show explicitly that the Hamiltonian (5.48) is invariant under the canonical transformations generated by the constant of motion

$$G = \left(\frac{x}{2} - \frac{c}{eB_0}p_y\right)\cos\omega_c t - \left(\frac{y}{2} + \frac{c}{eB_0}p_x\right)\sin\omega_c t,$$

where $\omega_c \equiv eB_0/mc$. (In terms of the kinematic momentum of the particle, G is given by the gauge-invariant expression $G = -(\pi_2\cos\omega_c t + \pi_1\sin\omega_c t)/eB_0$.)

Example 5.71. The function

$$G = -xp_y + yp_x - mgtx + \tfrac{1}{2}gt^2p_x \tag{5.162}$$

is a constant of motion if the Hamiltonian is given by (5.159). Solving Equations (5.128) one finds that the one-parameter group of canonical transformations generated by G is given by

$$\begin{aligned}
Q_1 &= x\cos s + (y + \tfrac{1}{2}gt^2)\sin s, \quad Q_2 = -\tfrac{1}{2}gt^2 - x\sin s + (y + \tfrac{1}{2}gt^2)\cos s, \\
P_1 &= p_x\cos s + (p_y + mgt)\sin s, \quad P_2 = -mgt - p_x\sin s + (p_y + mgt)\cos s.
\end{aligned} \tag{5.163}$$

Instead of showing that the Hamiltonian is invariant under these transformations, it is simpler to show that the form of the equations of motion is invariant under (5.163). In fact, the Hamilton equations given by the Hamiltonian (5.159) are

$$\dot{x} = \frac{p_x}{m}, \qquad \dot{y} = \frac{p_y}{m}, \qquad \dot{p}_x = 0, \qquad \dot{p}_y = -mg. \tag{5.164}$$

Then, making use of these equations and (5.163) we have

$$\dot{Q}_1 = \dot{x}\cos s + (\dot{y} + gt)\sin s = \frac{p_x}{m}\cos s + \left(\frac{p_y}{m} + gt\right)\sin s = \frac{P_1}{m},$$

$$\dot{Q}_2 = -gt - \dot{x}\sin s + (\dot{y} + gt)\cos s = -gt - \frac{p_x}{m}\sin s + \left(\frac{p_y}{m} + gt\right)\cos s = \frac{P_2}{m},$$

$$\dot{P}_1 = \dot{p}_x\cos s + (\dot{p}_y + mg)\sin s = 0,$$

$$\dot{P}_2 = -mg - \dot{p}_x\sin s + (\dot{p}_y + mg)\cos s = -mg,$$

which have the same form as Equations (5.164).

Exercise 5.72. The function

$$G = -\frac{t^2}{2m}({p_x}^2 + {p_y}^2) + t(xp_x + yp_y) - \frac{m}{2}(x^2 + y^2) - \frac{1}{2}gt^3 p_y + \frac{m}{2}gt^2 y - \frac{1}{8}mg^2 t^4$$

is a constant of motion if the Hamiltonian is the standard one of a particle of mass m in a uniform gravitational field (5.159). Find the one-parameter group of canonical transformations generated by G and show explicitly that the equations of motion (5.164) maintain their form under these transformations.

Exercise 5.73. Show explicitly that the Hamiltonian

$$H = \frac{p^2}{2m} + mgq$$

is invariant under the one-parameter group of canonical transformations generated by the function (5.129). (Cf. also Example 5.21.)

The invariance of the Hamiltonian under the transformations generated by a constant of motion is equivalent to the invariance of the form of the equations of motion under these transformations. In the following proposition we shall prove that if the equations of motion are given by

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = X_i(q_j, p_j, t), \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = Y_i(q_j, p_j, t), \tag{5.165}$$

where X_i and Y_i are convenient ways of denoting the partial derivatives of the Hamiltonian,

$$X_i \equiv \frac{\partial H}{\partial p_i}, \qquad Y_i \equiv -\frac{\partial H}{\partial p_i}, \tag{5.166}$$

then

$$\frac{\mathrm{d}Q_i}{\mathrm{d}t} = X_i(Q_j, P_j, t), \qquad \frac{\mathrm{d}P_i}{\mathrm{d}t} = Y_i(Q_j, P_j, t), \tag{5.167}$$

where $Q_i = Q_i(q_j, p_j, t, s)$ and $P_i = P_i(q_j, p_j, t, s)$ represent the one-parameter group of canonical transformations generated by a constant of motion G, according to Equations (5.128).

Proposition 5.74. *If G is a constant of motion, then the Hamilton equations are form-invariant under the one-parameter group of canonical transformations generated by G.*

Proof. Making use of the chain rule, the Hamilton equations, and Equations (5.128) we obtain

$$\begin{aligned}
\frac{\partial}{\partial s}\left(\frac{\mathrm{d}Q_i}{\mathrm{d}t}\right)_{s=0} &= \frac{\partial}{\partial s}\left(\frac{\partial Q_i}{\partial t} + \frac{\partial Q_i}{\partial q_j}\frac{\mathrm{d}q_j}{\mathrm{d}t} + \frac{\partial Q_i}{\partial p_j}\frac{\mathrm{d}p_j}{\mathrm{d}t}\right)_{s=0} \\
&= \frac{\partial}{\partial t}\left.\frac{\partial Q_i}{\partial s}\right|_{s=0} + \frac{\partial}{\partial q_j}\left.\frac{\partial Q_i}{\partial s}\right|_{s=0}\frac{\partial H}{\partial p_j} - \frac{\partial}{\partial p_j}\left.\frac{\partial Q_i}{\partial s}\right|_{s=0}\frac{\partial H}{\partial q_j} \\
&= \frac{\partial}{\partial t}\frac{\partial G}{\partial p_i} + \left(\frac{\partial}{\partial q_j}\frac{\partial G}{\partial p_i}\right)\frac{\partial H}{\partial p_j} - \left(\frac{\partial}{\partial p_j}\frac{\partial G}{\partial p_i}\right)\frac{\partial H}{\partial q_j} \\
&= \frac{\partial}{\partial p_i}\frac{\partial G}{\partial t} + \left(\frac{\partial}{\partial p_i}\frac{\partial G}{\partial q_j}\right)\frac{\partial H}{\partial p_j} - \left(\frac{\partial}{\partial p_i}\frac{\partial G}{\partial p_j}\right)\frac{\partial H}{\partial q_j} \\
&= \frac{\partial}{\partial p_i}\left(\frac{\partial G}{\partial t} + \frac{\partial G}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial G}{\partial p_j}\frac{\partial H}{\partial q_j}\right) \\
&\quad - \frac{\partial G}{\partial q_j}\left(\frac{\partial}{\partial p_i}\frac{\partial H}{\partial p_j}\right) + \frac{\partial G}{\partial p_j}\left(\frac{\partial}{\partial p_i}\frac{\partial H}{\partial q_j}\right).
\end{aligned}$$

Thus, making use of the hypothesis, Equations (5.128), the definitions (5.166), and the chain rule we finally have

$$\begin{aligned}
\frac{\partial}{\partial s}\left(\frac{\mathrm{d}Q_i}{\mathrm{d}t}\right)_{s=0} &= \left.\frac{\partial P_j}{\partial s}\right|_{s=0}\frac{\partial X_i}{\partial p_j} + \left.\frac{\partial Q_j}{\partial s}\right|_{s=0}\frac{\partial X_i}{\partial q_j} \\
&= \left.\frac{\partial}{\partial s}X_i(Q_j, P_j, t)\right|_{s=0}.
\end{aligned}$$

The remaining relations in (5.167) are proved in an analogous way. □

It may be noticed that following the steps above, if we assume that the form of the Hamilton equations is left invariant under the canonical transformations generated by a function G, then it follows that $\mathrm{d}G/\mathrm{d}t$ is a function of t only. Hence, by adding to G an appropriate function of t only, we could make $\mathrm{d}G/\mathrm{d}t$ equal to zero, without altering the transformations generated by G. (Cf. Proposition 5.55.)

5.5 Canonoid Transformations

As in the case of systems with one degree of freedom considered in Section 5.1, when the number of degrees of freedom is greater than 1, for each Hamiltonian, there exist coordinate transformations (known as canonoid transformations) that are not canonical but maintain the form of the Hamilton equations. In this section we shall extend the results established in Section 5.1, showing that each of these transformations leads to a set of constants of motion (which can be trivial).

A simple example of a canonoid transformation is provided by the Hamiltonian

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy, \tag{5.168}$$

which corresponds to a particle in a uniform gravitational field, in terms of Cartesian coordinates. The coordinate transformation

$$Q_1 = x, \qquad Q_2 = y, \qquad P_1 = p_y, \qquad P_2 = p_x, \tag{5.169}$$

is not canonical (e.g., $\{Q_1, P_1\} = \{x, p_y\} = 0$) and the time derivatives of the new coordinates are given by

$$\dot{Q}_1 = \frac{P_2}{m}, \qquad \dot{Q}_2 = \frac{P_1}{m}, \qquad \dot{P}_1 = -mg, \qquad \dot{P}_2 = 0.$$

Then, one can readily verify that these equations can be written in the form of the Hamilton equations with Hamiltonian $K = P_1 P_2/m + mgQ_1$. Thus, the coordinate transformation (5.169) is canonoid for the Hamiltonian (5.168).

We are going to establish necessary and sufficient conditions for a coordinate transformation to be a canonoid transformation for a given Hamiltonian.

As shown at the beginning of Section 5.2, if the coordinates q_i, p_i satisfy the Hamilton equations

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q_i} \tag{5.170}$$

($i = 1, 2, \ldots, n$), then the new coordinates $Q_i = Q_i(q_j, p_j, t)$, $P_i = P_i(q_j, p_j, t)$ satisfy the equations

$$\frac{\mathrm{d}Q_i}{\mathrm{d}t} = \frac{\partial K}{\partial P_i}, \qquad \frac{\mathrm{d}P_i}{\mathrm{d}t} = -\frac{\partial K}{\partial Q_i} \tag{5.171}$$

($i = 1, 2, \ldots, n$), for some function K, if and only if the equations

$$\begin{aligned}\frac{\partial K}{\partial q_i} &= [q_i, p_k]\frac{\partial H}{\partial q_k} - [q_i, q_k]\frac{\partial H}{\partial p_k} + [t, q_i],\\ \frac{\partial K}{\partial p_i} &= [p_i, p_k]\frac{\partial H}{\partial q_k} - [p_i, q_k]\frac{\partial H}{\partial p_k} + [t, p_i]\end{aligned} \tag{5.172}$$

hold. Making use of the notation employed in the proofs of Propositions 4.9 and 5.44, we can write Equations (5.172) in the compact form

$$\begin{aligned}\frac{\partial K}{\partial x_\alpha} &= [x_\alpha, p_k]\frac{\partial H}{\partial q_k} - [x_\alpha, q_k]\frac{\partial H}{\partial p_k} + [t, x_\alpha]\\ &= \varepsilon_{\mu\nu}[x_\alpha, x_\nu]\frac{\partial H}{\partial x_\mu} + [t, x_\alpha]\end{aligned} \tag{5.173}$$

($\alpha = 1, 2, \ldots, 2n$). (Note that when the transformation is canonical, $\varepsilon_{\mu\nu}[x_\alpha, x_\nu] = \delta_{\mu\alpha}$, and (5.173) reduces to (5.42).)

As usual, the local existence of a function K satisfying a system of first-order PDEs of the form (5.173) follows from the integrability conditions $\partial^2 K/\partial x_\beta \partial x_\alpha = \partial^2 K/\partial x_\alpha \partial x_\beta$. Making use of the antisymmetry of the Lagrange brackets and the identities (5.40) we find that

$$\begin{aligned}&\frac{\partial^2 K}{\partial x_\beta \partial x_\alpha} - \frac{\partial^2 K}{\partial x_\alpha \partial x_\beta}\\ &= \varepsilon_{\mu\nu}\frac{\partial [x_\alpha, x_\nu]}{\partial x_\beta}\frac{\partial H}{\partial x_\mu} + \varepsilon_{\mu\nu}[x_\alpha, x_\nu]\frac{\partial^2 H}{\partial x_\beta \partial x_\mu} + \frac{\partial [t, x_\alpha]}{\partial x_\beta}\\ &\quad - \varepsilon_{\mu\nu}\frac{\partial [x_\beta, x_\nu]}{\partial x_\alpha}\frac{\partial H}{\partial x_\mu} - \varepsilon_{\mu\nu}[x_\beta, x_\nu]\frac{\partial^2 H}{\partial x_\alpha \partial x_\mu} - \frac{\partial [t, x_\beta]}{\partial x_\alpha}\\ &= -\varepsilon_{\mu\nu}\frac{\partial [x_\beta, x_\alpha]}{\partial x_\nu}\frac{\partial H}{\partial x_\mu} + \varepsilon_{\mu\nu}[x_\alpha, x_\nu]\frac{\partial^2 H}{\partial x_\beta \partial x_\mu} - \varepsilon_{\mu\nu}[x_\beta, x_\nu]\frac{\partial^2 H}{\partial x_\alpha \partial x_\mu} + \frac{\partial [x_\beta, x_\alpha]}{\partial t}.\end{aligned}$$

Thus, if the equations of motion (5.170) can also be expressed in the form (5.171), then the Lagrange brackets $[x_\beta, x_\alpha]$ satisfy the system of linear PDEs

$$\frac{\partial [x_\beta, x_\alpha]}{\partial t} + \varepsilon_{\nu\mu}\frac{\partial [x_\beta, x_\alpha]}{\partial x_\nu}\frac{\partial H}{\partial x_\mu} + \varepsilon_{\mu\nu}[x_\alpha, x_\nu]\frac{\partial^2 H}{\partial x_\beta \partial x_\mu} - \varepsilon_{\mu\nu}[x_\beta, x_\nu]\frac{\partial^2 H}{\partial x_\alpha \partial x_\mu} = 0, \tag{5.174}$$

and K is determined by Equations (5.173), up to an arbitrary additive function of t only.

For the readers familiarized with the formalism of differential forms (also called exterior calculus), it may be pointed out that the Lagrange brackets can be readily calculated making use of the fact that

$$\mathrm{d}Q_i \wedge \mathrm{d}P_i = \tfrac{1}{2}[x_\alpha, x_\beta]\,\mathrm{d}x_\alpha \wedge \mathrm{d}x_\beta + [t, x_\alpha]\,\mathrm{d}t \wedge \mathrm{d}x_\alpha.$$

Constants of Motion Associated with a Canonoid Transformation
Now, it is convenient to define the two $2n \times 2n$ matrices, $M = (M_{\alpha\beta})$ and $\Phi = (\Phi_{\alpha\beta})$, with entries

$$M_{\alpha\beta} \equiv \varepsilon_{\alpha\gamma}[x_\beta, x_\gamma], \qquad \Phi_{\alpha\beta} \equiv \varepsilon_{\alpha\gamma}\frac{\partial^2 H}{\partial x_\gamma \partial x_\beta}. \tag{5.175}$$

(Note that M is the unit matrix if and only if $[x_\beta, x_\gamma] = \varepsilon_{\beta\gamma}$, which means that the coordinates Q_i, P_i are related to q_i, p_i by means of a canonical transformation.) Then, multiplying both sides of (5.174) by $\varepsilon_{\gamma\alpha}$, using the antisymmetry of $\varepsilon_{\alpha\beta}$ and of $[x_\alpha, x_\beta]$, we find that Equations (5.174) can be written in the equivalent form

$$\frac{\mathrm{d}M_{\gamma\beta}}{\mathrm{d}t} + M_{\gamma\nu}\Phi_{\nu\beta} - \Phi_{\gamma\mu}M_{\mu\beta} = 0,$$

assuming that the coordinates x_α satisfy the equations of motion, i.e., Equations (5.170) are satisfied. This last equation amounts to the matrix equation

$$\frac{\mathrm{d}M}{\mathrm{d}t} = \Phi M - M\Phi \tag{5.176}$$

and from Equation (5.176) one readily finds that, for $k = \pm 1, \pm 2, \ldots,$

$$\frac{\mathrm{d}M^k}{\mathrm{d}t} = \Phi M^k - M^k\Phi \tag{5.177}$$

(the matrix M must be invertible as a consequence of the invertibility of the coordinate transformation). Taking the trace on both sides of (5.177) it follows that the trace of any integral power of M is a constant of motion

$$\frac{\mathrm{d}(\operatorname{tr} M^k)}{\mathrm{d}t} = 0.$$

Of course, these constants of motion cannot be all functionally independent, and may be trivial constants [as in the case of the transformation (5.169)]. Since the determinant of a matrix is a polynomial of the traces of its integral powers, the determinant of M is also a constant of motion. (See also Exercise 5.76, below.)

Exercise 5.75. Show directly from (5.176) that $\det M$ is a constant of motion.

When $n = 1$ (the case already considered in Section 5.1), the antisymmetry of $\varepsilon_{\alpha\beta}$ and of $[x_\alpha, x_\beta]$ under the interchange of the indices α and β imply that M is a multiple of the 2×2 identity matrix. In fact, from (5.105) and (5.175) we have

$$M = \begin{pmatrix} [x_1, x_2] & 0 \\ 0 & [x_1, x_2] \end{pmatrix} = \begin{pmatrix} [q_1, p_1] & 0 \\ 0 & [q_1, p_1] \end{pmatrix}.$$

Since a multiple of the identity matrix commutes with any matrix, from (5.176) we only obtain the condition that $[x_1, x_2]$ (i.e., $[q, p]$, or $\{Q, P\}$) is a constant of motion, which is just Proposition 5.1.

In a similar way, from Equations (5.105) and (5.175), making use of the antisymmetry of the Lagrange brackets, one finds that when $n = 2$, M is the 4×4 matrix

$$M = \begin{pmatrix} [q_1, p_1] & [q_2, p_1] & 0 & -[p_1, p_2] \\ [q_1, p_2] & [q_2, p_2] & [p_1, p_2] & 0 \\ 0 & [q_1, q_2] & [q_1, p_1] & [q_1, p_2] \\ -[q_1, q_2] & 0 & [q_2, p_1] & [q_2, p_2] \end{pmatrix}. \tag{5.178}$$

Exercise 5.76. Evidently, when $n = 1$, the matrix M has only one repeated eigenvalue (equal to $[q_1, p_1]$). Show that, similarly, when $n = 2$, each eigenvalue of the matrix M has even multiplicity. Show that the eigenvalues of M are constants of motion. (As a consequence of the antisymmetry of $[x_\alpha, x_\beta]$ and $\varepsilon_{\alpha\beta}$, for an arbitrary number, n, of degrees of freedom, the characteristic polynomial of M is the square of a polynomial of degree n (see, e.g., Eves [7, Sect. 3.6A]) and, therefore, each eigenvalue of M has even multiplicity.)

Now we can readily demonstrate that, as claimed at the beginning of Section 5.2, the only coordinate transformations $Q_i = Q_i(q_j, p_j, t)$, $P_i = P_i(q_j, p_j, t)$ that preserve the form of the Hamilton equations, for *any* Hamiltonian, are essentially the canonical transformations. Indeed, from Equations (5.174) one concludes that there exists a new Hamiltonian K for any function H if and only if the terms with the first and second partial derivatives of H separately vanish. Since the matrix $(\varepsilon_{\mu\nu})$ is non-singular, this amounts to say that all the Lagrange brackets $[x_\beta, x_\alpha]$ must be (trivial) constants, and that the entries of the (constant) matrix M must satisfy

$$M_{\mu\alpha} \frac{\partial^2 H}{\partial x_\beta \partial x_\mu} = M_{\mu\beta} \frac{\partial^2 H}{\partial x_\alpha \partial x_\mu},$$

for any function H. This last condition holds if and only if M is a multiple of the unit matrix and, therefore (see (5.175) and recall that $\varepsilon_{\alpha\beta}\varepsilon_{\gamma\beta} = \delta_{\alpha\gamma}$),

$$[x_\alpha, x_\beta] = \lambda \varepsilon_{\alpha\beta},$$

where λ is a nonzero constant. By rescaling the new coordinates we can eliminate λ from the last equation and in that way we recover Equations (5.41), which was the definition of a canonical transformation.

Canonoid Transformations and Alternative Poisson Brackets
As we have seen, given a set of first-order ODEs expressed in the form of the Hamilton equations in terms of coordinates (q_i, p_i, t), it is possible to find other coordinate systems, (Q_i, P_i, t), not related to (q_i, p_i, t) by canonical transformations, in terms of which the given set of equations can also be written in the form of the Hamilton equations and, in the same way as the canonical transformations can be used to simplify one of such sets, or to relate two different problems, the canonoid transformations can be employed to convert a system with a given Hamiltonian into another with a new Hamiltonian.

If the coordinate systems (q_i, p_i, t) and (Q_i, P_i, t) are not related by a canonical transformation, then the Poisson brackets

$$\{f, g\}_1 \equiv \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}, \qquad \{f, g\}_2 \equiv \frac{\partial f}{\partial Q_i}\frac{\partial g}{\partial P_i} - \frac{\partial f}{\partial P_i}\frac{\partial g}{\partial Q_i}$$

do not coincide. Thus, with a canonoid transformation we obtain a new Poisson bracket (with all the usual properties of a Poisson bracket). The Poisson bracket $\{\,,\,\}_2$ is, in turn, invariant under the canonical transformations that lead from (Q_i, P_i, t) to another coordinate system of the extended phase space.

Given a Hamiltonian, H, in terms of some coordinate system (q_i, p_i, t), each canonoid transformation admitted by H corresponds to a nontrivial solution of the system of equations (5.174) (that is, $[x_\alpha, x_\beta] \neq \lambda\varepsilon_{\alpha\beta}$, with a constant factor λ, and $\det([x_\alpha, x_\beta]) \neq 0$) that satisfies the conditions (5.40). Conversely, given a set of functions, $[x_\alpha, x_\beta] = -[x_\beta, x_\alpha]$, $[x_\alpha, t]$, satisfying the conditions (5.40), $\det([x_\alpha, x_\beta]) \neq 0$, and (5.174), then, by virtue of the Darboux theorem (see, e.g., Crampin and Pirani [5]), there exist local coordinates (Q_i, P_i, t) such that the Lagrange brackets $[x_\alpha, x_\beta]_{Q,P}$ and $[x_\alpha, t]_{Q,P}$ coincide with the given functions and the transformation $(q_i, p_i, t) \mapsto (Q_i, P_i, t)$ preserves the form of the Hamilton equations; an example of this fact is given in Example 5.78, below.

Exercise 5.77. Show that, in terms of the coordinates (q_i, p_i, t), the Poisson bracket $\{\,,\,\}_2$ is given by

$$\{f, g\}_2 = (M^{-1})_{\alpha\gamma}\varepsilon_{\gamma\beta}\frac{\partial f}{\partial x_\alpha}\frac{\partial g}{\partial x_\beta},$$

where $(M^{-1})_{\alpha\gamma}$ are the entries of the inverse of the matrix M (that is, $(M^{-1})_{\alpha\gamma}M_{\gamma\beta} = \delta_{\alpha\beta}$).

A locally equivalent approach to the subject of this section is presented in Das [6], but only for the restricted case where the Hamiltonians do not depend explicitly on the time and the coordinate transformations do not involve the time.

Instead of considering canonoid transformations explicitly, in Das [6] it is assumed that a given *autonomous* system of $2n$ first-order ODEs can be written in the form of the Hamilton equations, making use of two different Poisson brackets and two (possibly different) time-independent Hamiltonians, making use of a single coordinate system. (Equation (5.174) reduces to Equation (9.41) of Das [6] in the time-independent case.)

Example 5.78 (Some canonoid transformations for the two-dimensional isotropic harmonic oscillator). We consider the Hamiltonian

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + \frac{m\omega^2}{2}(x^2 + y^2),$$

corresponding to a two-dimensional isotropic harmonic oscillator. Instead of attempting to find the general solution of the system of PDEs (5.174), we restrict ourselves to coordinate transformations such that the Lagrange brackets $[x_\alpha, x_\beta]$ are all constant and $[t, x_\alpha] = 0$, then the conditions (5.40) are trivially satisfied, the matrix M is constant [see (5.175)] and from Equation (5.176) it follows that M must commute with the matrix Φ, which is given by

$$\Phi = \begin{pmatrix} 0 & 0 & 1/m & 0 \\ 0 & 0 & 0 & 1/m \\ -m\omega^2 & 0 & 0 & 0 \\ 0 & -m\omega^2 & 0 & 0 \end{pmatrix} \tag{5.179}$$

[see (5.175)]. This condition, together with the fact that M must be of the form (5.178), imply that

$$M = \begin{pmatrix} a & b & 0 & -d \\ b & c & d & 0 \\ 0 & e & a & b \\ -e & 0 & b & c \end{pmatrix}, \tag{5.180}$$

with $e = m^2\omega^2 d$. The four constants a, b, c, and d are only restricted by the condition $\det M = (b^2 - ac + ed)^2 \neq 0$.

Making use of the assumption that $[t, x_\alpha] = 0$, from Equations (5.173) we obtain the expression for the new Hamiltonian in terms of the original coordinates

$$K = a\left(\frac{{p_x}^2}{2m} + \frac{m\omega^2}{2}x^2\right) + b\left(\frac{p_x p_y}{m} + m\omega^2 xy\right) + c\left(\frac{{p_y}^2}{2m} + \frac{m\omega^2}{2}y^2\right) - dm\omega^2(xp_y - yp_x). \tag{5.181}$$

Since K does not depend on the time explicitly, it is a constant of motion for all values of a, b, c, and d, which are arbitrary (except for the condition $b^2 - ac +$

$ed \neq 0$), hence each expression inside the parentheses in (5.181) is a constant of motion. (However, only three of them are functionally independent.) Thus, even though the determinant and the trace of all powers of M are trivial constants, the new Hamiltonian leads to three constants of motion.

It only remains to find the coordinate transformation explicitly. Taking, for instance, $a = b = c = 0$, $d = 1$ (and, hence, $e = m^2\omega^2$), the only Lagrange brackets different from zero are $[p_x, p_y] = 1$ and $[x, y] = m^2\omega^2$. Up to a canonical transformation, these conditions give

$$Q_1 = m\omega x, \qquad Q_2 = p_x, \qquad P_1 = m\omega y, \qquad P_2 = p_y.$$

Then, making use of (5.181), one readily finds that $K = -\omega(Q_1 P_2 - Q_2 P_1)$.

Exercise 5.79. Find the Hamiltonians, K, that lead to the same equations of motion as

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy,$$

assuming that the entries of the matrix M are trivial constants and that the Lagrange brackets $[t, x_\alpha]$ are equal to zero. What are the constants of motion that can be identified in this manner?

Exercise 5.80 (Two-dimensional isotropic harmonic oscillator). Verify that, for any value of the constant λ, the matrix

$$M = I + \lambda \begin{pmatrix} 0 & yp_x & 0 & p_x p_y/\beta \\ -xp_y & 0 & -p_x p_y/\beta & 0 \\ 0 & -\beta xy & 0 & -xp_y \\ \beta xy & 0 & yp_x & 0 \end{pmatrix}, \tag{5.182}$$

where I is the 4×4 unit matrix and $\beta \equiv m^2\omega^2$, has the form (5.178), satisfies Equations (5.176), with Φ given by (5.179), and conditions (5.40), assuming that the coordinate transformation does not involve t. Find the expression for the new Hamiltonian, K, in terms of the original coordinates.

Example 5.81 (A canonoid transformation for the Toda lattice). A more involved example is the so-called Toda lattice, which in its simplest nontrivial case ($n = 2$) is defined by the system of ODEs

$$\dot{q}_1 = p_1, \qquad \dot{q}_2 = p_2, \qquad \dot{p}_1 = -e^{q_1 - q_2}, \qquad \dot{p}_2 = e^{q_1 - q_2}. \tag{5.183}$$

One can readily verify that these equations are the Hamilton equations for

$$H = \tfrac{1}{2}({p_1}^2 + {p_2}^2) + e^{q_1 - q_2}. \tag{5.184}$$

A straightforward computation shows that the matrix

$$M = \begin{pmatrix} p_1 & 0 & 0 & 1 \\ 0 & p_2 & -1 & 0 \\ 0 & -e^{q_1-q_2} & p_1 & 0 \\ e^{q_1-q_2} & 0 & 0 & p_2 \end{pmatrix}, \tag{5.185}$$

which is of the form (5.178), satisfies Equation (5.176) as a consequence of (5.183). (Making use of (5.184) and (5.175) one finds that, in this case

$$\Phi = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -e^{q_1-q_2} & e^{q_1-q_2} & 0 & 0 \\ e^{q_1-q_2} & -e^{q_1-q_2} & 0 & 0 \end{pmatrix}.)$$

On the other hand, one readily verifies that the Lagrange brackets given by (5.185) satisfy (5.40), which guarantees the existence of coordinates (Q_i, P_i) such that the Lagrange brackets $[x_\alpha, x_\beta]$ coincide with the expressions determined by (5.185).

Assuming, as in Example 5.78, that the coordinate transformation does not involve t, making use of (5.173) we find that the new Hamiltonian is (in terms of the original coordinates)

$$K = \tfrac{1}{3}({p_1}^3 + {p_2}^3) + (p_1 + p_2)\, e^{q_1-q_2}.$$

A simple computation shows that $\operatorname{tr} M = 2(p_1+p_2)$ and $\det M = (p_1 p_2 - e^{q_1-q_2})^2$, which are constants of motion. (The conservation of $p_1 + p_2$ follows directly from (5.183) or (5.184).) Since H and K do not depend on t, they are also constants of motion, but one can readily see that they are functions of $p_1 + p_2$ and $p_1 p_2 - e^{q_1-q_2}$. (As we shall see in Section 6.2, these two constants of motion are sufficient to obtain a complementary pair of constants of motion and, therefore, to find the solution of the equations of motion.)

A more advanced and detailed study of this important system, with an arbitrary number of degrees of freedom, can be found, e.g., in Das [6, Chap. 10] and Perelomov [13, Chap. 4].

Chapter 6
The Hamilton–Jacobi Formalism

In the preceding chapters we have studied convenient forms of expressing some systems of ordinary differential equations, most of them related to mechanical systems. In this chapter we shall see that each of these systems of equations can be translated into a single partial differential equation, known as the Hamilton–Jacobi equation, which is constructed out of the Hamiltonian. A complete solution of this equation is the generating function of a canonical transformation that relates the coordinates being employed with another set of canonical coordinates which are all constants of motion.

As we have seen in the preceding chapter, the canonical transformations can be employed to simplify the solution of the Hamilton equations. However, we do not have a systematic method to find a convenient canonical transformation for any given Hamiltonian. As we shall show in this chapter, by finding a complete solution of a certain first-order partial differential equation (the Hamilton–Jacobi equation, or HJ equation, for short) one obtains the generating function of a local canonical transformation such that the new Hamiltonian is equal to zero.

In Section 6.1 we present the HJ equation and we give several standard examples of its application, finding complete solutions of the HJ equation by means of the method of separation of variables. In Section 6.1.1 we study the relationship between different complete solutions of the HJ equation. In Section 6.1.2, we consider alternative expressions for the HJ equation, which are useful in some cases, but not usually discussed in the standard textbooks. In Section 6.1.3 we show that in some problems where the method of separation of variables is not applicable, it may be possible to obtain R-separable solutions, which are sums of a fixed function that may depend on all the variables, and separated one-variable functions. In Section 6.2 we give a simple proof and several applications of the Liouville theorem, which enables us to find complete solutions of the HJ equation, making use of an adequate set of constants of motion.

G. F. Torres del Castillo, *An Introduction to Hamiltonian Mechanics*, Birkhäuser Advanced Texts Basler Lehrbücher, https://doi.org/10.1007/978-3-319-95225-3_6

In Section 6.3 we show how to map the solutions of the HJ equation corresponding to a Hamiltonian H into solutions of the HJ equation corresponding to the new Hamiltonian K obtained by a canonical transformation. This mapping is then applied in Section 6.3.1 to find the solutions of the HJ equation with a specified initial condition. In Section 6.4 we show that with any point transformation in the extended configuration space, in which the time may be also transformed, and any Hamiltonian, we can obtain new Hamiltonians such that the solutions of the corresponding HJ equations are related in a simple way.

In Section 6.5 we apply the Lagrangian and the Hamiltonian formalisms to the study of geometrical optics and we show that the HJ equation leads to the eikonal equation.

6.1 The Hamilton–Jacobi Equation

As we have seen in Section 5.2, any real-valued function of $2n + 1$ variables, $F_2(q_i, P_i, t)$, such that

$$\det\left(\frac{\partial^2 F_2}{\partial q_i \partial P_j}\right) \neq 0, \tag{6.1}$$

defines a canonical transformation, $Q_i = Q_i(q_j, p_j, t)$, $P_i = P_i(q_j, p_j, t)$, given implicitly by

$$p_i = \frac{\partial F_2}{\partial q_i}, \qquad Q_i = \frac{\partial F_2}{\partial P_i}. \tag{6.2}$$

Then, the Hamilton equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \tag{6.3}$$

for an arbitrary Hamiltonian, H, are equivalent to

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i}, \tag{6.4}$$

where

$$K = H + \frac{\partial F_2}{\partial t}, \tag{6.5}$$

but Equations (6.4) need not be simpler than (6.3).

However, if we find a generating function F_2 such that the new Hamiltonian, K, is equal to zero, then the equations of motion (6.4) are trivially integrated, yielding $Q_i = \text{const.}$, $P_i = \text{const.}$ (that is, the new canonical coordinates, Q_i, P_i, are $2n$, locally defined, constants of motion) (cf. Proposition 5.42). In that case, by combining (6.5) with the first equations in (6.2), and denoting the generating function F_2 by S, one obtains

$$H\left(q_i, \frac{\partial S}{\partial q_i}, t\right) + \frac{\partial S}{\partial t} = 0. \tag{6.6}$$

Equation (6.6) is a first-order partial differential equation (PDE) for $S(q_i, P_i, t)$, known as the *Hamilton–Jacobi* (HJ) equation, and the function S will be called *Hamilton's principal function*. It should be noted that this equation does not contain the variables P_i explicitly, so that, in order to satisfy the condition

$$\det\left(\frac{\partial^2 S}{\partial q_i \partial P_j}\right) \neq 0 \tag{6.7}$$

[see Equation (6.1)], the function S must contain the n variables P_i as parameters. Any solution of the HJ equation satisfying (6.7) is called a *complete solution.* (A first-order linear PDE possesses a *general* solution that contains an arbitrary *function*. See, e.g., Example 6.29.)

Since the HJ equation does not contain S explicitly, but only its partial derivatives, given a solution, S, of the HJ equation, if c is an arbitrary constant, then $S+c$ is also a solution of the same equation. However, such a trivial constant cannot be one of the n parameters P_i contained in a complete solution because it would produce an entire row or column of zeroes in the matrix (6.7).

As we have seen in Proposition 5.42, making use of the explicit form of the solution of the Hamilton equations, one can find the generating function of a canonical transformation such that the new Hamiltonian is equal to zero (that is, a complete solution of the HJ equation) (see Examples 5.40 and 5.43). What is desirable is to find complete solutions of the HJ equation without knowing beforehand the solution of the Hamilton equations. Unfortunately, we do not have an alternative method to solve the HJ equation in general.

Once we have a complete solution, $S(q_i, P_i, t)$, of the HJ equation, we substitute it into the equations

$$p_i = \frac{\partial S}{\partial q_i}, \qquad Q_i = \frac{\partial S}{\partial P_i} \tag{6.8}$$

[see Equation (6.2)] in order to obtain the canonical transformation generated by S. If we make use of these equations to find Q_i and P_i in terms of (q_j, p_j, t), we obtain $2n$ (functionally independent) constants of motion, and if we use them to express q_i and p_i in terms of (Q_j, P_j, t), we obtain the solution of the Hamilton equations (6.3), using the fact that the solution of the Hamilton equations (6.4), with K being equal to zero, is $Q_i(t) = \text{const.}$, $P_i(t) = \text{const.}$

Example 6.1 (One-dimensional harmonic oscillator). By means of a direct substitution one can readily verify that

$$S(q, P, t) = -\left(\frac{P^2}{2m} + \frac{m\omega^2}{2}q^2\right)\frac{\tan\omega t}{\omega} + Pq\sec\omega t$$

is a solution of the HJ equation

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{m\omega^2}{2}q^2 + \frac{\partial S}{\partial t} = 0, \tag{6.9}$$

which corresponds to the standard Hamiltonian of a one-dimensional harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2$$

[cf. Equation (6.6)]. (This solution of Equation (6.9) will be obtained in Example 6.18, below.) Then, Equations (6.8) yield

$$p = -m\omega q\tan\omega t + P\sec\omega t, \qquad Q = -\frac{P}{m\omega}\tan\omega t + q\sec\omega t$$

and, therefore, the canonical transformation generated by S is given by

$$Q = q\cos\omega t - \frac{p}{m\omega}\sin\omega t, \qquad P = m\omega q\sin\omega t + p\cos\omega t, \tag{6.10}$$

or

$$q = Q\cos\omega t + \frac{P}{m\omega}\sin\omega t, \qquad p = -m\omega Q\sin\omega t + P\cos\omega t. \tag{6.11}$$

According to the discussion above, Equations (6.10) give two constants of motion, while Equations (6.11) give the solution of the Hamilton equations (in this example, Q and P happen to be the values of q and p at $t = 0$, respectively). It may be noticed that, by virtue of the Hamilton equations, $p = m\dot{q}$ and, therefore, the second equation in (6.11) can be obtained by differentiating the first one with respect to the time, but this *is not necessary* (though not wrong, either); the canonical transformation generated by any complete solution of the HJ equation yields the entire solution of the Hamilton equations.

Separation of Variables
The method regularly employed to find complete solutions of the HJ equation (and in most textbooks the only one mentioned) is the method of *separation of variables*. In this method one looks for solutions of Equation (6.6) that can be written as the sum of $n+1$ one-variable functions, $S = S_1(q_1) + S_2(q_2) + \cdots + S_n(q_n) + S_{n+1}(t)$. When the method is applicable, one obtains $n+1$ first-order ODEs (for the functions $S_1, S_2, \ldots, S_{n+1}$), and in the process of separating the variables one has to introduce n constants of separation, which can be taken as the parameters P_i (see the examples below).

Example 6.2 (Particle in a uniform gravitational field). A very simple, but illustrative, example is given by the Hamiltonian

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy, \tag{6.12}$$

corresponding to a particle of mass m in a uniform gravitational field. The HJ equation is given by

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2\right] + mgy + \frac{\partial S}{\partial t} = 0 \tag{6.13}$$

and we look for a separable solution of (6.13), that is, a solution of the form

$$S = A(x) + B(y) + C(t), \tag{6.14}$$

where A, B, and C are real-valued functions of a single variable. Substituting (6.14) into (6.13) we obtain, after rearrangement of the terms,

$$\frac{1}{2m}\left[\left(\frac{\mathrm{d}A}{\mathrm{d}x}\right)^2 + \left(\frac{\mathrm{d}B}{\mathrm{d}y}\right)^2\right] + mgy = -\frac{\mathrm{d}C}{\mathrm{d}t},$$

which must hold for all values of x, y, and t, in some open subset of $\mathbb{R}^3$. The left-hand side of this last equation does not depend on t, while the right-hand side does not depend on x and y; hence, the two sides of the equation do not depend on x, y, or t, and therefore must be equal to some constant, P_1, say. Hence, up to an irrelevant constant term,

$$C(t) = -P_1 t \tag{6.15}$$

and

$$\frac{1}{2m}\left[\left(\frac{\mathrm{d}A}{\mathrm{d}x}\right)^2 + \left(\frac{\mathrm{d}B}{\mathrm{d}y}\right)^2\right] + mgy = P_1.$$

Rewriting the last equation in the form

$$\left(\frac{\mathrm{d}A}{\mathrm{d}x}\right)^2 = 2mP_1 - 2m^2gy - \left(\frac{\mathrm{d}B}{\mathrm{d}y}\right)^2,$$

we obtain an equation such that the left-hand side does not depend on y, and the right-hand side does not depend on x, thus, each side must be a constant. Hence,

$$A(x) = P_2 x, \tag{6.16}$$

where P_2 is a constant, and

$$\frac{\mathrm{d}B}{\mathrm{d}y} = \pm\sqrt{2mP_1 - P_2{}^2 - 2m^2gy}.$$

(In what follows there is no need to consider the two signs in the square root, since we only require *one* complete solution of the HJ equation.) In this manner, we have obtained a solution of the HJ equation (6.13),

$$S(x, y, P_1, P_2, t) = P_2 x + \int \sqrt{2mP_1 - P_2{}^2 - 2m^2gy}\,\mathrm{d}y - P_1 t, \tag{6.17}$$

that contains two parameters (the *constants of separation* P_1 and P_2) which have been identified with the new momenta, in order to emphasize the role of S as the (type F_2) generating function of a canonical transformation.

Making use of Equations (6.8) we have

$$p_x = P_2, \qquad p_y = \sqrt{2mP_1 - P_2{}^2 - 2m^2gy},$$

and

$$Q_1 = -t + \int \frac{m\,\mathrm{d}y}{\sqrt{2mP_1 - P_2{}^2 - 2m^2gy}} = -t - \frac{1}{mg}\sqrt{2mP_1 - P_2{}^2 - 2m^2gy},$$

$$Q_2 = x - \int \frac{P_2\,\mathrm{d}y}{\sqrt{2mP_1 - P_2{}^2 - 2m^2gy}} = x + \frac{P_2}{m^2g}\sqrt{2mP_1 - P_2{}^2 - 2m^2gy}.$$

By combining these last expressions, we obtain the coordinate transformation

$$Q_1 = -t - \frac{p_y}{mg}, \qquad Q_2 = x + \frac{p_x p_y}{m^2 g}, \qquad P_1 = \frac{p_x{}^2 + p_y{}^2}{2m} + mgy, \qquad P_2 = p_x, \tag{6.18}$$

and its inverse

$$x = Q_2 + \frac{P_2(t + Q_1)}{m}, \quad y = \frac{P_1}{mg} - \frac{{P_2}^2}{2m^2 g} - \frac{g}{2}(t + Q_1)^2, \tag{6.19}$$
$$p_x = P_2, \qquad p_y = -mg(t + Q_1).$$

Since the new Hamiltonian is equal to zero, Equations (6.18) give four constants of motion, while Equations (6.19) give the solution of the Hamilton equations in the original variables. (In this example not all of the Q_i and P_i coincide with the initial values of q_i and p_i; however, the constants of motion Q_i and P_i can be expressed in terms of the initial values of q_i and p_i by simply setting $t = 0$, or any other initial value of t, in Equations (6.18).)

For future convenience, it is useful to note that the functions $A(x)$ and $C(t)$, which depend on the variables that do not appear in the Hamiltonian (6.12), are *linear functions* [see Equations (6.15) and (6.16)]. One can convince oneself that this is a general rule: if a coordinate q_K does not appear in the Hamiltonian (but its conjugate momentum, p_K, does appear in H), then, in a separable solution of the HJ equation, the function depending on q_K must be a linear function. Similarly, if t does not appear in the Hamiltonian, then, in a separable solution of the HJ equation, the function depending on t must be a linear function of t.

Example 6.3 (Particle in a central field of force). One of the standard examples of the application of the HJ equation is that of a particle in a central field of force. Using the fact that the orbit must lie on a plane passing through the center of force, we consider a particle moving on the Euclidean plane under the influence of a potential $V(r)$, hence

$$H = \frac{1}{2m}\left({p_r}^2 + \frac{{p_\theta}^2}{r^2}\right) + V(r), \tag{6.20}$$

in terms of the polar coordinates (r, θ) [see (4.7)]. Thus, the HJ equation is given by

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \theta}\right)^2\right] + V(r) + \frac{\partial S}{\partial t} = 0 \tag{6.21}$$

and, taking into account that θ and t do not appear in the Hamiltonian (or in the HJ equation), we look for a separable solution of (6.21) of the form

$$S = A(r) + P_2\theta - P_1 t, \tag{6.22}$$

where A is a real-valued function of a single variable, and P_1, P_2 are separation constants. Substituting (6.22) into (6.21) we obtain

$$\frac{\mathrm{d}A}{\mathrm{d}r} = \pm\sqrt{2m\left[P_1 - V(r) - \frac{{P_2}^2}{2mr^2}\right]}.$$

Thus, we have a solution of the HJ equation (6.21),

$$S(r, \theta, P_1, P_2, t) = \int \sqrt{2m\left[P_1 - V(r) - \frac{{P_2}^2}{2mr^2}\right]}\,\mathrm{d}r + P_2\theta - P_1 t. \qquad (6.23)$$

that contains two parameters (P_1 and P_2), identified with the new momenta.

The canonical transformation generated by S is implicitly given by [see Equations (6.8)]

$$p_r = \frac{\partial S}{\partial r} = \sqrt{2m\left[P_1 - V(r) - \frac{{P_2}^2}{2mr^2}\right]}, \qquad p_\theta = \frac{\partial S}{\partial \theta} = P_2 \qquad (6.24)$$

and

$$Q_1 = \frac{\partial S}{\partial P_1} = -t + \int \frac{m\,\mathrm{d}r}{\sqrt{2m\left[P_1 - V(r) - \frac{{P_2}^2}{2mr^2}\right]}}, \qquad (6.25)$$

$$Q_2 = \frac{\partial S}{\partial P_2} = \theta - \int \frac{P_2}{\sqrt{2m\left[P_1 - V(r) - \frac{{P_2}^2}{2mr^2}\right]}}\frac{\mathrm{d}r}{r^2}. \qquad (6.26)$$

From Equations (6.24) we see that the new momenta are related to the original coordinates by

$$P_1 = \frac{1}{2m}\left({p_r}^2 + \frac{{p_\theta}^2}{r^2}\right) + V(r), \qquad P_2 = p_\theta,$$

that is, P_1 and P_2 are the Hamiltonian and the angular momentum about the origin, respectively. (We already knew that these two quantities are conserved because the Hamiltonian does not depend on t or θ.) Equation (6.26) yields the equation of the orbit.

Equations (6.25) and (6.26) are essentially Equations (2.14) and (2.15), respectively.

Example 6.4 (Kepler problem in parabolic coordinates). The Hamiltonian for the two-dimensional Kepler problem, expressed in parabolic coordinates, is given by [see Equation (4.43)]

$$H = \frac{1}{2m} \frac{{p_u}^2 + {p_v}^2}{u^2 + v^2} - \frac{2k}{u^2 + v^2}.$$

Hence, the corresponding HJ equation is

$$\frac{1}{2m(u^2 + v^2)} \left[\left(\frac{\partial S}{\partial u} \right)^2 + \left(\frac{\partial S}{\partial v} \right)^2 \right] - \frac{2k}{u^2 + v^2} + \frac{\partial S}{\partial t} = 0. \tag{6.27}$$

A separable solution of this equation has the form

$$S = A(u) + B(v) - P_1 t,$$

where A and B are functions of one variable, and P_1 is a separation constant. Substituting this last expression into the HJ equation (6.27), after some rearrangement we obtain

$$\left(\frac{\mathrm{d}A}{\mathrm{d}u} \right)^2 - 2mk - 2mP_1 u^2 = -\left(\frac{\mathrm{d}B}{\mathrm{d}v} \right)^2 + 2mk + 2mP_1 v^2.$$

Since the left-hand side does not depend on v and the right-hand side does not depend on u, both sides must be equal to some constant, P_2, say. Hence,

$$S = \int \sqrt{P_2 + 2mk + 2mP_1 u^2}\, \mathrm{d}u + \int \sqrt{-P_2 + 2mk + 2mP_1 v^2}\, \mathrm{d}v - P_1 t,$$

is a separable solution of the HJ equation which leads to the expressions

$$p_u = \sqrt{P_2 + 2mk + 2mP_1 u^2}, \qquad p_v = \sqrt{-P_2 + 2mk + 2mP_1 v^2}.$$

By combining these two equations one readily finds that $P_1 = H$ and that

$$v^2 {p_u}^2 - u^2 {p_v}^2 = (u^2 + v^2) P_2 + 2mk(v^2 - u^2),$$

hence, making use of (4.40), (4.41), and (4.51),

$$\begin{aligned} P_2 &= \frac{v^2 {p_u}^2 - u^2 {p_v}^2 + 2mk(u^2 - v^2)}{u^2 + v^2} \\ &= \frac{v^2 (u p_x + v p_y)^2 - u^2(-v p_x + u p_y)^2}{u^2 + v^2} + 2mk \frac{x}{r} \\ &= -2x {p_y}^2 + 2y p_x p_y + 2mk \frac{x}{r}, \end{aligned}$$

i.e., the constant of motion P_2 is equal to $-2A_1$, where A_1 is the x-component of the Laplace–Runge–Lenz vector (4.52).

The equation of the orbit is obtained from

$$Q_2 = \frac{\partial S}{\partial P_2} = \int \frac{\mathrm{d}u}{2\sqrt{P_2 + 2mk + 2mP_1u^2}} - \int \frac{\mathrm{d}v}{2\sqrt{-P_2 + 2mk + 2mP_1v^2}}, \tag{6.28}$$

using the fact that Q_2 is a constant of motion, and the dependence of the coordinates on the time is determined by

$$Q_1 = \frac{\partial S}{\partial P_1} = \int \frac{mu^2\mathrm{d}u}{\sqrt{P_2 + 2mk + 2mP_1u^2}} + \int \frac{mv^2\mathrm{d}v}{\sqrt{-P_2 + 2mk + 2mP_1v^2}} - t, \tag{6.29}$$

using the fact that Q_1 is a constant of motion.

In order to obtain the solution of the equations of motion, it is convenient to introduce an auxiliary parameter, τ, in the following way. Since Q_2 is a constant of motion, Equation (6.28) is equivalent to the ODE

$$\frac{\mathrm{d}u}{\sqrt{P_2 + 2mk + 2mP_1u^2}} = \frac{\mathrm{d}v}{\sqrt{-P_2 + 2mk + 2mP_1v^2}}.$$

Introducing the parameter τ by means of

$$\frac{\mathrm{d}\tau}{m} = \frac{\mathrm{d}u}{\sqrt{P_2 + 2mk + 2mP_1u^2}} = \frac{\mathrm{d}v}{\sqrt{-P_2 + 2mk + 2mP_1v^2}}, \tag{6.30}$$

where the constant factor $1/m$ is included in order to get agreement with the definition given in Section 4.3, from (6.29) we have

$$\mathrm{d}t = \frac{mu^2\mathrm{d}u}{\sqrt{P_2 + 2mk + 2mP_1u^2}} + \frac{mv^2\mathrm{d}v}{\sqrt{-P_2 + 2mk + 2mP_1v^2}} = (u^2 + v^2)\,\mathrm{d}\tau, \tag{6.31}$$

which coincides with Equation (4.47).

From Equations (6.30) one can readily get u and v as functions of τ and substituting the expressions thus obtained into (6.31) one obtains the relation between t and τ (see Section 4.3.1).

When the Hamiltonian does not depend explicitly on the time, the HJ equation (6.6) admits partially separable solutions of the form

$$S(q_i, t) = W(q_i) - Et,$$

where E is a separation constant and $W(q_i)$ obeys the equation

$$H\left(q_i, \frac{\partial W}{\partial q_i}, t\right) = E.$$

The function W is known as *Hamilton's characteristic function* and the equation satisfied by W is usually called time-independent Hamilton–Jacobi equation.

Example 6.5 (Charged particle in the field of a point electric dipole). Another well-known example of a Hamiltonian that leads to a separable HJ equation is the one corresponding to a charged particle in the field of a point electric dipole, expressed in spherical coordinates (r, θ, ϕ),

$$H = \frac{1}{2m}\left({p_r}^2 + \frac{{p_\theta}^2}{r^2} + \frac{{p_\phi}^2}{r^2 \sin^2\theta}\right) + \frac{k\cos\theta}{r^2},$$

where k is a constant. The HJ equation is

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \theta}\right)^2 + \frac{1}{r^2\sin^2\theta}\left(\frac{\partial S}{\partial \phi}\right)^2\right] + \frac{k\cos\theta}{r^2} + \frac{\partial S}{\partial t} = 0 \qquad (6.32)$$

and, taking into account that ϕ and t do not appear explicitly in H [but the partial derivatives of S with respect to ϕ and t do appear in (6.32)], we look for a separable solution of this equation of the form

$$S = A(r) + B(\theta) + P_2\phi - P_1 t,$$

where P_1 and P_2 are separation constants. Substituting this expression into (6.32) and multiplying by $2mr^2$ we obtain

$$r^2\left(\frac{\mathrm{d}A}{\mathrm{d}r}\right)^2 + \left(\frac{\mathrm{d}B}{\mathrm{d}\theta}\right)^2 + \frac{{P_2}^2}{\sin^2\theta} + 2mk\cos\theta - 2mP_1 r^2 = 0.$$

Hence,

$$\left(\frac{\mathrm{d}B}{\mathrm{d}\theta}\right)^2 + \frac{{P_2}^2}{\sin^2\theta} + 2mk\cos\theta = P_3 \qquad (6.33)$$

and

$$r^2\left(\frac{\mathrm{d}A}{\mathrm{d}r}\right)^2 - 2mP_1 r^2 = -P_3, \qquad (6.34)$$

where P_3 is a third separation constant. Thus, the HJ equation admits separable solutions given by

$$S = \int\sqrt{2mP_1 - \frac{P_3}{r^2}}\,\mathrm{d}r + \int\sqrt{P_3 - \frac{{P_2}^2}{\sin^2\theta} - 2mk\cos\theta}\,\mathrm{d}\theta + P_2\phi - P_1 t,$$

and the canonical transformation generated by S is implicitly given by

$$p_r = \sqrt{2mP_1 - \frac{P_3}{r^2}}, \qquad p_\theta = \sqrt{P_3 - \frac{{P_2}^2}{\sin^2\theta} - 2mk\cos\theta}, \qquad p_\phi = P_2, \tag{6.35}$$

and

$$Q_1 = -t + \int \frac{m\,\mathrm{d}r}{\sqrt{2mP_1 - \frac{P_3}{r^2}}}, \tag{6.36}$$

$$Q_2 = \phi - \int \frac{P_2\,\mathrm{d}\theta}{\sin^2\theta\sqrt{P_3 - \frac{{P_2}^2}{\sin^2\theta} - 2mk\cos\theta}}, \tag{6.37}$$

$$Q_3 = -\int \frac{\mathrm{d}r}{2r^2\sqrt{2mP_1 - \frac{P_3}{r^2}}} + \int \frac{\mathrm{d}\theta}{2\sqrt{P_3 - \frac{{P_2}^2}{\sin^2\theta} - 2mk\cos\theta}}. \tag{6.38}$$

From Equations (6.35) we find that the new momenta, P_i, are the constants of motion

$$P_1 = H, \qquad P_2 = p_\phi, \qquad P_3 = {p_\theta}^2 + \frac{{p_\phi}^2}{\sin^2\theta} + 2mk\cos\theta.$$

The conservation of P_1 and P_2 are related to the obvious symmetries of the Hamiltonian (i.e., t and ϕ do not appear in the Hamiltonian), while the conservation of P_3 is related to a "hidden" symmetry of H (cf. Exercise 4.22).

With Equations (6.36)–(6.38) the solution of the equations of motion has been reduced to quadratures. It should be kept in mind that Equations (6.35)–(6.38) only give a coordinate transformation, and that the equations of motion are $\dot{Q}_i = 0 = \dot{P}_i$. Thus, for example, Equation (6.36) amounts to the equation of motion

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{\sqrt{2mP_1 - \frac{P_3}{r^2}}}{m},$$

which makes sense also in the case where the constants of motion P_1 (the total energy) and P_3 are equal to zero.

Exercise 6.6. As shown in Exercise 4.18, the Hamiltonian for a particle of mass m moving on the Euclidean plane subject to the gravitational attraction of two fixed centers separated by a distance $2c$ can be written as [see Equation (4.57)]

$$H = \frac{{p_u}^2 + {p_v}^2}{2mc^2(\cosh^2 u - \cos^2 v)} - \frac{(k_1 + k_2)\cosh u + (k_1 - k_2)\cos v}{c\,(\cosh^2 u - \cos^2 v)}. \tag{6.39}$$

Show that the HJ equation admits separable solutions in these coordinates, find the explicit expressions of two constants of motion and reduce to quadratures the equation of the orbit.

Exercise 6.7. Show that the HJ equation for a two-dimensional isotropic harmonic oscillator can be solved by separation of variables in elliptic (or confocal) coordinates (see Exercise 4.18) and identify the constants of motion P_1, P_2.

Exercise 6.8. Show that the HJ equation for the Hamiltonian

$$K = \frac{p^2}{2t} + \frac{q^6}{6t} + \frac{pq}{2t},$$

which is related to the Emden–Fowler equation (see Example 5.17), can be solved by separation of variables and reduce the solution of the Hamilton equations to quadratures. (See also Example 5.18.)

Exercise 6.9. Show that the HJ equation for a Hamiltonian of the form

$$H = \frac{1}{2}\frac{\mathscr{P} {p_x}^2 + \mathscr{Q} {p_y}^2}{X + Y} + \frac{\xi + \eta}{X + Y},$$

where $\mathscr{P}$, X, ξ are functions of x only, and $\mathscr{Q}$, Y, η are functions of y only, can be solved by separation of variables.

A Multiplicatively Separable Solution

In some exceptional cases, the HJ equation admits multiplicatively separable complete solutions. A simple example of this is provided by the HJ equation

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{\partial S}{\partial t} = 0, \tag{6.40}$$

which corresponds to a free particle. Looking for a solution of the form $S(q, t) = A(q)B(t)$ we obtain

$$\frac{1}{2mA}\left(\frac{\mathrm{d}A}{\mathrm{d}q}\right)^2 = -\frac{1}{B^2}\frac{\mathrm{d}B}{\mathrm{d}t}$$

and, in the usual manner, we conclude that both sides of the last equation must be equal to some constant, a, say. Solving the resulting ODEs, we readily obtain

$$S(q, t) = \frac{ma}{2}(q + b)^2 \frac{1}{at + c},$$

where b and c are integration constants. However, rewriting the solution thus obtained in the equivalent form

$$S(q,t) = \frac{m(q+b)^2}{2(t+c/a)}, \tag{6.41}$$

we see that it only depends on two arbitrary constants (b and c/a).

Setting $c/a = 0$ or $b = 0$ (but not both) in (6.41) we obtain a complete solution of the HJ equation (6.40). This shows that the frequently found assertion that any additional constant in a complete solution of the HJ equation must be an additive constant is wrong. (See also Example 6.10, below.)

6.1.1 Relation Between Complete Solutions of the HJ Equation

The HJ equation for a given Hamiltonian, as any other first-order PDE, possesses an infinite number of complete solutions. As we shall show now, any two complete solutions of the HJ equation are related by means of a time-independent canonical transformation [cf. Equation (5.99)]. Indeed, if $S(q_i, P_i, t)$ is a complete solution of the HJ equation corresponding to a Hamiltonian $H(q_i, p_i, t)$, we have

$$p_i \mathrm{d}q_i - H \mathrm{d}t + Q_i \mathrm{d}P_i = \mathrm{d}S$$

[see Equation (5.54)]. Similarly, if $\tilde{S}(q_i, \tilde{P}_i, t)$ is any other complete solution of the same equation (in the same coordinates q_i), then

$$p_i \mathrm{d}q_i - H \mathrm{d}t + \tilde{Q}_i \mathrm{d}\tilde{P}_i = \mathrm{d}\tilde{S},$$

hence,

$$Q_i \mathrm{d}P_i - \tilde{Q}_i \mathrm{d}\tilde{P}_i = \mathrm{d}F, \tag{6.42}$$

where

$$F \equiv S - \tilde{S}. \tag{6.43}$$

Equation (6.42) explicitly shows that (Q_i, P_i) and $(\tilde{Q}_i, \tilde{P}_i)$ are related by means of a time-independent canonical transformation [cf. Equation (5.46)]. *If* the set $(P_i, \tilde{P}_i)$ is functionally independent, then the function F defined in (6.43) is a generating function of this canonical transformation.

Since $p_i = \partial S/\partial q_i$ and, also, $p_i = \partial \tilde{S}/\partial q_i$, we have

$$\frac{\partial (S - \tilde{S})}{\partial q_i} = 0. \tag{6.44}$$

Making use of these n conditions one can eliminate the q_i from the right-hand side of (6.43) (the dependence on t automatically disappears as a consequence of (6.44); no additional conditions come from $\partial(S - \tilde{S})/\partial t = 0$ since, by hypothesis, S and $\tilde{S}$ satisfy the same HJ equation).

Similarly, given a complete solution, $S(q_i, P_i, t)$, of the HJ equation for some Hamiltonian, and a function $F(P_i, \tilde{P}_i)$ that defines a canonical transformation,

$$\tilde{S}(q_i, \tilde{P}_i, t) = S(q_i, P_i, t) - F(P_i, \tilde{P}_i), \tag{6.45}$$

is also a complete solution of the same HJ equation. The dependence on the parameters P_i is eliminated from the expression on the right-hand side of (6.45) with the aid of the n conditions

$$\frac{\partial (S - F)}{\partial P_i} = 0. \tag{6.46}$$

(Cf. Calkin [2, pp. 148–150].)

Example 6.10. As pointed out in Example 6.1, the function

$$S(q, P, t) = -\left(\frac{P^2}{2m} + \frac{m\omega^2}{2} q^2\right) \frac{\tan \omega t}{\omega} + Pq \sec \omega t$$

is a complete solution of the HJ equation in the case of the standard Hamiltonian of a one-dimensional harmonic oscillator. With the aid of the function

$$F(P, \tilde{P}) = \frac{P^2}{2m\omega} \tan \tilde{P},$$

we can obtain another complete solution of the same HJ equation. Indeed, the condition (6.46) reads

$$0 = \frac{\partial (S - F)}{\partial P} = -\frac{P}{m\omega} \tan \omega t + q \sec \omega t - \frac{P}{m\omega} \tan \tilde{P},$$

that is,

$$P = \frac{m\omega q \sec \omega t}{\tan \omega t + \tan \tilde{P}}$$

and, therefore,

$$\tilde{S} = S - F = \frac{m\omega q^2}{2} \cot(\omega t + \tilde{P})$$

is also a complete solution of the HJ equation for the standard Hamiltonian of a one-dimensional harmonic oscillator. (It may be noticed that this solution is the *product* of separated functions of q and t.)

Exercise 6.11. Find a generating function of the canonical transformation that leads from the complete, separable, solution of the HJ equation

$$S(x, y, P_1, P_2, t) = P_1 x + P_2 y - \frac{{P_1}^2 + {P_2}^2}{2m} t$$

to the non-separable complete solution

$$\tilde{S}(x, y, \tilde{P}_1, \tilde{P}_2, t) = \frac{m}{2t}\big[(x - \tilde{P}_1)^2 + (y - \tilde{P}_2)^2\big].$$

What is the Hamiltonian?

Since any complete solution of the HJ equation leads to the solution of the Hamilton equations, it is not necessary to find a second complete solution of the HJ equation. In the context of classical mechanics, we make use of a complete solution of the HJ equation only as a means to find the solution of the Hamilton equations. However, in geometrical optics the function S is interesting in itself and it is highly relevant to find different solutions of the appropriate version of the HJ equation, which correspond to different trains of wavefronts (see Section 6.5, below).

Other Special Generating Functions

In the same manner as we can look for a canonical transformation that produces a new Hamiltonian equal to zero, we can also look for canonical transformations that take, locally, a given Hamiltonian H into any other specified Hamiltonian K (the only restriction is that the number of degrees of freedom in both Hamiltonians be the same). Making use again of Equations (6.2) we find that the required generating function must satisfy the PDE

$$K\left(\frac{\partial F_2}{\partial P_i}, P_i, t\right) = H\left(q_i, \frac{\partial F_2}{\partial q_i}, t\right) + \frac{\partial F_2}{\partial t} \tag{6.47}$$

and the condition (6.1). For instance, in the case of the Hamiltonians

$$H(q, p, t) = \frac{p^2}{2m} + mgq, \qquad K(Q, P, t) = \frac{P^2}{2m},$$

where g is a constant, one can verify that the function

$$F_2(q, P, t) = (P - mgt)q + \frac{1}{2}gt^2 P - \frac{m}{6}g^2 t^3$$

satisfies (6.47) and (6.1). In fact, substituting this expression into (6.2) one finds that F_2 generates the canonical transformation

$$q = Q - \tfrac{1}{2}gt^2, \qquad p = P - mgt, \tag{6.48}$$

and that H and K are related by (6.5). Hence, the coordinate transformation (6.48) maps the phase space trajectories of a free particle into those of a particle in free fall. Cf. Exercise 5.6. It may be noticed that, when $g = 0$, the transformation (6.48) reduces to the identity. (The coordinate transformation (6.48) gives the relation between two reference frames, one of which has an acceleration equal to g with respect to the other.)

These transformations can be constructed by finding separately solutions of the HJ equations for H and K, and combining them or composing the coordinate transformations generated by them (see also Section 6.3).

6.1.2 Alternative Expressions of the HJ Equation

It should be clear that, in order to find new canonical coordinates with a Hamiltonian equal to zero, instead of a (type F_2) generating function $S(q_i, P_i, t)$, that depends on the n original coordinates q_i, we can also look for generating functions that depend on other combinations of the original variables, replacing one or several coordinates q_i by its conjugate momentum p_i, which gives a total of 2^n different possibilities. Some of these alternatives are sometimes mentioned [e.g., Corben and Stehle ([4], Sect. 61), Greenwood [9, Sect. 6–1]], without actually using them, claiming that they are of little interest. However, in some cases, the dependence of H on the coordinates may be simpler than that on the momenta.

Example 6.12 (Particle in a uniform gravitational field). Let us consider again the simple case corresponding to the Hamiltonian

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy,$$

where m and g are constants. We look for a type F_4 generating function (which we shall denote also by S) such that $K = 0$. According to Equations (5.57), S must be a complete solution of

$$\frac{{p_x}^2 + {p_y}^2}{2m} - mg\frac{\partial S}{\partial p_y} + \frac{\partial S}{\partial t} = 0, \tag{6.49}$$

and the canonical transformation is implicitly given by

$$x = -\frac{\partial S}{\partial p_x}, \qquad y = -\frac{\partial S}{\partial p_y}, \qquad Q_i = \frac{\partial S}{\partial P_i}. \tag{6.50}$$

Since H does not depend on t, we look for a solution of (6.49) of the form

$$S = W(p_x, p_y) - P_1 t,$$

where P_1 is a separation constant. Then, the "characteristic function," W, has to satisfy

$$\frac{\partial W}{\partial p_y} = -\frac{P_1}{mg} + \frac{{p_x}^2 + {p_y}^2}{2m^2 g}.$$

The *general* solution of this PDE is readily found:

$$W = -\frac{P_1 p_y}{mg} + \frac{{p_x}^2 p_y}{2m^2 g} + \frac{{p_y}^3}{6m^2 g} + f(p_x),$$

where $f(p_x)$ is an *arbitrary function* of p_x only. Choosing $f(p_x) = P_2 p_x$, where P_2 is a constant, we obtain

$$S(p_x, p_y, P_1, P_2, t) = -\frac{P_1 p_y}{mg} + \frac{{p_x}^2 p_y}{2m^2 g} + \frac{{p_y}^3}{6m^2 g} + P_2 p_x - P_1 t.$$

It may be noticed that this solution of the HJ equation (6.49) is not the sum of separate functions of p_x, p_y and t. (This is an example of an R-separable solution, to be discussed in Section 6.1.3.) According to (6.50), S generates the canonical transformation given by

$$x = -\frac{p_x p_y}{m^2 g} - P_2, \qquad y = \frac{P_1}{mg} - \frac{{p_x}^2 + {p_y}^2}{2m^2 g}, \qquad Q_1 = -\frac{p_y}{mg} - t, \qquad Q_2 = p_x,$$

i.e.,

$$Q_1 = -\frac{p_y}{mg} - t, \quad Q_2 = p_x, \quad P_1 = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy, \quad P_2 = -\frac{p_x p_y}{m^2 g} - x,$$

which are the constants of motion obtained in Example 6.2. (See also Example 6.24.) The solution of the Hamilton equations in the original variables is obtained writing x, y, p_x, p_y in terms of Q_i, P_i.

6.1.3 R-Separable Solutions of the HJ Equation

As pointed out above, the method commonly employed to solve the HJ equation is the method of separation of variables, but, in many cases, this method may not work. For example, in the case of the Hamiltonian

$$H = \frac{p^2}{2m} - ktq, \tag{6.51}$$

which corresponds to a particle of mass m subjected to a variable force kt, where k is some constant, the HJ equation

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 - ktq + \frac{\partial S}{\partial t} = 0 \tag{6.52}$$

does *not* admit separable solutions owing to the presence of the term ktq. However, noting that the last two terms on the left-hand side of this equation can be written as

$$-ktq + \frac{\partial S}{\partial t} = \frac{\partial}{\partial t}\left(-\frac{kt^2 q}{2} + S\right),$$

we introduce $\widetilde{S} \equiv S - kt^2 q/2$, and we find that (6.52) amounts to

$$\frac{1}{2m}\left(\frac{\partial \widetilde{S}}{\partial q} + \frac{kt^2}{2}\right)^2 + \frac{\partial \widetilde{S}}{\partial t} = 0.$$

By contrast with (6.52), this last equation admits separable solutions, and since q does not appear explicitly in the equation, its separable solutions are of the form $\widetilde{S} = Pq + F(t)$, where P is a separation constant, with

$$\frac{1}{2m}\left(P + \frac{kt^2}{2}\right)^2 + \frac{\mathrm{d}F}{\mathrm{d}t} = 0.$$

Thus,

$$S = \frac{kt^2 q}{2} + Pq - \frac{1}{2m}\left(P^2 t + \frac{1}{3}Pkt^3 + \frac{1}{20}k^2 t^5\right) \tag{6.53}$$

is a (complete) solution of (6.52). This solution is the sum of a function of q and t (the term $kt^2 q/2$), that does not contain the separation constant P, and one-variable functions. Such solutions are called *R-separable solutions*.

Thus, we have the generating function of a canonical transformation implicitly given by [see Equations (6.8)],

$$Q = \frac{\partial S}{\partial P} = q - \frac{Pt}{m} - \frac{kt^3}{6m}, \qquad p = \frac{\partial S}{\partial q} = \frac{kt^2}{2} + P.$$

The original variables are given by

$$p = P + \frac{kt^2}{2}, \qquad q = Q + \frac{Pt}{m} + \frac{kt^3}{6m}.$$

Since Q and P are constants of motion, these expressions constitute the solution of the Hamilton equations. From these expressions we see that Q and P correspond to the values of q and p at $t = 0$, respectively. (Cf. Example 5.40.)

Example 6.13 (Charged particle in a uniform magnetic field). The Hamiltonian

$$H = \frac{1}{2m}\left[\left(p_x + \frac{eB_0 y}{2c}\right)^2 + \left(p_y - \frac{eB_0 x}{2c}\right)^2 + {p_z}^2\right], \tag{6.54}$$

corresponds to a charged particle of mass m and electric charge e in a uniform magnetic field $\mathbf{B} = B_0\mathbf{k}$, if the vector potential is chosen according to the rule $\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}$, which is applicable for a uniform magnetic field $\mathbf{B}$ [see (4.14)]. The resulting HJ equation

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial x} + \frac{eB_0 y}{2c}\right)^2 + \left(\frac{\partial S}{\partial y} - \frac{eB_0 x}{2c}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2\right] + \frac{\partial S}{\partial t} = 0 \tag{6.55}$$

does not admit separable solutions but, letting $\widetilde{S} \equiv S + eB_0 xy/2c$, we obtain

$$\frac{1}{2m}\left[\left(\frac{\partial \widetilde{S}}{\partial x}\right)^2 + \left(\frac{\partial \widetilde{S}}{\partial y} - \frac{eB_0 x}{c}\right)^2 + \left(\frac{\partial \widetilde{S}}{\partial z}\right)^2\right] + \frac{\partial \widetilde{S}}{\partial t} = 0. \tag{6.56}$$

Since y, z, and t do not appear explicitly in this equation, it admits separable solutions of the form $\widetilde{S} = F(x) + P_2 y + P_3 z - P_1 t$, where P_1, P_2, and P_3 are constants, and F satisfies the separated equation

$$\frac{1}{2m}\left[\left(\frac{\mathrm{d}F}{\mathrm{d}x}\right)^2 + \left(P_2 - \frac{eB_0 x}{c}\right)^2 + {P_3}^2\right] = P_1.$$

Thus,

$$S = -\frac{eB_0 xy}{2c} + P_2 y + P_3 z - P_1 t + \int \sqrt{2mP_1 - {P_3}^2 - \left(P_2 - eB_0 x/c\right)^2}\,\mathrm{d}x$$

and making use of Equations (6.8) we obtain

$$p_x = -\frac{eB_0 y}{2c} + \sqrt{2mP_1 - P_3{}^2 - \big(P_2 - eB_0 x/c\big)^2}, \quad p_y = -\frac{eB_0 x}{2c} + P_2, \quad p_z = P_3,$$

and

$$\begin{aligned} Q_1 &= -t + \int \frac{m\,\mathrm{d}x}{\sqrt{2mP_1 - P_3{}^2 - \big(P_2 - eB_0 x/c\big)^2}}, \\ Q_2 &= y - \int \frac{(P_2 - eB_0 x/c)\,\mathrm{d}x}{\sqrt{2mP_1 - P_3{}^2 - \big(P_2 - eB_0 x/c\big)^2}}, \\ Q_3 &= z - \int \frac{P_3\,\mathrm{d}x}{\sqrt{2mP_1 - P_3{}^2 - \big(P_2 - eB_0 x/c\big)^2}}. \end{aligned} \tag{6.57}$$

By combining these equations one obtains the constants of motion

$$P_1 = H, \qquad P_2 = p_y + \frac{eB_0 x}{2c}, \qquad P_3 = p_z, \qquad Q_2 = -\frac{c}{eB_0}\left(p_x - \frac{eB_0 y}{2c}\right).$$

With the aid of the change of variable

$$P_2 - \frac{eB_0 x}{c} = \sqrt{2mP_1 - P_3{}^2}\,\cos\theta,$$

from the first two equations in (6.57) we obtain

$$x = \frac{cP_2}{eB_0} - \frac{c}{eB_0}\sqrt{2mP_1 - P_3{}^2}\,\cos\omega_c(t + Q_1),$$

where $\omega_c \equiv eB_0/mc$ (the cyclotron frequency), and

$$y = Q_2 + \frac{c}{eB_0}\sqrt{2mP_1 - P_3{}^2}\,\sin\omega_c(t + Q_1),$$

respectively. These last expressions show that the projection of the orbit on the xy-plane is a circle whose center and radius are given in terms of the constants of motion Q_i and P_i (cf. Example 1.19).

The vector potential $\mathbf{A}' = B_0 x\,\mathbf{j}$ also yields the uniform magnetic field $\mathbf{B} = B_0\,\mathbf{k}$, and leads to a separable HJ equation. (In fact, the difference $\mathbf{A}' - \mathbf{A} = B_0 x\,\mathbf{j} - \frac{1}{2}B_0(-y\,\mathbf{i} + x\,\mathbf{j}) = \frac{1}{2}B_0(y\,\mathbf{i} + x\,\mathbf{j})$, which is the gradient of $\frac{1}{2}B_0 xy$.) (See also the discussion at the end of this section.)

Exercise 6.14. Show that the HJ equation for a charged particle in a uniform magnetic field, with $\mathbf{A} = \frac{1}{2}\mathbf{B}\times\mathbf{r}$, can be solved by separation of variables in circular cylindrical coordinates and identify the new momenta.

Example 6.15. We consider the HJ equation

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2\right] + mgy + \frac{\partial S}{\partial t} = 0, \tag{6.58}$$

which *does* admit separable solutions as a consequence of the fact that x and t do not appear explicitly in the equation (see Example 6.2). Noting that

$$mgy + \frac{\partial S}{\partial t} = \frac{\partial}{\partial t}(mgyt + S),$$

we introduce $\widetilde{S} \equiv S + mgyt$, and we have

$$\frac{1}{2m}\left[\left(\frac{\partial \widetilde{S}}{\partial x}\right)^2 + \left(\frac{\partial \widetilde{S}}{\partial y} - mgt\right)^2\right] + \frac{\partial \widetilde{S}}{\partial t} = 0. \tag{6.59}$$

This equation admits separable solutions as a consequence of the fact that x and y do not appear explicitly in it. Indeed, looking for solutions of the form

$$\widetilde{S} = P_1 x + P_2 y + F(t),$$

where P_1 and P_2 are constants, we obtain

$$\frac{1}{2m}\left[{P_1}^2 + (P_2 - mgt)^2\right] + \frac{\mathrm{d}F}{\mathrm{d}t} = 0.$$

Hence,

$$F(t) = -\frac{1}{2m}\left[({P_1}^2 + {P_2}^2)t - P_2 mgt^2 + \frac{1}{3}m^2g^2t^3\right]$$

and, therefore,

$$S = -mgyt + P_1 x + P_2 y - \frac{1}{2m}\left[\left({P_1}^2 + {P_2}^2\right)t - P_2 mgt^2 + \frac{1}{3}m^2g^2t^3\right] \tag{6.60}$$

is a complete R-separable solution of (6.58). Thus, the HJ equation (6.58) admits both separable and R-separable solutions in the Cartesian coordinates (x, y).

Substitution of (6.60) into Equations (6.8) yields the canonical transformation

$$p_x = \frac{\partial S}{\partial x} = P_1, \qquad p_y = \frac{\partial S}{\partial y} = -mgt + P_2,$$

and

$$Q_1 = \frac{\partial S}{\partial P_1} = x - \frac{P_1 t}{m}, \qquad Q_2 = \frac{\partial S}{\partial P_2} = y - \frac{P_2 t}{m} + \frac{1}{2} g t^2.$$

That is, we have four constants of motion

$$Q_1 = x - \frac{t p_x}{m}, \qquad Q_2 = y - \frac{t p_y}{m} - \frac{1}{2} g t^2, \qquad P_1 = p_x, \qquad P_2 = p_y + mgt,$$

and the solution of the Hamilton equations

$$x = Q_1 + \frac{t P_1}{m}, \qquad y = Q_2 + \frac{t P_2}{m} - \frac{1}{2} g t^2, \qquad p_x = P_1, \qquad p_y = P_2 - mgt.$$

The constants of motion Q_1, Q_2, P_1, and P_2 correspond to the values at $t = 0$ of x, y, p_x, and p_y, respectively.

It may be noticed that finding R-separable solutions of the HJ equation for a Hamiltonian H is equivalent to finding *separable* solutions of the HJ equation for another Hamiltonian, H', obtained from H by means of a canonical transformation of the form

$$q'_i = q_i, \qquad p'_i = p_i + \frac{\partial R}{\partial q_i}, \tag{6.61}$$

where R is a function of q_i and t only [see (5.67)]. For instance, Equation (6.59) is the HJ equation corresponding to the Hamiltonian (5.69), which is obtained from (5.68) by means of the canonical transformation (6.61) with $R = mgyt$.

Exercise 6.16. Show that if the Hamiltonian has the form

$$H = \frac{p^2}{2m} - \phi(t) q,$$

where $\phi(t)$ is a given function of t only, then the corresponding HJ equation admits R-separable complete solutions. This result is applicable to the problem of a rocket in a uniform gravitational field, for which the Hamiltonian can be taken as

$$H = \frac{p^2}{2} + \left(u \frac{\mathrm{d} \ln m}{\mathrm{d}t} + g \right) q,$$

where $m(t)$ is the mass of the rocket at time t and u is the speed of the exhaust gases with respect to the rocket (see Example 2.13).

6.2 The Liouville Theorem on Solutions of the HJ Equation

Apart from the method of separation of variables, there exist some other methods for solving first-order PDEs (see, e.g., Sneddon [14]). In one of these lesser-known methods, when applied to the HJ equation, one has to express the canonical momenta in terms of the coordinates and n constants of motion; a complete solution, S, of the HJ equation can then be obtained from $\mathrm{d}S = p_i \mathrm{d}q_i - H\mathrm{d}t$. However, it turns out that $p_i \mathrm{d}q_i - H\mathrm{d}t$ is an exact differential if and only if the constants of motion employed in this process are in involution, that is, their Poisson brackets are all equal to zero, and this result is known as Liouville's Theorem. In the case where there is only one degree of freedom, the Liouville Theorem can be applied making use of a single arbitrary constant of motion, since the Poisson bracket of a function with itself is trivially equal to zero.

The application of the Liouville theorem requires the knowledge of n constants of motion in involution, but is not linked to some specific coordinate system; the complete solutions of the HJ equation obtained in this manner need not be separable or R-separable.

Proposition 6.17 (Liouville's Theorem). *If $P_i = P_i(q_j, p_j, t)$, $i = 1, 2, \ldots, n$, are n functionally independent constants of motion in involution such that the momenta p_i can be written in terms of q_j, P_j, and t, then, locally, there exists a function $S(q_i, t)$, depending parametrically on the P_i, such that*

$$p_i(q_j, P_j, t)\,\mathrm{d}q_i - H\big(q_i, p_i(q_j, P_j, t), t\big)\,\mathrm{d}t = \mathrm{d}S \tag{6.62}$$

and S is a complete *solution of the HJ equation.*

Note that if the momenta p_i can be written in terms of q_j, P_j, and t, then the constants of motion $P_i = P_i(q_j, p_j, t)$, $i = 1, 2, \ldots, n$, have to be functionally independent.

Proof. According to the hypotheses, from the expressions $P_i = P_i(q_j, p_j, t)$, we can find the p_i as functions of q_j, P_j, and t, hence

$$\det\left(\frac{\partial P_i}{\partial p_j}\right) \neq 0. \tag{6.63}$$

Substituting

$$\mathrm{d}p_k = \frac{\partial p_k}{\partial q_j}\mathrm{d}q_j + \frac{\partial p_k}{\partial P_j}\mathrm{d}P_j + \frac{\partial p_k}{\partial t}\mathrm{d}t$$

into

$$\mathrm{d}P_i = \frac{\partial P_i}{\partial q_j}\mathrm{d}q_j + \frac{\partial P_i}{\partial p_j}\mathrm{d}p_j + \frac{\partial P_i}{\partial t}\mathrm{d}t$$

one obtains

$$\mathrm{d}P_i = \frac{\partial P_i}{\partial q_j}\mathrm{d}q_j + \frac{\partial P_i}{\partial p_k}\left(\frac{\partial p_k}{\partial q_j}\mathrm{d}q_j + \frac{\partial p_k}{\partial P_j}\mathrm{d}P_j + \frac{\partial p_k}{\partial t}\mathrm{d}t\right) + \frac{\partial P_i}{\partial t}\mathrm{d}t$$

which implies the *identities*

$$\frac{\partial P_i}{\partial q_j} = -\frac{\partial P_i}{\partial p_k}\frac{\partial p_k}{\partial q_j}, \qquad \frac{\partial P_i}{\partial t} = -\frac{\partial P_i}{\partial p_k}\frac{\partial p_k}{\partial t} \tag{6.64}$$

and

$$\frac{\partial P_i}{\partial p_k}\frac{\partial p_k}{\partial P_j} = \delta_{ij}. \tag{6.65}$$

(Note that in the partial derivatives of the P_i, P_i is a function of q_j, p_j, and t, that is

$$\frac{\partial P_i}{\partial q_j} = \left(\frac{\partial P_i}{\partial q_j}\right)_{q,p,t}, \qquad \frac{\partial P_i}{\partial p_j} = \left(\frac{\partial P_i}{\partial p_j}\right)_{q,p,t},$$

while in the partial derivatives of the p_i, p_i is a function of q_j, P_j, and t,

$$\frac{\partial p_i}{\partial P_j} = \left(\frac{\partial p_i}{\partial P_j}\right)_{q,P,t}, \qquad \frac{\partial p_i}{\partial q_j} = \left(\frac{\partial p_i}{\partial q_j}\right)_{q,P,t}.$$

Equations (6.64) and (6.65) can also be derived with the aid of the chain rule.)

Making use of the first equation in (6.64) we find that

$$\begin{aligned}
\{P_i, P_j\} &= \frac{\partial P_i}{\partial q_k}\frac{\partial P_j}{\partial p_k} - \frac{\partial P_j}{\partial q_k}\frac{\partial P_i}{\partial p_k} \\
&= -\frac{\partial P_i}{\partial p_m}\frac{\partial p_m}{\partial q_k}\frac{\partial P_j}{\partial p_k} + \frac{\partial P_j}{\partial p_m}\frac{\partial p_m}{\partial q_k}\frac{\partial P_i}{\partial p_k} \\
&= -\frac{\partial P_i}{\partial p_m}\frac{\partial p_m}{\partial q_k}\frac{\partial P_j}{\partial p_k} + \frac{\partial P_j}{\partial p_k}\frac{\partial p_k}{\partial q_m}\frac{\partial P_i}{\partial p_m} \\
&= \frac{\partial P_i}{\partial p_m}\frac{\partial P_j}{\partial p_k}\left(\frac{\partial p_k}{\partial q_m} - \frac{\partial p_m}{\partial q_k}\right).
\end{aligned}$$

Thus, taking into account (6.63), it follows that $\{P_i, P_j\} = 0$ if and only if

$$\frac{\partial p_k}{\partial q_m} = \frac{\partial p_m}{\partial q_k}. \tag{6.66}$$

On the other hand, making use of the fact that each P_i is a constant of motion, and of (6.64) and (6.66),

$$
\begin{aligned}
0 &= \frac{\partial P_i}{\partial t} + \frac{\partial P_i}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial H}{\partial q_j}\frac{\partial P_i}{\partial p_j} \\
&= -\frac{\partial P_i}{\partial p_k}\frac{\partial p_k}{\partial t} - \frac{\partial P_i}{\partial p_k}\frac{\partial p_k}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial H}{\partial q_j}\frac{\partial P_i}{\partial p_j} \\
&= -\frac{\partial P_i}{\partial p_k}\left(\frac{\partial p_k}{\partial t} + \frac{\partial p_k}{\partial q_j}\frac{\partial H}{\partial p_j} + \frac{\partial H}{\partial q_k}\right) \\
&= -\frac{\partial P_i}{\partial p_k}\left(\frac{\partial p_k}{\partial t} + \frac{\partial p_j}{\partial q_k}\frac{\partial H}{\partial p_j} + \frac{\partial H}{\partial q_k}\right).
\end{aligned}
$$

As a consequence of (6.63) and the chain rule, this amounts to

$$
\frac{\partial p_k}{\partial t} = -\frac{\partial H}{\partial q_k} - \frac{\partial H}{\partial p_j}\frac{\partial p_j}{\partial q_k} = -\frac{\partial}{\partial q_k} H\big(q_j, p_j(q_k, P_k, t), t\big),
$$

and these conditions together with (6.66) imply that the left-hand side of (6.62) is locally exact.

Finally, from (6.62) it follows that S is a solution of the HJ equation, which is complete by virtue of (6.63), in fact,

$$
\det\left(\frac{\partial^2 S}{\partial P_i \partial q_j}\right) = \det\left(\frac{\partial p_j}{\partial P_i}\right) = \left[\det\left(\frac{\partial P_i}{\partial p_j}\right)\right]^{-1} \neq 0.
$$

□

In some textbooks this result is called *Liouville's integrability theorem*.

Example 6.18. One can readily verify that the function $P = m\omega q \sin \omega t + p \cos \omega t$ is a constant of motion for the one-dimensional harmonic oscillator, that is, if

$$
H = \frac{p^2}{2m} + \frac{m\omega^2}{2} q^2.
$$

Since, $p = P \sec \omega t - m\omega q \tan \omega t$, expressing $p\mathrm{d}q - H\mathrm{d}t$ in terms of q, P, t, and treating P as a parameter, we obtain

$$
\begin{aligned}
&p\mathrm{d}q - H\mathrm{d}t \\
&\quad = \big(P \sec \omega t - m\omega q \tan \omega t\big)\mathrm{d}q - \left[\frac{(P \sec \omega t - m\omega q \tan \omega t)^2}{2m} + \frac{m\omega^2}{2} q^2\right]\mathrm{d}t \\
&\quad = \mathrm{d}\left(Pq \sec \omega t - \frac{1}{2} m\omega q^2 \tan \omega t\right) - \omega P q \sec \omega t \tan \omega t \,\mathrm{d}t + \frac{1}{2} m\omega^2 q^2 \sec^2 \omega t \,\mathrm{d}t
\end{aligned}
$$

$$- \left[\frac{(P \sec \omega t - m\omega q \tan \omega t)^2}{2m} + \frac{m\omega^2}{2} q^2 \right] \mathrm{d}t$$

$$= \mathrm{d} \left(Pq \sec \omega t - \frac{1}{2} m\omega q^2 \tan \omega t - \frac{P^2}{2m\omega} \tan \omega t \right).$$

According to Proposition 6.17, the expression inside the parenthesis must be a complete solution of the HJ equation for the Hamiltonian H and, following the standard procedure, it can be used to find the solution of the equations of motion.

Example 6.19. The HJ equation for the Kepler problem in two dimensions, which corresponds to the Hamiltonian

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} - \frac{k}{\sqrt{x^2 + y^2}},$$

expressed in Cartesian coordinates x, y, where m is the mass of the particle and k is a positive constant, is separable in polar and parabolic coordinates (see Examples 6.3 and 6.4, respectively) but *is not* separable in Cartesian coordinates.

Since H is time-independent and invariant under rotations about the origin,

$$P_1 \equiv H, \qquad P_2 \equiv x p_y - y p_x$$

(the total energy and the angular momentum about the origin) are constants of motion, which are in involution (as can be seen from the fact that the angular momentum is a constant of motion). Inverting these expressions one finds

$$p_x = \frac{-P_2 y \pm x\sqrt{2mP_1 r^2 + 2mkr - {P_2}^2}}{r^2}, \qquad p_y = \frac{P_2 x \pm y\sqrt{2mP_1 r^2 + 2mkr - {P_2}^2}}{r^2},$$

where the signs in front of the square roots have to be chosen both plus or both minus and $r^2 \equiv x^2 + y^2$, which give the p_i in terms of q_j and P_j. Thus, the left-hand side of Equation (6.62) becomes

$$P_2 \frac{(-y\mathrm{d}x + x\mathrm{d}y)}{r^2} \pm \frac{\sqrt{2mP_1 r^2 + 2mkr - {P_2}^2}}{r^2} (x\mathrm{d}x + y\mathrm{d}y) - P_1 \mathrm{d}t$$

or, equivalently,

$$P_2 \, \mathrm{d} \left(\arctan \frac{y}{x} \right) \pm \sqrt{2mP_1 + \frac{2mk}{r} - \frac{{P_2}^2}{r^2}} \, \mathrm{d}r - P_1 \mathrm{d}t.$$

This last expression is indeed, *locally*, the differential of a function, which must be a complete solution of the HJ equation. It may be noticed that this function turns out to be the sum of separate functions of the polar coordinates θ, r, though we started with the Hamiltonian in Cartesian coordinates.

Example 6.20. Another simple example is given by the Hamiltonian

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy, \tag{6.67}$$

for which all constants of motion are readily obtained (here m and g are constants). In fact, the HJ equation is separable in the coordinates (x, y) of the configuration space, and the separation constants are the values of H and p_x, which are constants of motion as a consequence of the fact that t and x do not appear in the Hamiltonian (see Example 6.2).

Another constant of motion (which is related to a "hidden" symmetry of the Hamiltonian) is

$$\frac{p_x p_y}{m} + mgx \tag{6.68}$$

(see, e.g., Example 5.35), therefore, H and $\frac{1}{m} p_x p_y + mgx$ are in involution and can be taken as P_1 and P_2, respectively. A straightforward computation leads to the expressions

$$\begin{aligned} p_x + p_y &= \pm\sqrt{2m(P_1 + P_2) - 2m^2 g\,(x + y)}, \\ p_x - p_y &= \pm\sqrt{2m(P_1 - P_2) + 2m^2 g\,(x - y)}, \end{aligned} \tag{6.69}$$

where the signs in front of the square roots have to be chosen both plus or both minus. Hence, by writing Equation (6.62) in the form

$$\tfrac{1}{2}(p_x + p_y)\,\mathrm{d}(x + y) + \tfrac{1}{2}(p_x - p_y)\,\mathrm{d}(x - y) - P_1 \mathrm{d}t = \mathrm{d}S, \tag{6.70}$$

and taking into account that $p_x \pm p_y$ is a function of $x \pm y$ only, we see that the function S is the sum of three one-variable functions that depend on $x + y$, $x - y$, and t (with P_1 and P_2 being treated as parameters). In other words, the HJ equation corresponding to the Hamiltonian (6.67) admits separable solutions in the coordinates (u, v) defined by

$$u \equiv x + y, \qquad v \equiv x - y,$$

and the separation constants are the values of H and $\frac{1}{m} p_x p_y + mgx$.

Exercise 6.21. Find a complete solution of the HJ equation corresponding to the Hamiltonian (6.67) making use of the constants of motion in involution $P_1 = p_x$ and $P_2 = p_y + mgt$.

Exercise 6.22. Making use of the fact that the coordinates P_1 and P_2 defined in Example 5.45 are constants of motion in involution, find a complete solution of the HJ equation for a charged particle in a uniform magnetic field and use it to find a second pair of constants of motion.

Exercise 6.23 (Toda lattice). In Example 5.81 the system with Hamiltonian

$$H = \tfrac{1}{2}({p_1}^2 + {p_2}^2) + e^{q_1 - q_2},$$

was considered and it was shown that the functions

$$P_1 \equiv p_1 + p_2, \qquad P_2 \equiv p_1 p_2 - e^{q_1 - q_2}$$

are constants of motion. Prove that these two functions are in involution, find a complete solution of the HJ equation and use it to solve the equations of motion.

Example 6.24. It should be clear that in order to find a complete solution of the HJ equation starting from n functionally independent constants of motion in involution, $P_1, P_2, \ldots, P_n$, it is not indispensable to solve the equations $P_i = P_i(q_j, p_j, t)$ for $p_1, p_2, \ldots, p_n$; instead of p_i we can employ the corresponding conjugate coordinate q_i (that is, in place of, say, p_5, we can make use of q_5, and so on). For instance, in the case of the Hamiltonian (6.67), the functions

$$P_1 = p_y + mgt, \qquad P_2 = \frac{p_x p_y}{m} + mgx \tag{6.71}$$

are two functionally independent constants of motion in involution. Even though we can solve (6.71) for p_x and p_y, we shall make use of x instead of p_x. From (6.71) we obtain

$$x = \frac{P_2}{mg} - \frac{(P_1 - mgt)p_x}{m^2 g}, \qquad p_y = P_1 - mgt$$

and instead of (6.62), we have

$$-x\,dp_x + p_y dy - H dt = dS,$$

with x and p_y expressed in terms of y, p_x, P_1, P_2, and t. Hence,

$$\begin{aligned} dS &= -\left(\frac{P_2}{mg} - \frac{P_1 p_x}{m^2 g} + \frac{t p_x}{m}\right) dp_x + \big(P_1 - mgt\big) dy - \left[\frac{{p_x}^2}{2m} + \frac{(P_1 - mgt)^2}{2m} + mgy\right] dt \\ &= d\left[-\frac{P_2 p_x}{mg} + \frac{P_1 {p_x}^2}{2m^2 g} - \frac{t {p_x}^2}{2m} + P_1 y - mgty + \frac{(P_1 - mgt)^3}{6m^2 g}\right] \end{aligned}$$

and we can take

$$S(y, p_x, P_1, P_2, t) = -\frac{P_2 p_x}{mg} + \frac{P_1 {p_x}^2}{2m^2 g} - \frac{t {p_x}^2}{2m} + P_1 y - mgty + \frac{(P_1 - mgt)^3}{6m^2 g}.$$

Note that the function S thus obtained is R-separable and is a complete solution of the HJ equation

$$\frac{1}{2m}\left[{p_x}^2 + \left(\frac{\partial S}{\partial y}\right)^2\right] + mgy + \frac{\partial S}{\partial t} = 0,$$

which does not contain the partial derivative of S with respect to p_x.

Exercise 6.25. Show that, in the case of the constants of motion considered in Example 6.20, the coordinates x and y can be expressed in terms of p_x, p_y, P_1, P_2, and t, and use those expressions to find a (type F_4) complete solution of the appropriate HJ equation.

Exercise 6.26. Show that

$$Q = q - \frac{pt}{m} + \frac{kt^3}{3m}$$

is a constant of motion if the Hamiltonian is given by

$$H = \frac{p^2}{2m} - ktq.$$

Use the expression for q in terms of p and Q to find a (type F_3) complete solution, $S(p, Q, t)$, of the HJ equation

$$H\left(-\frac{\partial S}{\partial p}, p, t\right) + \frac{\partial S}{\partial t} = 0$$

and use it to find a second constant of motion.

Exercise 6.27. Show that

$$P = p - \int^t \phi(u)\,\mathrm{d}u$$

is a constant of motion if the Hamiltonian is given by

$$H = \frac{p^2}{2m} - \phi(t)q,$$

where $\phi(t)$ is a given function of t only, and use it to find a complete solution of the HJ equation (cf. Exercise 6.16).

6.3 Mapping of Solutions of the HJ Equation Under Canonical Transformations

The form of the Hamiltonian of a given system can be modified by a canonical transformation and, therefore, the expression of the HJ equation and its solutions can also be modified by these transformations. As we shall see now, there is a simple way of relating a solution of the HJ equation corresponding to a Hamiltonian H with a solution of the HJ equation corresponding to the Hamiltonian K, obtained by means of a canonical transformation. We begin by pointing out a relation between solutions of the HJ equation and certain subsets of the extended phase space [10, 17].

Proposition 6.28. *Any solution, $S(q_i, t)$, of the HJ equation defines a surface (a submanifold), N, of the extended phase space, given by the n equations*

$$p_i = \frac{\partial S}{\partial q_i} \tag{6.72}$$

($i = 1, 2, \ldots, n$), on which the linear differential form $p_i \mathrm{d}q_i - H \mathrm{d}t$ is exact; in fact,

$$p_i \mathrm{d}q_i - H \mathrm{d}t = \mathrm{d}S, \qquad \text{on } N. \tag{6.73}$$

Conversely, an $(n+1)$-dimensional submanifold, N, of the extended phase space, on which the differential form $p_i \mathrm{d}q_i - H \mathrm{d}t$ is exact, defines (up to an additive constant) a solution of the HJ equation. (The solution in question is the function S determined by Equation (6.73).)

The function S appearing in Equations (6.72) and (6.73) may contain some parameters (as in the case of a complete solution), but this is not essential at this point. For example, if the Hamiltonian is taken as

$$H = \frac{p^2}{2m} + \frac{m\omega^2 q^2}{2}, \tag{6.74}$$

then, on the two-dimensional submanifold of the extended phase space defined by

$$p = -m\omega q \tan \omega t,$$

we have

$$\begin{aligned} p\mathrm{d}q - H\mathrm{d}t &= -m\omega q \tan \omega t \, \mathrm{d}q - \left(\frac{m^2\omega^2 q^2 \tan^2 \omega t}{2m} + \frac{m\omega^2 q^2}{2} \right) \mathrm{d}t \\ &= -m\omega \tan \omega t \, \mathrm{d}\left(\frac{q^2}{2}\right) - \frac{m\omega^2}{2} q^2 \sec^2 \omega t \, \mathrm{d}t \\ &= \mathrm{d}\big(-\tfrac{1}{2} m\omega q^2 \tan \omega t \big). \end{aligned}$$

Hence, the function

$$S = -\tfrac{1}{2} m\omega q^2 \tan \omega t$$

is a solution of the HJ equation corresponding to the Hamiltonian (6.74), which does not contain arbitrary parameters.

However, for each value of the parameter P, the equation

$$p = -m\omega q \tan \omega(t + P) \tag{6.75}$$

defines a two-dimensional submanifold of the extended phase space, on which the differential form $p\mathrm{d}q - H\mathrm{d}t$ is exact. In fact, one finds that [on the surface defined by (6.75)]

$$p\mathrm{d}q - H\mathrm{d}t = \mathrm{d}\big[-\tfrac{1}{2} m\omega q^2 \tan \omega(t + P)\big]$$

and this time we have a complete solution of the HJ equation

$$S(q, P, t) = -\tfrac{1}{2} m\omega q^2 \tan \omega(t + P), \tag{6.76}$$

which is not (additively) separable. The completeness of the solution (6.76) is related to the fact that the family of submanifolds defined by (6.75) fills the extended phase space.

Returning to the problem of finding the effect of a canonical transformation on the solutions of the HJ equation, we recall that if the coordinate transformation

$$Q_i = Q_i(q_j, p_j, t), \qquad P_i = P_i(q_j, p_j, t), \tag{6.77}$$

is canonical, then

$$p_i \mathrm{d}q_i - H\mathrm{d}t - (P_i \mathrm{d}Q_i - K\mathrm{d}t) = \mathrm{d}F_1,$$

for some real-valued function F_1 defined in a [$(2n + 1)$-dimensional] region of the extended phase space [see Equation (5.46)]. By contrast with the differential form $p_i \mathrm{d}q_i - H\mathrm{d}t$ (and, similarly, $P_i \mathrm{d}Q_i - K\mathrm{d}t$), which is exact only on some submanifolds of the extended phase space, the combination $p_i \mathrm{d}q_i - H\mathrm{d}t - (P_i \mathrm{d}Q_i - K\mathrm{d}t)$ is exact everywhere (or in some open neighborhood of each point of the extended phase space). Hence, if $p_i \mathrm{d}q_i - H\mathrm{d}t$ is an exact differential on some submanifold of the extended phase space, then $P_i \mathrm{d}Q_i - K\mathrm{d}t$ is also exact on that submanifold.

Thus, if $S(q_i, t)$ is a solution of the HJ equation, then, on the submanifold N defined by (6.72),

$$P_i \mathrm{d}Q_i - K \mathrm{d}t = p_i \mathrm{d}q_i - H \mathrm{d}t - \mathrm{d}F_1$$
$$= \mathrm{d}(S - F_1),$$

which means that

$$S' = S - F_1 \tag{6.78}$$

is a solution of the HJ equation corresponding to K, provided that it is expressed in terms of Q_i and t, making use of Equations (6.72) and (6.77). (Cf. Equations (6.45)–(6.46).) By construction, the solution of the Hamilton equations obtained from S' is the image under the canonical transformation (6.77) of the solution of the Hamilton equations obtained from S.

Example 6.29. A simple and illustrative example is given by the standard Hamiltonian of a one-dimensional harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{m\omega^2 q^2}{2}.$$

The coordinate transformation

$$q = \frac{1}{\omega}\sqrt{\frac{2Q}{m}}\,\cos\omega P, \qquad p = \sqrt{2mQ}\,\sin\omega P, \tag{6.79}$$

is canonical and we can take $K = H$. In fact,

$$p\mathrm{d}q - P\mathrm{d}Q = \sqrt{2mQ}\,\sin\omega P\left(-\sqrt{\frac{2Q}{m}}\,\sin\omega P\,\mathrm{d}P + \frac{\cos\omega P}{\omega\sqrt{2mQ}}\,\mathrm{d}Q\right) - P\mathrm{d}Q$$
$$= -2Q\sin^2\omega P\,\mathrm{d}P + \frac{1}{\omega}\sin\omega P\,\cos\omega P\,\mathrm{d}Q - P\mathrm{d}Q$$
$$= \mathrm{d}\left(-PQ + \frac{Q}{\omega}\sin\omega P\,\cos\omega P\right),$$

hence, up to an additive trivial constant, $F_1 = -PQ + (Q/\omega)\sin\omega P\,\cos\omega P$.

Since $K = H = Q$ [see (6.79)], the HJ equation for K is given by

$$Q + \frac{\partial S'}{\partial t} = 0, \tag{6.80}$$

whose *general* solution is $S' = -Qt + f(Q)$, where $f(Q)$ is an *arbitrary* function of Q only. In order to simplify the computations below, we choose $f(Q) = t_0 Q$, where t_0 is a constant; thus

$$S' = -Q(t - t_0),$$

which constitutes a complete solution of the HJ equation (6.80). Then, from Equation (6.78), taking into account that $P = \partial S'/\partial Q = t_0 - t$, with the aid of the first equation in (6.79) we obtain

$$\begin{aligned} S &= S' + F_1 \\ &= -Q(t - t_0) - PQ + \frac{Q}{\omega} \sin \omega P \, \cos \omega P \\ &= \frac{Q}{\omega} \sin \omega P \, \cos \omega P \\ &= -\tfrac{1}{2} m\omega q^2 \tan \omega (t - t_0), \end{aligned}$$

which is, therefore, the complete solution of the HJ equation corresponding to H [cf. Equation (6.76)].

Example 6.30 (Damped harmonic oscillator). The Hamiltonian

$$H = e^{-2\gamma t} \frac{p^2}{2m} + e^{2\gamma t} \frac{m\omega^2}{2} q^2,$$

where γ is a positive constant, corresponds to a damped harmonic oscillator (see Example 2.6). Making use of (5.15) one finds that the coordinate transformation

$$Q = e^{\gamma t} q, \qquad P = e^{-\gamma t} p$$

is canonical (cf. Example 5.33) and that the new Hamiltonian can be taken as

$$K = \frac{P^2}{2m} + \frac{m\omega^2}{2} Q^2 + \gamma P Q,$$

with $F_1 = 0$. By contrast with H, the Hamiltonian K does not depend explicitly on t and therefore the HJ equation for K admits separable solutions of the form

$$S' = -\tilde{P} t + f(Q),$$

where $\tilde{P}$ is a separation constant and f satisfies

$$\frac{df}{dQ} = -m\gamma Q \pm \sqrt{2m\tilde{P} - m^2(\omega^2 - \gamma^2) Q^2}.$$

Thus,

$$S' = -\tilde{P}t - \tfrac{1}{2}m\gamma Q^2 + \int^{Q} \sqrt{2m\tilde{P} - m^2(\omega^2 - \gamma^2)u^2}\, \mathrm{d}u$$

is a complete solution of the HJ equation for K and, according to Equation (6.78), the function

$$S(q, \tilde{P}, t) = -\tilde{P}t - \tfrac{1}{2}m\gamma\, \mathrm{e}^{2\gamma t}q^2 + \int^{\mathrm{e}^{\gamma t}q} \sqrt{2m\tilde{P} - m^2(\omega^2 - \gamma^2)u^2}\, \mathrm{d}u$$

is the corresponding solution of the HJ equation for H. It may be noticed that this function is neither separable nor R-separable. (Note that in this example we are considering two canonical transformations; the first one relates the original coordinates, q, p, with a second set of canonical coordinates, Q, P. A second canonical transformation is generated by S', leading to a third set of canonical coordinates, $\tilde{Q}, \tilde{P}$, which are constants of motion.)

Hence,

$$p = -m\gamma\, \mathrm{e}^{2\gamma t}q + \mathrm{e}^{\gamma t}\sqrt{2m\tilde{P} - m^2(\omega^2 - \gamma^2)\, \mathrm{e}^{2\gamma t}q^2},$$

and from this equation we can obtain the constant of motion $\tilde{P}$ in terms of (q, p, t),

$$\tilde{P} = \mathrm{e}^{-2\gamma t}\frac{p^2}{2m} + \mathrm{e}^{2\gamma t}\frac{m\omega^2}{2}q^2 + \gamma pq$$

(which coincides with the Hamiltonian K). The second constant of motion,

$$\tilde{Q} = \frac{\partial S}{\partial \tilde{P}} = -t + \int^{\mathrm{e}^{\gamma t}q} \frac{m\, \mathrm{d}u}{\sqrt{2m\tilde{P} - m^2(\omega^2 - \gamma^2)u^2}},$$

gives q as a function of the time (and the constants of motion $\tilde{P}$ and $\tilde{Q}$). For instance, in the case where $\gamma < \omega$ (the so-called underdamped motion), one readily finds

$$q = \mathrm{e}^{-\gamma t}\sqrt{\frac{m(\omega^2 - \gamma^2)}{2\tilde{P}}}\cos\sqrt{\omega^2 - \gamma^2}\big(t + \tilde{Q}\big)$$

(cf. Example 5.13).

Exercise 6.31. The canonical transformation

$$Q = q + \tfrac{1}{2}gt^2, \qquad P = p + mgt$$

relates the Hamiltonians

$$H = \frac{p^2}{2m} + mgq \qquad \text{and} \qquad K = \frac{P^2}{2m}$$

[cf. Equations (6.48)]. Using the fact that

$$S' = \frac{m}{2t}(Q - a)^2,$$

where a is a constant, is a (complete) solution of the HJ equation for K, find the corresponding solution for the HJ equation for H.

Exercise 6.32. The canonical transformation

$$Q = q(\cos\omega t + \omega t \sin\omega t) + \frac{p}{m\omega}(\omega t \cos\omega t - \sin\omega t),$$
$$P = m\omega q \sin\omega t + p \cos\omega t$$

relates the Hamiltonians

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2 \qquad \text{and} \qquad K = \frac{P^2}{2m}$$

(see Exercise 5.6). Making use of the fact that

$$S' = \frac{m}{2t}(Q - a)^2,$$

where a is a constant, is a solution of the HJ equation for K (corresponding to a free particle), find the corresponding solution of the HJ equation for H. (*Hint*: the results of Example 5.7 may be useful.)

Exercise 6.33. Consider the canonical transformation

$$Q = q - vt, \qquad P = p - mv,$$

where v is a constant and m is the mass of a particle. *Assuming* that the new Hamiltonian is given by

$$K = H - vp + \tfrac{1}{2}mv^2$$

(cf. Example 5.63), show that if $S(q, t)$ is a solution of the HJ equation for H, then

$$S'(Q, t) = S(Q + vt, t) - mvQ - \tfrac{1}{2}mv^2 t$$

is the corresponding solution of the HJ equation for K. (This relationship has an analog in quantum mechanics in the transformation of a wavefunction under a Galilean transformation, see, e.g., Torres del Castillo and Nájera Salazar [19].)

Exercise 6.34. Show that in the case of a (passive) translation, $Q = q - s$, $P = p$, assuming that $K = H$, we have $S'(Q, t) = S(Q + s, t)$. Similarly, show that for a translation in the momentum, $Q = q$, $P = p - s$, choosing $K = H$, it follows that $S'(Q, t) = S(Q, t) - sQ$. (Note that the assumption $K = H$ is consistent with the fact that the transformations considered here do not involve the time. Note also that H need not be invariant under these translations.)

Exercise 6.35 (Transformation of the principal function under gauge transformations). As shown in Section 5.2, a gauge transformation

$$\mathbf{A} \mapsto \mathbf{A} + \nabla\xi, \qquad \varphi \mapsto \varphi - \frac{1}{c}\frac{\partial\xi}{\partial t},$$

where ξ is some function of the coordinates and the time, corresponds to a canonical transformation given by

$$Q_i = q_i, \qquad P_i = p_i + \frac{\partial(e\xi/c)}{\partial q_i}.$$

Show that if $K = H - \partial(e\xi/c)/\partial t$, then $S' = S + e\xi/c$. (This result also has a well-known analog in quantum mechanics.)

Covariance of the HJ Equation

The HJ equation is a partial differential equation somewhat similar to other scalar PDEs of mathematical physics, such as the Laplace equation for the electrostatic potential, or the wave equation for the fractional change of the density of the air, in the case of the sound waves. However, apart from the fact that the HJ equation is of first order and not necessarily linear, an important difference between the HJ equation and the other equations just mentioned is that, under a change of coordinates of the configuration space, the solutions of the HJ equation may require an additional term [see Equation (6.78)].

However, according to the discussion presented in Example 5.36, a time-independent coordinate transformation in the configuration space,

$$Q_i = Q_i(q_j),$$

together with the implicit relation

$$p_i = P_j \frac{\partial Q_j}{\partial q_i}, \tag{6.81}$$

constitute a canonical transformation. If we choose $K = H$, i.e.,

$$K(Q_i, P_i, t) = H\big(q_i(Q_j), p_i(Q_j, P_j), t\big) = H\Big(q_i(Q_j), P_j \frac{\partial Q_j}{\partial q_i}, t\Big),$$

then the function F_1 can be taken equal to zero and, according to (6.78),

$$S'(Q_i, t) = S(q_i(Q_j), t) \tag{6.82}$$

is the solution of the HJ equation for the Hamiltonian K corresponding to a solution, $S(q_i, t)$, of the HJ equation for H.

Thus, if we have a (not necessarily complete) solution of the HJ equation, in terms of some coordinates, $S(q_i, t)$, by simply substituting the coordinates q_i by any other set of coordinates of the configuration space, $q_i = q_i(Q_j)$, we obtain a solution of the HJ equation for the same Hamiltonian, provided that the momenta are related by (6.81). This means that the HJ equation is *covariant* under this restricted class of coordinate transformations. (See also Section 6.4.)

6.3.1 The HJ Equation as an Evolution Equation

The HJ equation can be seen as an *evolution equation*, which determines the function $S(q_i, t)$ that reduces to a given function, $f(q_i)$, for $t = 0$ (or any other initial value, t_0, of t). According to the results of the previous section, if we have the solution of the Hamilton equations, we can find the solution of the HJ equation satisfying any initial condition, $S(q_i, t_0) = f(q_i)$, making use of the fact that the time evolution from $t = t_0$ to an arbitrary value of t is a canonical transformation, with the Hamiltonian corresponding to the initial coordinates equal to zero [see Equation (5.97)]. The initial condition $f(q_i)$ can be chosen arbitrarily because *any function*, $f(q_i)$, that does not depend on t, is trivially a solution of the HJ equation if the Hamiltonian is equal to zero.

In the following examples, we obtain the function F_1 appearing in Equation (6.78), corresponding to the time evolution, making use of the explicit solution of the Hamilton equations, while the function S' is the initial condition. If the function $f(q_i)$ contains arbitrary parameters, then the solution $S(q_i, t)$ of the HJ equation will also contain those parameters.

A different approach to the problem of finding the solution of a PDE passing through a given curve or surface can be found, e.g., in Sneddon [14, Sect. 12]; one advantage of the method presented there is that the initial condition need not be the value of S at some particular value of t.

Example 6.36. In the case of the Hamiltonian

$$H = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy, \tag{6.83}$$

where m and g are constants, the solution of the corresponding Hamilton equations can be readily obtained and is given by

$$\begin{aligned} x &= Q_1 + \frac{P_1 t}{m}, \quad & y &= Q_2 + \frac{P_2 t}{m} - \frac{g t^2}{2}, \\ p_x &= P_1, & p_y &= P_2 - mgt, \end{aligned} \tag{6.84}$$

where Q_1, Q_2, P_1, and P_2 are the values of x, y, p_x, and p_y, respectively, at $t = 0$. As the initial condition we choose

$$S(x, y, 0) = \alpha_1 x + \alpha_2 y, \tag{6.85}$$

where α_1, α_2 are two arbitrary constants, that is, as the initial function in terms of the initial coordinates, we take

$$S'(Q_1, Q_2, 0) = \alpha_1 Q_1 + \alpha_2 Q_2. \tag{6.86}$$

Note that, as pointed out above, S' is a solution of the HJ equation for $K = 0$.

Making use of the expressions (6.84) we obtain

$$p_i \mathrm{d}q_i - H \mathrm{d}t - P_i \mathrm{d}Q_i = \mathrm{d}\left(\frac{{p_x}^2 + {p_y}^2}{2m} t - mgty + gt^2 p_y + \frac{m g^2 t^3}{3} \right),$$

while from (6.86) it follows that $\mathrm{d}S' = \alpha_1 \mathrm{d}Q_1 + \alpha_2 \mathrm{d}Q_2$, that is, $P_1 = \alpha_1$, $P_2 = \alpha_2$. Then, from Equations (6.78) and (6.84), expressing all coordinates in terms of x, y, α_1, α_2, and t, we have

$$\begin{aligned} S &= S' + F_1 \\ &= \alpha_1 Q_1 + \alpha_2 Q_2 + \frac{{p_x}^2 + {p_y}^2}{2m} t - mgty + gt^2 p_y + \frac{m g^2 t^3}{3} \\ &= \alpha_1 \left(x - \frac{\alpha_1 t}{m} \right) + \alpha_2 \left(y - \frac{\alpha_2 t}{m} + \frac{g t^2}{2} \right) + \frac{{\alpha_1}^2 + (\alpha_2 - mgt)^2}{2m} t - mgty \\ &\quad + gt^2 (\alpha_2 - mgt) + \frac{m g^2 t^3}{3}, \end{aligned}$$

i.e.,

$$S(x, y, t) = \alpha_1 x + \alpha_2 y - mgty - \frac{{\alpha_1}^2 t}{2m} + \frac{(\alpha_2 - mgt)^3 - {\alpha_2}^3}{6 m^2 g} \tag{6.87}$$

[cf. Equation (6.60)]. The expression (6.87) is a (complete, R-separable) solution of the HJ equation that reduces to the specified function (6.85) for $t = 0$.

Exercise 6.37. Find the solution of the Hamilton equations for the time-dependent Hamiltonian

$$H = \frac{p^2}{2m} - ktq, \tag{6.88}$$

where m and k are constants, and use it to find the solution of the corresponding the HJ equation such that $S(q,0) = aq$, where a is an arbitrary constant. Compare the result with (6.53).

We might consider expressions for $S(q_i, 0)$ more complicated than (6.86) and the one given in Exercise 6.37, but these simple expressions are enough to obtain complete solutions of the HJ equation and to illustrate the procedure.

Example 6.38. We shall consider again the HJ equation

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2\right] + mgy + \frac{\partial S}{\partial t} = 0 \tag{6.89}$$

corresponding to the Hamiltonian (6.83) but, instead of (6.85), we take as the initial condition

$$S(x, y, 0) = k\big[(x - \alpha_1)^2 + (y - \alpha_2)^2\big], \tag{6.90}$$

where k is a constant with the appropriate dimensions, and α_1, α_2 are two arbitrary parameters. Thus,

$$S'(Q_1, Q_2, 0) = k\big[(Q_1 - \alpha_1)^2 + (Q_2 - \alpha_2)^2\big] \tag{6.91}$$

and, therefore,

$$\mathrm{d}S' = 2k\big[(Q_1 - \alpha_1)\,\mathrm{d}Q_1 + (Q_2 - \alpha_2)\,\mathrm{d}Q_2\big]$$

i.e., $P_1 = 2k(Q_1 - \alpha_1)$ and $P_2 = 2k(Q_2 - \alpha_2)$. Proceeding as above, from (6.78), (6.84), and (6.91) we have

$$\begin{aligned} S &= S' + F_1 \\ &= k\big[(Q_1 - \alpha_1)^2 + (Q_2 - \alpha_2)^2\big] + \frac{{p_x}^2 + {p_y}^2}{2m}t - mgty + gt^2 p_y + \frac{mg^2t^3}{3} \\ &= \frac{k(m + 2kt)}{m}\big[(Q_1 - \alpha_1)^2 + (Q_2 - \alpha_2)^2\big] - mgty - \frac{mg^2t^3}{6}. \end{aligned}$$

Hence, after the elimination of the Q_i we find that the solution of the HJ equation that reduces to (6.90) at $t = 0$, is given by

$$S(x, y, t) = \frac{km\big[(x - \alpha_1)^2 + (y - \alpha_2 + gt^2/2)^2\big]}{m + 2kt} - mgty - \frac{mg^2t^3}{6}. \qquad (6.92)$$

Exercise 6.39. Show that the solution of the HJ equation corresponding to the standard Hamiltonian of a one-dimensional harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2$$

that satisfies the initial condition $S(q, 0) = \alpha q$, where α is an arbitrary constant, is given by

$$S(q, t) = -\left(\frac{\alpha^2}{2m} + \frac{m\omega^2}{2}q^2\right)\frac{\tan\omega t}{\omega} + \alpha q \sec\omega t.$$

Of course, if we already have the solution of the Hamilton equations, it does not seem necessary to find a complete solution of the HJ equation. However, the construction presented in this section explicitly shows that given the solution of the Hamilton equations, one can find any complete solution of the HJ equation, and that the general solution of the HJ equation involves an arbitrary function of n variables. On the other hand, in geometrical optics, each solution (complete or not) of the eikonal equation corresponds to a wavefront train and the procedure developed in this section allows us to find the evolution of a given wavefront (see Section 6.5, below).

6.4 Transformation of the HJ Equation Under Arbitrary Point Transformations

With the aid of Proposition 6.28 we can see that under an *arbitrary* point transformation,

$$q'_i = q'_i(q_j, t), \qquad t' = t'(q_j, t), \qquad (6.93)$$

the HJ equation for a Hamiltonian $H(q_i, p_i, t)$ is transformed into the HJ equation for a Hamiltonian H', possibly different from H. (These transformations differ from those considered in Section 6.3, because in the latter the time is not transformed.) Indeed, $S(q_i, t)$ is a solution of the HJ equation for H (containing arbitrary parameters or not) if and only if

$$\mathrm{d}S = p_i \mathrm{d}q_i - H\mathrm{d}t$$

on the submanifold of the extended phase space defined by $p_i = \partial S/\partial q_i$. Inverting the formulas (6.93) we obtain q_i and t as functions of the q'_i and t', then if F_1 is any function of q'_i and t' only (or, equivalently, of q_i and t only), defining

$$S' \equiv S - F_1 \tag{6.94}$$

we have

$$\begin{aligned} \mathrm{d}S' &= p_i \left(\frac{\partial q_i}{\partial q'_j} \mathrm{d}q'_j + \frac{\partial q_i}{\partial t'} \mathrm{d}t' \right) - H \left(\frac{\partial t}{\partial q'_j} \mathrm{d}q'_j + \frac{\partial t}{\partial t'} \mathrm{d}t' \right) - \frac{\partial F_1}{\partial q'_j} \mathrm{d}q'_j - \frac{\partial F_1}{\partial t'} \mathrm{d}t' \\ &= \left(\frac{\partial q_i}{\partial q'_j} p_i - \frac{\partial t}{\partial q'_j} H - \frac{\partial F_1}{\partial q'_j} \right) \mathrm{d}q'_j - \left(\frac{\partial t}{\partial t'} H - \frac{\partial q_i}{\partial t'} p_i + \frac{\partial F_1}{\partial t'} \right) \mathrm{d}t'. \end{aligned}$$

Making use again of Proposition 6.28, the last equation shows that S' is a solution of the HJ equation for the Hamiltonian

$$H' = \frac{\partial t}{\partial t'} H - \frac{\partial q_i}{\partial t'} p_i + \frac{\partial F_1}{\partial t'} \tag{6.95}$$

with

$$p'_j = \frac{\partial q_i}{\partial q'_j} p_i - \frac{\partial t}{\partial q'_j} H - \frac{\partial F_1}{\partial q'_j} \tag{6.96}$$

[cf. Equations (5.65) and (5.66)]. In conclusion, given a Hamiltonian H and an arbitrary function, $F_1(q_i, t)$, any coordinate transformation (6.93), in which the time may be also transformed, leads to a new Hamiltonian (6.95) in such a way that any solution, S, of the HJ equation for H produces a solution, S', of the HJ equation for H' given by (6.94). If S contains arbitrary parameters, so will do S'.

Example 6.40. We shall consider the point transformation

$$q = q' \sec \omega t', \qquad t = \frac{\tan \omega t'}{\omega} \tag{6.97}$$

where ω is a constant. Then, from Equation (6.96), we obtain

$$p' = p \sec \omega t' - \frac{\partial F_1}{\partial q'},$$

which substituted into (6.95) gives

$$\begin{aligned} H' &= H \sec^2 \omega t' - p q' \omega \sec \omega t' \tan \omega t' + \frac{\partial F_1}{\partial t'} \\ &= H \sec^2 \omega t' - \omega q' \tan \omega t' \left(p' + \frac{\partial F_1}{\partial q'} \right) + \frac{\partial F_1}{\partial t'}. \end{aligned}$$

Hence, taking $H = p^2/2m$, and expressing the result in terms of the primed variables

$$\begin{aligned} H' &= \frac{p^2}{2m}\sec^2\omega t' - \omega q' \tan\omega t' \left(p' + \frac{\partial F_1}{\partial q'}\right) + \frac{\partial F_1}{\partial t'} \\ &= \frac{1}{2m}\left(p' + \frac{\partial F_1}{\partial q'}\right)^2 - \omega q' \tan\omega t' \left(p' + \frac{\partial F_1}{\partial q'}\right) + \frac{\partial F_1}{\partial t'} \\ &= \frac{p'^2}{2m} + \left(\frac{\partial F_1}{\partial q'} - m\omega q' \tan\omega t'\right)\frac{p'}{m} + \frac{1}{2m}\left(\frac{\partial F_1}{\partial q'}\right)^2 - \omega q' \tan\omega t' \frac{\partial F_1}{\partial q'} + \frac{\partial F_1}{\partial t'}. \end{aligned}$$

In order to eliminate the term linear in p' we take

$$\frac{\partial F_1}{\partial q'} = m\omega q' \tan\omega t',$$

which implies that $F_1 = \frac{1}{2}m\omega q'^2 \tan\omega t' + f(t')$, where $f(t')$ is some function of t' only. In this manner, H' reduces to

$$H' = \frac{p'^2}{2m} + \frac{m\omega^2}{2}q'^2 + \frac{\mathrm{d}f}{\mathrm{d}t'}.$$

Thus, choosing $f = 0$ it follows that if S is a solution for the HJ equation corresponding to the standard Hamiltonian of a free particle, then

$$S' = S - F_1 = S - \tfrac{1}{2}m\omega q'^2 \tan\omega t' \tag{6.98}$$

is a solution of the HJ equation corresponding to the standard Hamiltonian of a harmonic oscillator.

For instance,

$$S = \frac{m(q-a)^2}{2t},$$

where a is a constant, is a solution of the HJ equation for a free particle, which substituted into (6.98) yields

$$S' = \frac{m\omega[(q'^2 + a^2)\cos\omega t' - 2aq']}{2\sin\omega t'}.$$

Exercise 6.41. Apply the transformation

$$q' = q\mathrm{e}^s + \tfrac{1}{2}gt^2(\mathrm{e}^s - \mathrm{e}^{4s}), \qquad t' = t\mathrm{e}^{2s},$$

where g and s are constants, to the Hamiltonian

$$H = \frac{p^2}{2m} + mgq.$$

Show that by suitably choosing the function F_1 appearing in Equations (6.94)–(6.96) one obtains

$$H' = \frac{p'^2}{2m} + mgq'$$

(that is, the Hamiltonian is form-invariant). This means that, except for the substitution of the coordinates (q, t) by (q', t'), the HJ equation for H' is the same as that for H and, therefore, by means of (6.94), from a given solution of the HJ equation for H we obtain a possibly different solution of the same equation.

6.5 Geometrical Optics

The Hamiltonian formulation of classical mechanics arose from the study of geometrical optics (see, e.g., Whittaker [22, Chap. XI]) and, as we shall see in this section, it is very instructive to apply the formalism developed in this chapter to geometrical optics.

Fermat's Principle. The Ray Equation

In geometrical optics it is assumed that the light propagates along curves, which are called light rays. The basic equations of geometrical optics can be obtained from the *Fermat principle of least time*, which can be formulated in the following way. The speed of light in an isotropic medium, with refractive index n, is c/n, where c is the speed of light in vacuum; therefore, given two points of the three-dimensional Euclidean space, A and B, the time required for the light to go from A to B along a curve C is given by the integral

$$\frac{1}{c}\int_C n\,\mathrm{d}s, \tag{6.99}$$

where $\mathrm{d}s$ is the arclength element (see below). Of course, there are an infinite number of curves joining A and B; the Fermat principle states that the path actually followed by the light is the one that minimizes the integral (6.99). Since c is a constant, finding the curve corresponding to the least time is equivalent to finding the curve with the minimum *optical length* (or *optical path length*), defined as

$$\int_C n\,\mathrm{d}s. \tag{6.100}$$

If we consider curves that can be parameterized by one of the Cartesian coordinates, z say, the integral (6.100) can be expressed as

$$\int_{z_0}^{z_1} n(x, y, z) \left[1 + \left(\frac{\mathrm{d}x}{\mathrm{d}z} \right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}z} \right)^2 \right]^{1/2} \mathrm{d}z, \tag{6.101}$$

where z_0 and z_1 are the values of the coordinate z at the points A and B, respectively (see Figure 6.1). Hence, the light rays are determined by the Euler–Lagrange equations for the Lagrangian

$$L(x, y, x', y', z) = n(x, y, z)\sqrt{1 + x'^2 + y'^2}, \tag{6.102}$$

where $x' \equiv \mathrm{d}x/\mathrm{d}z$ and $y' \equiv \mathrm{d}y/\mathrm{d}z$. However, instead of writing down these equations and attempting to solve them, we shall be mainly interested in the Hamiltonian description.

Exercise 6.42. Making use of the fact that $\mathrm{d}s = \sqrt{1 + x'^2 + y'^2}\,\mathrm{d}z$, show that the Euler–Lagrange equations for the Lagrangian (6.102) amount to

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n \frac{\mathrm{d}x}{\mathrm{d}s} \right) = \frac{\partial n}{\partial x}, \qquad \frac{\mathrm{d}}{\mathrm{d}s}\left(n \frac{\mathrm{d}y}{\mathrm{d}s} \right) = \frac{\partial n}{\partial y},$$

and that, making use of the identity (1.92), one obtains the equation

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n \frac{\mathrm{d}z}{\mathrm{d}s} \right) = \frac{\partial n}{\partial z}.$$

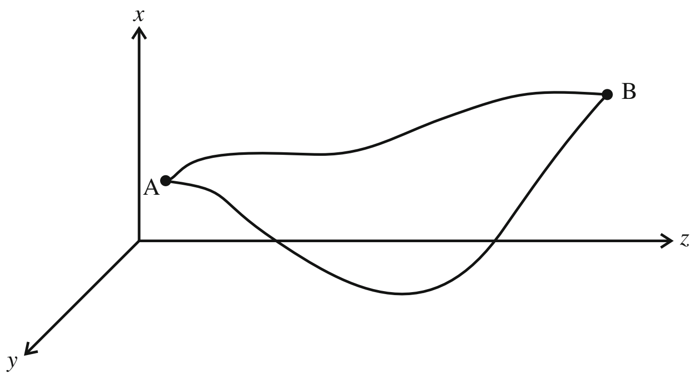


Fig. 6.1 The curves shown join the points A and B. In order to use z as a parameter for these curves, any plane $z =$ const must intersect each curve at most at one point

The last three equations are equivalent to the vector equation

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right) = \nabla n,$$

where $\mathbf{r} = (x, y, z)$ is the position vector of a point of the ray. This equation is known as the *ray equation*.

Exercise 6.43 (Spherically symmetric media). Show that if the refractive index is a function of the distance, r, to a fixed point (taken as the origin), $n = n(r)$, then

$$\mathbf{r} \times n\,\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}$$

is constant along each ray and show that this implies that each ray lies on a plane passing through the origin.

The optical system defined by the spherically symmetric refractive index

$$n = \frac{a}{b + r^2}, \tag{6.103}$$

where a and b are real constants, with $a > 0$, is known as *Maxwell's fish eye*. Several properties of this system can be derived from its relationship with the Kepler problem. Show that, in this case,

$$\mathbf{r} \times \left(\mathbf{r} \times n\,\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right) + \frac{a}{2}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}$$

is also constant along each ray and deduce from this that the rays are (arcs of) circles (just like the hodographs of the Kepler problem). (See also Exercise 6.45, below.)

The Eikonal Equation

The canonical momenta conjugate to x and y are

$$p_x = \frac{\partial L}{\partial x'} = \frac{nx'}{\sqrt{1 + x'^2 + y'^2}}, \qquad p_y = \frac{\partial L}{\partial y'} = \frac{ny'}{\sqrt{1 + x'^2 + y'^2}}, \tag{6.104}$$

respectively, and from these equations we obtain

$$\sqrt{n^2 - {p_x}^2 - {p_y}^2} = \frac{n}{\sqrt{1 + x'^2 + y'^2}}$$

and

$$x' = \frac{p_x}{\sqrt{n^2 - {p_x}^2 - {p_y}^2}}, \qquad y' = \frac{p_y}{\sqrt{n^2 - {p_x}^2 - {p_y}^2}}. \tag{6.105}$$

Thus, in this approach, we have a system with two degrees of freedom, with the coordinate z as the independent variable, and a Hamiltonian given by [see Equation (4.10)]

$$H = -\sqrt{n^2 - {p_x}^2 - {p_y}^2} \tag{6.106}$$

(hence, e.g., $\mathrm{d}p_x/\mathrm{d}z = -\partial H/\partial x$). Therefore, the corresponding HJ equation is

$$-\left[n^2 - \left(\frac{\partial S}{\partial x}\right)^2 - \left(\frac{\partial S}{\partial y}\right)^2\right]^{1/2} + \frac{\partial S}{\partial z} = 0, \tag{6.107}$$

which implies that

$$\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2 = n^2. \tag{6.108}$$

This last equation is known as the *eikonal equation* and, in this context, S is called the *eikonal* (or *eikonal function*). The eikonal equation can also be derived from the Huygens principle (see, e.g., Synge [15, Sect. 22]). Any complete solution of Equation (6.107) or (6.108) allows us to find all the light rays in the medium characterized by the refractive index n (that is, the solutions of the ray equation). Note that in the eikonal equation, the three coordinates (x, y, z) appear on an equal footing and it is optional which of them is taken as the independent variable.

For instance, if the refractive index is *constant*, Equation (6.107) admits separable solutions of the form

$$S(x, y, P_1, P_2, z) = P_1 x + P_2 y + \sqrt{n^2 - {P_1}^2 - {P_2}^2}\, z, \tag{6.109}$$

where P_1, P_2 are constants such that ${P_1}^2 + {P_2}^2 \leqslant n^2$. Making use of the standard formulas (6.8) we obtain the canonical transformation generated by (6.109) (treating again z as the independent variable)

$$p_x = P_1, \quad p_y = P_2, \quad Q_1 = x - \frac{P_1 z}{\sqrt{n^2 - {P_1}^2 - {P_2}^2}}, \quad Q_2 = y - \frac{P_2 z}{\sqrt{n^2 - {P_1}^2 - {P_2}^2}}. \tag{6.110}$$

The last two equations in (6.110) show that in this case the light rays are straight lines, as expected. In fact, in terms of the usual vector notation, we have

$$(x, y, z) = (Q_1, Q_2, 0) + \frac{z\,(P_1, P_2, \sqrt{n^2 - {P_1}^2 - {P_2}^2})}{\sqrt{n^2 - {P_1}^2 - {P_2}^2}}.$$

Thus, the constants P_1, P_2 determine the direction of the light ray, and $(Q_1, Q_2, 0)$ are the Cartesian coordinates of the intersection of the ray with the plane $z = 0$

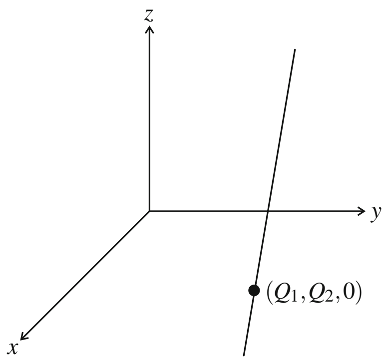


Fig. 6.2 Any light ray in a homogeneous isotropic medium is a straight line. With the exception of the light rays parallel to the xy-plane, any light ray can be specified by the four real numbers (Q_1, Q_2, P_1, P_2); P_1 and P_2 determine the direction of the ray, and $(Q_1, Q_2, 0)$ are the Cartesian coordinates of the intersection of the ray with the plane $z = 0$. The planes orthogonal to this straight line are the wavefronts defined by (6.109)

(see Figure 6.2). (Note also that the two last equations in (6.110) are obtained regardless of which of the coordinates (x, y, z) is taken as the independent variable.)

Exercise 6.44. In terms of the spherical coordinates (r, θ, ϕ), the optical length (6.100) is given by

$$\int_{r_1}^{r_2} n(r, \theta, \phi) \sqrt{1 + r^2\theta'^2 + r^2 \sin^2\theta\phi'^2}\, \mathrm{d}r,$$

with $\theta' \equiv \mathrm{d}\theta/\mathrm{d}r$, $\phi' \equiv \mathrm{d}\phi/\mathrm{d}r$, assuming that the curve C can be parameterized by r [cf. Equation (6.101)]. Starting from the Fermat principle, using this expression, show that the corresponding HJ equation leads to the equation

$$\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \theta}\right)^2 + \frac{1}{r^2 \sin^2\theta}\left(\frac{\partial S}{\partial \phi}\right)^2 = n^2 \tag{6.111}$$

which is just the eikonal equation (6.108) expressed in spherical coordinates. Even though we might expect this result, taking into account the meaning of the eikonal function, it does not follow from the discussion presented in the preceding sections (e.g., Section 6.3) because in the present case we are also changing the parameter of the light rays z in Equation (6.101), by r.

Exercise 6.45. Solving the eikonal equation, show that in the case of the Maxwell fish eye, the light rays are (arcs of) circles. (It is convenient to make use of the fact that each ray lies on a plane passing through the origin—see Exercise 6.43.)

Going back to the general case, where n is an arbitrary function, from Equations (6.105) and (6.107) we find that the vector with Cartesian components $(\mathrm{d}x/\mathrm{d}z, \mathrm{d}y/\mathrm{d}z, \mathrm{d}z/\mathrm{d}z)$, which is tangent to the light ray, is proportional to the gradient of S,

$$\left(\frac{\mathrm{d}x}{\mathrm{d}z}, \frac{\mathrm{d}y}{\mathrm{d}z}, \frac{\mathrm{d}z}{\mathrm{d}z}\right) = \frac{(p_x, p_y, \sqrt{n^2 - {p_x}^2 - {p_y}^2})}{\sqrt{n^2 - {p_x}^2 - {p_y}^2}} = \frac{1}{\partial S/\partial z}\left(\frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}, \frac{\partial S}{\partial z}\right),$$

which means that the light rays intersect orthogonally the level surfaces $S(x, y, z) =$ const. (see Figure 6.3). The surfaces $S =$ const. constitute the *wavefronts*.

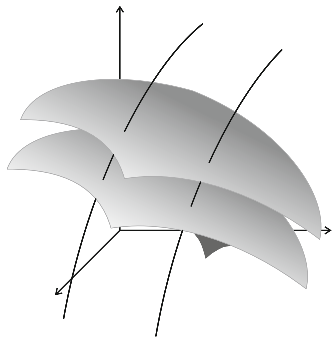


Fig. 6.3 The level surfaces of a solution $S(x, y, z)$ of the eikonal equation constitute a family of two-dimensional surfaces that fill the three-dimensional space. The curves orthogonal to these surfaces correspond to some of the possible rays of light in the medium

In the example (6.109), the wavefronts are the planes normal to the vector with Cartesian components $(P_1, P_2, \sqrt{n^2 - {P_1}^2 - {P_2}^2})$ (see Figure 6.2).

Exercise 6.46. Show that if $S(x, y, z)$ is a solution of the eikonal equation, which may not contain arbitrary parameters, then the curves orthogonal to the level surfaces $S =$ const. correspond to possible light rays. (*Hint*: if $x_i(s)$ are the Cartesian coordinates of a curve parameterized by its arclength, then the norm of its tangent vector, $(\mathrm{d}x/\mathrm{d}s, \mathrm{d}y/\mathrm{d}s, \mathrm{d}z/\mathrm{d}s)$, is equal to 1. On the other hand, if this curve is orthogonal to the level surfaces of S, then its tangent vector is proportional to ∇S, whose norm is equal to n.)

Exercise 6.47. Find the light rays determined by the eikonal function in two dimensions

$$S(x, y, P) = \tfrac{1}{2}a\big[(x^2 - y^2)\cos P + 2xy \sin P\big],$$

where a and P are constants. What is the refractive index?

Each solution, $S(x, y, z)$, of the eikonal equation defines a family of surfaces (its level surfaces) in such a way that if a_1 and a_2 are two real constants (such that the sets $\{(x, y, z) \in \mathbb{R}^3 \mid S(x, y, z) = a_1\}$ and $\{(x, y, z) \in \mathbb{R}^3 \mid S(x, y, z) = a_2\}$ are nonempty), the wavefront $S = a_2$ is obtained from $S = a_1$ by the propagation of the light by a time $(a_2 - a_1)/c$ (see Figure 6.4). In fact, if we consider a light ray, C, connecting a point belonging to the surface $S(x, y, z) = a_1$ with a point belonging to $S(x, y, z) = a_2$, assuming that this curve can be parameterized by z, we have [see Equations (6.107) and (6.104)]

$$
\begin{aligned}
a_2 - a_1 &= \int_C \mathrm{d}S = \int_{z_1}^{z_2} \left(\frac{\partial S}{\partial x} x' + \frac{\partial S}{\partial y} y' + \frac{\partial S}{\partial z} \right) \mathrm{d}z \\
&= \int_{z_1}^{z_2} \left(p_x x' + p_y y' + \sqrt{n^2 - {p_x}^2 - {p_y}^2} \right) \mathrm{d}z = \int_{z_1}^{z_2} \frac{n\,(x'^2 + y'^2 + 1)}{\sqrt{1 + x'^2 + y'^2}} \, \mathrm{d}z \\
&= \int_C n \, \mathrm{d}s.
\end{aligned}
$$

Thus, the set of level surfaces $S = \text{const.}$ represent the evolution of one of them (see Figure 6.4).

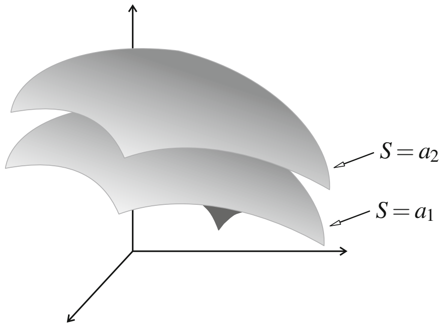


Fig. 6.4 Each level surface of $S(x, y, z)$ represents the wavefront at some particular time. The level surface $S = a_1$ evolves into the level surface $S = a_2$ after a time $(a_2 - a_1)/c$. The set of all the level surfaces of S represents the time evolution of anyone of them

Generation of Complete Solutions of the Eikonal Equation from a given Complete Solution

As shown in Section 6.1.1, from a given complete solution of the HJ equation one can obtain any other complete solution of the same equation. For instance, making use of the complete solution (6.109), and choosing the time-independent generating function $F(P_i, \tilde{P}_i) = P_1 \tilde{P}_1 + P_2 \tilde{P}_2$, according to Equations (6.46) we have

$$
0 = \frac{\partial (S - F)}{\partial P_1} = x - \frac{P_1 z}{\sqrt{n^2 - {P_1}^2 - {P_2}^2}} - \tilde{P}_1,
$$

hence,

$$
P_1 = \frac{(x - \tilde{P}_1)\sqrt{n^2 - {P_1}^2 - {P_2}^2}}{z},
$$

with a similar expression for P_2. Then, from the expressions thus obtained, we find that

$$P_1 = \frac{n\,(x - \tilde{P}_1)}{\sqrt{(x - \tilde{P}_1)^2 + (y - \tilde{P}_2)^2 + z^2}}, \qquad P_2 = \frac{n\,(y - \tilde{P}_2)}{\sqrt{(x - \tilde{P}_1)^2 + (y - \tilde{P}_2)^2 + z^2}}.$$

With the aid of these formulas we can eliminate the parameters P_i appearing in the right-hand side of (6.45) and in this manner we find a second complete solution of the eikonal equation

$$\tilde{S}(x, y, \tilde{P}_1, \tilde{P}_2, z) = n\sqrt{(x - \tilde{P}_1)^2 + (y - \tilde{P}_2)^2 + z^2}, \tag{6.112}$$

$\tilde{P}_1, \tilde{P}_2 \in (-\infty, \infty)$. The wavefronts $\tilde{S} = \text{const.}$ are spheres [centered at $(\tilde{P}_1, \tilde{P}_2, 0)$].

Solutions

Exercises of Chapter 1

1.2 From Equations (1.1) we have

$$\ddot{\mathbf{r}}_1 = \ddot{x}\,\mathbf{i}, \qquad \ddot{\mathbf{r}}_2 = (\ddot{x} + \ddot{y}\,\cot\theta_0)\,\mathbf{i} + \ddot{y}\,\mathbf{j}$$

and

$$\frac{\partial \mathbf{r}_1}{\partial x} = \mathbf{i}, \qquad \frac{\partial \mathbf{r}_1}{\partial y} = \mathbf{0}, \qquad \frac{\partial \mathbf{r}_2}{\partial x} = \mathbf{i}, \qquad \frac{\partial \mathbf{r}_2}{\partial y} = \cot\theta_0 \mathbf{i} + \mathbf{j}.$$

On the other hand, $\mathbf{F}_1^{(\text{appl})} = -m_1 g\mathbf{j}$ and $\mathbf{F}_2^{(\text{appl})} = -m_2 g\mathbf{j}$. Hence, substituting into Equations (1.8) we obtain

$$\begin{aligned} (m_1\ddot{x}\,\mathbf{i} + m_1 g\mathbf{j}) \cdot \mathbf{i} + \{m_2[(\ddot{x} + \ddot{y}\,\cot\theta_0)\,\mathbf{i} + \ddot{y}\,\mathbf{j}] + m_2 g\mathbf{j}\} \cdot \mathbf{i} &= 0, \\ (m_1\ddot{x}\,\mathbf{i} + m_1 g\mathbf{j}) \cdot \mathbf{0} + \{m_2[(\ddot{x} + \ddot{y}\,\cot\theta_0)\,\mathbf{i} + \ddot{y}\,\mathbf{j}] + m_2 g\mathbf{j}\} \cdot (\cot\theta_0\mathbf{i} + \mathbf{j}) &= 0, \end{aligned}$$

which, using the fact that **i** and **j** are orthonormal vectors, reduce to

$$\begin{aligned} m_1\ddot{x} + m_2(\ddot{x} + \ddot{y}\,\cot\theta_0) &= 0, \\ m_2(\ddot{x} + \ddot{y}\,\cot\theta_0)\cot\theta_0 + m_2\ddot{y} + m_2 g &= 0. \end{aligned}$$

The solution of this system of algebraic equations is readily found to be

$$\ddot{x} = \frac{m_2 g \cot\theta_0}{m_2 + m_1 \csc^2\theta_0}, \qquad \ddot{y} = \frac{-(m_1 + m_2)g}{m_2 + m_1 \csc^2\theta_0}.$$

G. F. Torres del Castillo, *An Introduction to Hamiltonian Mechanics*, Birkhäuser Advanced Texts Basler Lehrbücher, https://doi.org/10.1007/978-3-319-95225-3

1.6 With the aid of Figure 1.8, we find that

$$\mathbf{r}_1 = x\mathbf{i}, \qquad \mathbf{r}_2 = x\mathbf{i} + l(\sin\theta\mathbf{i} - \cos\theta\mathbf{j}),$$

where l is the length of the pendulum. Eliminating the parameters x and θ, or from Figure 1.8, one finds that these expressions represent the general solutions of the constraint equations

$$\mathbf{r}_1 \cdot \mathbf{j} = 0, \qquad |\mathbf{r}_2 - \mathbf{r}_1| = l.$$

The constraint forces are given by $\mathbf{F}_1^{(\text{constr})} = N\mathbf{j} + T(\sin\theta\mathbf{i} - \cos\theta\mathbf{j})$ and $\mathbf{F}_2^{(\text{constr})} = -T(\sin\theta\mathbf{i} - \cos\theta\mathbf{j})$, where N is the magnitude of the normal force on the block and T is the tension of the rod. Hence,

$$\sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial x} = \mathbf{F}_1^{(\text{constr})} \cdot \mathbf{i} + \mathbf{F}_2^{(\text{constr})} \cdot \mathbf{i} = N\mathbf{j} \cdot \mathbf{i} = 0$$

and

$$\sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial \theta} = \mathbf{F}_2^{(\text{constr})} \cdot (l\cos\theta\mathbf{i} + l\sin\theta\mathbf{j}) = 0$$

which implies that the equations of motion can be obtained from Equations (1.8).

Taking into account Hooke's law, the applied forces are $\mathbf{F}_1^{(\text{appl})} = -kx\,\mathbf{i} - m_1 g\mathbf{j}$ and $\mathbf{F}_2^{(\text{appl})} = -m_2 g\mathbf{j}$, and Equations (1.8) give

$$(m_1\ddot{x}\,\mathbf{i} + kx\,\mathbf{i} + m_1 g\mathbf{j}) \cdot \mathbf{i} \\ + \{m_2[(\ddot{x} + l\ddot{\theta}\cos\theta - l\dot{\theta}^2\sin\theta)\mathbf{i} + (l\ddot{\theta}\sin\theta + l\dot{\theta}^2\cos\theta)\mathbf{j}] + m_2 g\mathbf{j}\} \cdot \mathbf{i} = 0,$$

and

$$\{m_2[(\ddot{x} + l\ddot{\theta}\cos\theta - l\dot{\theta}^2\sin\theta)\mathbf{i} \\ + (l\ddot{\theta}\sin\theta + l\dot{\theta}^2\cos\theta)\mathbf{j}] + m_2 g\mathbf{j}\} \cdot (l\cos\theta\mathbf{i} + l\sin\theta\mathbf{j}) = 0,$$

which reduce to the second-order ODEs

$$m_1\ddot{x} + kx + m_2(\ddot{x} + l\ddot{\theta}\cos\theta - l\dot{\theta}^2\sin\theta) = 0, \qquad \ddot{x}\cos\theta + l\ddot{\theta} + g\sin\theta = 0.$$

1.8 According to the definition of the angle ϕ, the position vector of the bead can be expressed in the form

$$\mathbf{r} = a(\cos\omega t\mathbf{i} + \sin\omega t\mathbf{j}) + a[\cos(\omega t + \phi)\mathbf{i} + \sin(\omega t + \phi)\mathbf{j}].$$

The only constraint force on the bead is a normal force produced by the hoop, hence

$$\mathbf{F}^{(\text{constr})} \cdot \frac{\partial \mathbf{r}}{\partial \phi} = \mathbf{F}^{(\text{constr})} \cdot a[-\sin(\omega t + \phi)\mathbf{i} + \cos(\omega t + \phi)\mathbf{j}],$$

which is equal to zero because the vector $\partial \mathbf{r}/\partial \phi$ is tangent to the hoop.

The applied force is $\mathbf{F}^{(\text{appl})} = -mg\mathbf{k}$ and from Newton's second law we obtain

$$\begin{aligned} 0 &= \left(m\ddot{\mathbf{r}} - \mathbf{F}^{(\text{appl})}\right) \cdot \frac{\partial \mathbf{r}}{\partial \phi} \\ &= m\Big\{ -\omega^2 a(\cos \omega t \mathbf{i} + \sin \omega t \mathbf{j}) + \ddot{\phi} a[-\sin(\omega t + \phi)\mathbf{i} + \cos(\omega t + \phi)\mathbf{j}] \\ &\quad -(\omega+\dot{\phi})^2 a[\cos(\omega t+\phi)\mathbf{i}+ \sin(\omega t+\phi)\mathbf{j}]\Big\} \cdot a[-\sin(\omega t+\phi)\mathbf{i}+ \cos(\omega t+\phi)\mathbf{j}] \\ &= ma^2(\ddot{\phi} + \omega^2 \sin \phi), \end{aligned}$$

that is,

$$\ddot{\phi} + \omega^2 \sin \phi = 0$$

[cf. Equation (1.25)].

1.9 The position vectors shown in Figure 1.11 are given by

$$\mathbf{r}_1 = x\mathbf{i}, \qquad \mathbf{r}_2 = h\mathbf{j} + x\mathbf{i} + (l - x)(\cos \theta_0 \mathbf{i} - \sin \theta_0 \mathbf{j}),$$

and also with the aid of Figure 1.11 one finds that the constraint forces are

$$\begin{aligned} \mathbf{F}_1^{(\text{constr})} &= N_1 \mathbf{j} - N_2(\sin \theta_0 \mathbf{i} + \cos \theta_0 \mathbf{j}) - T\mathbf{i} + T(\cos \theta_0 \mathbf{i} - \sin \theta_0 \mathbf{j}), \\ \mathbf{F}_2^{(\text{constr})} &= N_2(\sin \theta_0 \mathbf{i} + \cos \theta_0 \mathbf{j}) - T(\cos \theta_0 \mathbf{i} - \sin \theta_0 \mathbf{j}), \end{aligned}$$

where N_1 is the magnitude of the normal force exerted by the horizontal surface on the wedge, N_2 is the normal force applied by the block above on the wedge, and T is the tension of the rope. Hence, combining these expressions one finds

$$\begin{aligned} &\sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\text{constr})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial x} \\ &\quad = \mathbf{F}_1^{(\text{constr})} \cdot \mathbf{i} + \mathbf{F}_2^{(\text{constr})} \cdot [\mathbf{i} - (\cos \theta_0 \mathbf{i} - \sin \theta_0 \mathbf{j})] \\ &\quad = (\mathbf{F}_1^{(\text{constr})} + \mathbf{F}_2^{(\text{constr})}) \cdot \mathbf{i} - \mathbf{F}_2^{(\text{constr})} \cdot (\cos \theta_0 \mathbf{i} - \sin \theta_0 \mathbf{j}) \\ &\quad = (N_1 \mathbf{j} - T\mathbf{i}) \cdot \mathbf{i} - [N_2(\sin \theta_0 \mathbf{i} + \cos \theta_0 \mathbf{j}) - T(\cos \theta_0 \mathbf{i} - \sin \theta_0 \mathbf{j})] \cdot (\cos \theta_0 \mathbf{i} - \sin \theta_0 \mathbf{j}) \\ &\quad = -T + T \\ &\quad = 0. \end{aligned}$$

The applied forces are the weights of the blocks: $\mathbf{F}_1^{(\text{appl})} = -m_1 g\mathbf{j}$, $\mathbf{F}_2^{(\text{appl})} = -m_2 g\mathbf{j}$, hence, the equation of motion for this system can be obtained from

$$\begin{aligned}0 &= \sum_{\alpha=1}^{2} \left(m_\alpha \ddot{\mathbf{r}}_\alpha - \mathbf{F}_\alpha^{(\text{appl})}\right) \cdot \frac{\partial \mathbf{r}_\alpha}{\partial x} \\ &= (m_1 \ddot{x}\mathbf{i} + m_1 g\mathbf{j}) \cdot \mathbf{i} \\ &\quad + \{m_2[\ddot{x}\mathbf{i} - \ddot{x}(\cos\theta_0 \mathbf{i} - \sin\theta_0 \mathbf{j})] + m_2 g\mathbf{j}\} \cdot [\mathbf{i} - (\cos\theta_0 \mathbf{i} - \sin\theta_0 \mathbf{j})] \\ &= m_1 \ddot{x} + 2m_2 \ddot{x}(1 - \cos\theta_0) + m_2 g \sin\theta_0,\end{aligned}$$

i.e.,

$$\ddot{x} = -\frac{m_2 g \sin\theta_0}{m_1 + 2m_2(1 - \cos\theta_0)}.$$

Multiplying this equation by $\dot{x}$ we find that

$$\tfrac{1}{2}[m_1 + 2m_2(1 - \cos\theta_0)]\dot{x}^2 + m_2 g \sin\theta_0 x$$

is a constant of motion, which is the total energy. Thus, the phase curves in the $x\dot{x}$-plane are parabolas (see Figure S.1).

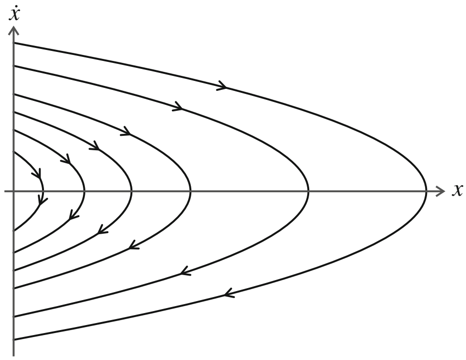


Fig. S.1 The phase curves are parabolas. The are no equilibrium points in the domain $x > 0$

1.12 Making use of the expressions for the position vectors

$$\mathbf{r}_1 = x\mathbf{i}, \qquad \mathbf{r}_2 = x\mathbf{i} + l(\sin\theta \mathbf{i} - \cos\theta \mathbf{j}),$$

employed in the solution of Exercise 1.6, we have

$$\dot{\mathbf{r}}_1 = \dot{x}\mathbf{i}, \qquad \dot{\mathbf{r}}_2 = \dot{x}\mathbf{i} + l\dot{\theta}(\cos\theta \mathbf{i} + \sin\theta \mathbf{j}).$$

Thus, the kinetic energy is given by

$$T = \tfrac{1}{2}m_1\dot{x}^2 + \tfrac{1}{2}m_2(\dot{x}^2 + l^2\dot{\theta}^2 + 2l\dot{x}\dot{\theta}\cos\theta).$$

The applied forces were also given in the solution of Exercise 1.6, they are $\mathbf{F}_1^{(\mathrm{appl})} = -kx\,\mathbf{i} - m_1 g\mathbf{j}$ and $\mathbf{F}_2^{(\mathrm{appl})} = -m_2 g\mathbf{j}$. Therefore, the generalized forces are

$$Q_1 = \sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\mathrm{appl})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial x} = (-kx\,\mathbf{i} - m_1 g\mathbf{j}) \cdot \mathbf{i} + (-m_2 g\mathbf{j}) \cdot \mathbf{i} = -kx$$

and

$$\begin{aligned} Q_2 &= \sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\mathrm{appl})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial \theta} = (-kx\,\mathbf{i} - m_1 g\mathbf{j}) \cdot \mathbf{0} + (-m_2 g\mathbf{j}) \cdot l(\cos\theta\mathbf{i} + \sin\theta\mathbf{j}) \\ &= -m_2 gl\sin\theta. \end{aligned}$$

These generalized forces are derivable from the potential $V = \frac{1}{2}kx^2 - m_2 gl\cos\theta$ [cf. Equation (1.48)]. (The term $\frac{1}{2}kx^2$ can be recognized as the potential energy of a spring with stiffness k, while the second term, $-m_2 gl\cos\theta$, has the well-known form mgh of the gravitational potential energy of a body of mass m in a uniform gravitational field with acceleration g, and height h measured from an arbitrarily selected level.)

Hence, the standard Lagrangian is

$$L = \tfrac{1}{2}\big[(m_1 + m_2)\dot{x}^2 + m_2(l^2\dot{\theta}^2 + 2l\dot{x}\dot{\theta}\cos\theta)\big] - \tfrac{1}{2}kx^2 + m_2 gl\cos\theta,$$

which, substituted into the Lagrange equations, yields

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = \frac{\mathrm{d}}{\mathrm{d}t}\big[(m_1 + m_2)\dot{x} + m_2 l\dot{\theta}\cos\theta\big] + kx \\ &= (m_1 + m_2)\ddot{x} + m_2 l\ddot{\theta}\cos\theta - m_2 l\dot{\theta}^2\sin\theta + kx, \\ 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = \frac{\mathrm{d}}{\mathrm{d}t}\big[m_2(l^2\dot{\theta} + l\dot{x}\cos\theta)\big] + m_2 l\dot{x}\dot{\theta}\sin\theta + m_2 gl\sin\theta \\ &= m_2(l^2\ddot{\theta} + l\ddot{x}\cos\theta) + m_2 gl\sin\theta. \end{aligned}$$

These equations coincide with the equations of motion obtained in Exercise 1.6.

1.13 Making use of the parametrization of the position vector of the bead given in the solution of Exercise 1.8 we find that the velocity of the bead is given by

$$\dot{\mathbf{r}} = a\omega(-\sin\omega t\mathbf{i} + \cos\omega t\mathbf{j}) + a(\omega + \dot{\phi})[-\sin(\omega t + \phi)\mathbf{i} + \cos(\omega t + \phi)\mathbf{j}]$$

and, therefore, the kinetic energy is

$$T = \frac{m}{2}[a^2\omega^2 + 2a^2\omega(\omega + \dot{\phi})\cos\phi + a^2(\omega + \dot{\phi})^2].$$

On the other hand, the only applied force is $-mg\mathbf{k}$ (assuming that the z-axis points upwards) and the generalized force is

$$Q = (-mg\mathbf{k}) \cdot \frac{\partial \mathbf{r}}{\partial \phi} = (-mg\mathbf{k}) \cdot a[-\sin(\omega t + \phi)\mathbf{i} + \cos(\omega t + \phi)\mathbf{j}] = 0.$$

Thus, we can chose $V = 0$ and the standard Lagrangian is given by

$$L = \frac{m}{2}[a^2\omega^2 + 2a^2\omega(\omega + \dot{\phi})\cos\phi + a^2(\omega + \dot{\phi})^2].$$

Substituting this Lagrangian in the Lagrange equations we obtain

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}[ma^2\omega\cos\phi + ma^2(\omega + \dot{\phi})] + ma^2\omega(\omega + \dot{\phi})\sin\phi \\ &= -ma^2\omega\dot{\phi}\sin\phi + ma^2\ddot{\phi} + ma^2\omega(\omega + \dot{\phi})\sin\phi \\ &= ma^2(\ddot{\phi} + \omega^2\sin\phi), \end{aligned}$$

which is the equation of motion found in the solution of Exercise 1.8.

1.14 Making use of the parametrization of the position vectors employed in the solution of Exercise 1.9,

$$\mathbf{r}_1 = x\mathbf{i}, \qquad \mathbf{r}_2 = h\mathbf{j} + x\mathbf{i} + (l - x)(\cos\theta_0\mathbf{i} - \sin\theta_0\mathbf{j}),$$

we find that the kinetic energy of the system is

$$\begin{aligned} T &= \tfrac{1}{2}m_1\dot{\mathbf{r}}_1{}^2 + \tfrac{1}{2}m_2\dot{\mathbf{r}}_2{}^2 = \tfrac{1}{2}m_1\dot{x}^2 + \tfrac{1}{2}m_2[\dot{x}\mathbf{i} - \dot{x}(\cos\theta_0\mathbf{i} - \sin\theta_0\mathbf{j})]^2 \\ &= \tfrac{1}{2}[m_1 + 2m_2(1 - \cos\theta_0)]\dot{x}^2. \end{aligned}$$

The generalized force is

$$Q = \sum_{\alpha=1}^{2} \mathbf{F}_\alpha^{(\mathrm{appl})} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial x} = -m_1 g\mathbf{j}\cdot\mathbf{i} - m_2 g\mathbf{j}\cdot[\mathbf{i} - (\cos\theta_0\mathbf{i} - \sin\theta_0\mathbf{j})] = -m_2 g\sin\theta_0,$$

which is derivable from the potential $V = m_2 g x \sin\theta_0$. Hence, the standard Lagrangian is

$$L = \tfrac{1}{2}[m_1 + 2m_2(1 - \cos\theta_0)]\dot{x}^2 - m_2 g x \sin\theta_0,$$

and the corresponding Lagrange equation yields the equation of motion

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\big\{[m_1+2m_2(1-\cos\theta_0)]\dot{x}\big\}+m_2 g\theta_0 = [m_1+2m_2(1-\cos\theta_0)]\ddot{x}+m_2 g \sin\theta_0,$$

in agreement with the result in the solution of Exercise 1.9.

1.15 The angles θ and ϕ, defined in Figure 1.12, are related by $b\theta = a\phi$, and the angle rotated by the cylinder (with respect to the coordinate axes), from the equilibrium position, is $\phi - \theta = (b/a - 1)\theta$. On the other hand, the tangential velocity of the center of mass of the cylinder is $(b - a)\dot{\theta}$. Hence, recalling that the kinetic energy of the cylinder is equal to the sum of the kinetic energy of the translation of the center of mass plus the kinetic energy of rotation about the axis passing through the center of mass, we have

$$T = \frac{1}{2}m[(b-a)\dot{\theta}]^2 + \frac{1}{2}I\left[\left(\frac{b}{a} - 1\right)\dot{\theta}\right]^2 = \frac{3}{4}m(b-a)^2\dot{\theta}^2,$$

where we have used the fact that the moment of inertia of a cylinder of mass m and radius a about its axis is $ma^2/2$. Thus, the standard Lagrangian is

$$L = \frac{3}{4}m(b-a)^2\dot{\theta}^2 + mg(b-a)\cos\theta.$$

From this Lagrangian we obtain the equation of motion

$$0 = \frac{3}{2}m(b-a)^2\ddot{\theta} + mg(b-a)\sin\theta,$$

which has the form of the equation of motion of a simple pendulum [cf. Equation (1.25)]. Hence, the period of the small oscillations is

$$2\pi\sqrt{\frac{3(b-a)}{2g}}.$$

1.17 Substituting Equations (1.61) into Equations (1.62) one readily verifies that these conditions are identically satisfied assuming, as usual, that the partial derivatives commute.

Conversely, if conditions (1.62) hold, letting

$$M_{ij} \equiv \frac{\partial Q_i}{\partial \dot{q}_j},$$

from the first two equations in (1.62) it follows that the M_{ij} are functions of (q_k, t) only, with $M_{ij} = -M_{ji}$. Then, the second line of (1.62) reads

$$\frac{\partial M_{ij}}{\partial q_k} + \frac{\partial M_{jk}}{\partial q_i} + \frac{\partial M_{ki}}{\partial q_j} = 0,$$

which is equivalent to the local existence of functions $\alpha_i(q_j, t)$ such that

$$M_{ij} = \frac{\partial \alpha_i}{\partial q_j} - \frac{\partial \alpha_j}{\partial q_i}.$$

Thus,

$$Q_i = M_{ik}\dot{q}_k + \mu_i = \left(\frac{\partial \alpha_i}{\partial q_k} - \frac{\partial \alpha_k}{\partial q_i}\right)\dot{q}_k + \mu_i, \tag{S.1}$$

where the μ_i are functions of (q_j, t) only. Substituting (S.1) into the last line of conditions (1.62) one obtains

$$\frac{\partial}{\partial q_j}\left(\mu_i - \frac{\partial \alpha_i}{\partial t}\right) = \frac{\partial}{\partial q_i}\left(\mu_j - \frac{\partial \alpha_j}{\partial t}\right),$$

which is equivalent to the existence of a function $\beta(q_j, t)$ such that

$$\mu_i = \frac{\partial \alpha_i}{\partial t} - \frac{\partial \beta}{\partial q_i}.$$

Substituting this last expression into (S.1) one obtains (1.61).

1.20 Making use of a set of Cartesian axes such that the gravitational acceleration is directed along the negative z-axis, the natural Lagrangian is given by

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz,$$

where g is the acceleration of gravity. The coordinates x and y are ignorable and therefore, their conjugate momenta, $p_x = m\dot{x}$ and $p_y = m\dot{y}$, are constants of motion. The coordinate z is not ignorable, but

$$\frac{\partial L}{\partial z} = -mg,$$

which is of the form (1.83) with $G = -mgt$, hence $p_z - G = m\dot{z} + mgt$ is also a constant of motion.

1.22 The Lagrangian (1.88) admits a constant of motion associated with the translations along the z-axis if the coordinate z is ignorable or if there exists a function $G(\rho, \phi, z, t)$ such that

$$\frac{\partial L}{\partial z} = \frac{\partial G}{\partial q_i}\dot{q}_i + \frac{\partial G}{\partial t},$$

that is, making use of the explicit expression of L,

$$\frac{e}{c}\frac{\partial A_i}{\partial z}\dot{q}_i = \frac{\partial G}{\partial q_i}\dot{q}_i + \frac{\partial G}{\partial t}.$$

The validity of this equation for all values of $(q_i, \dot{q}_i, t)$ is equivalent to the condition

$$\mathrm{d}G = \frac{e}{c}\frac{\partial A_i}{\partial z}\mathrm{d}q_i = \frac{e}{c}\left(\frac{\partial A_1}{\partial z}\mathrm{d}\rho + \frac{\partial A_2}{\partial z}\mathrm{d}\phi + \frac{\partial A_3}{\partial z}\mathrm{d}z\right),$$

which amounts to

$$\begin{aligned}\mathrm{d}\left(G - \frac{e}{c}A_3\right) &= \frac{e}{c}\left[\left(\frac{\partial A_1}{\partial z} - \frac{\partial A_3}{\partial \rho}\right)\mathrm{d}\rho + \left(\frac{\partial A_2}{\partial z} - \frac{\partial A_3}{\partial \phi}\right)\mathrm{d}\phi\right] \\ &= \frac{e}{c}(B_\phi \mathrm{d}\rho - \rho B_\rho \mathrm{d}\phi),\end{aligned}$$

where, as in Example 1.21, B_ρ, B_ϕ, B_z are the components of the magnetic field with respect to the orthonormal basis $(\hat{\rho}, \hat{\phi}, \hat{z})$ defined by the circular cylindrical coordinates.

Thus, the existence of G is equivalent to the exactness of $B_\phi \mathrm{d}\rho - \rho B_\rho \mathrm{d}\phi$. Making use of the fact that the divergence of $\mathbf{B}$ must be equal to zero (see Example 1.21), one finds that the exactness of $B_\phi \mathrm{d}\rho - \rho B_\rho \mathrm{d}\phi$ is equivalent to the condition that the components B_ρ, B_ϕ, B_z be independent of z. The corresponding constant of motion is

$$\frac{\partial L}{\partial \dot{z}} - G = m\dot{z} + \frac{e}{c}A_3 - G.$$

An example of a magnetic field invariant under translations along the z-axis is that of a straight wire along the z-axis, carrying a current I. Making use, e.g., of Ampère's law, one finds that the only nonzero cylindrical component of this field is $B_\phi = 2I/c\rho$. Thus,

$$\mathrm{d}\left(G - \frac{e}{c}A_3\right) = \frac{e}{c}B_\phi \mathrm{d}\rho = \frac{2eI}{c^2}\frac{\mathrm{d}\rho}{\rho},$$

and the constant of motion is $m\dot{z} - (2eI/c^2)\ln(\rho/\rho_0)$, where ρ_0 is a constant. (Note that it was not necessary to specify a vector potential.)

1.23 The Lagrangian (1.77) does not contain the variables ϕ, z, and t; therefore we immediately have three constants of motion,

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = m\rho^2\dot{\phi} + \frac{eB_0}{2c}\rho^2, \qquad p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z},$$

which correspond to the z-component of the angular momentum and the z-component of the linear momentum, respectively, and [see Equation (1.93)]

$$\dot{q}_i\frac{\partial L}{\partial \dot{q}_i} - L = \frac{m}{2}(\dot{\rho}^2 + \rho^2\dot{\phi}^2 + \dot{z}^2)$$

which is the particle's kinetic energy, T. By combining these equations one gets

$$\frac{2T}{m} - \frac{{p_z}^2}{m^2} = \dot{\rho}^2 + \left(\frac{p_\phi}{m\rho} - \frac{eB_0\rho}{2mc}\right)^2 \tag{S.2}$$

and with the aid of this equation we can find t as a function of ρ.

Alternatively, we can obtain first the orbit, eliminating the time by means of the chain rule

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{\mathrm{d}\rho}{\mathrm{d}\phi}\frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{\mathrm{d}\rho}{\mathrm{d}\phi}\left(\frac{p_\phi}{m\rho^2} - \frac{eB_0}{2mc}\right). \tag{S.3}$$

Substituting (S.3) into (S.2) we obtain

$$\frac{2T}{m} - \frac{{p_z}^2}{m^2} = \left(\frac{\mathrm{d}\rho}{\mathrm{d}\phi}\right)^2\left(\frac{p_\phi}{m\rho^2} - \frac{eB_0}{2mc}\right)^2 + \left(\frac{p_\phi}{m\rho} - \frac{eB_0\rho}{2mc}\right)^2,$$

which leads to

$$\pm\int \mathrm{d}\phi = \int \frac{\left(p_\phi - \dfrac{eB_0\rho^2}{2c}\right)\mathrm{d}\rho}{\rho\sqrt{(2mT - {p_z}^2)\rho^2 - \left(p_\phi - \dfrac{eB_0\rho^2}{2c}\right)^2}}$$

$$= \int \frac{\left(p_\phi - \dfrac{eB_0\rho^2}{2c}\right)\mathrm{d}\rho}{\rho\sqrt{\left(2mT - {p_z}^2 + \dfrac{2eB_0p_\phi}{c}\right)\rho^2 - \left(p_\phi + \dfrac{eB_0\rho^2}{2c}\right)^2}}$$

and with the change of variable

$$\cos\beta = \frac{p_\phi + \dfrac{eB_0\rho^2}{2c}}{\rho\sqrt{2mT - p_z{}^2 + \dfrac{2eB_0 p_\phi}{c}}} \tag{S.4}$$

it follows that

$$\sin\beta\,\mathrm{d}\beta = \frac{\left(p_\phi - \dfrac{eB_0\rho^2}{2c}\right)\mathrm{d}\rho}{\rho^2\sqrt{2mT - p_z{}^2 + \dfrac{2eB_0 p_\phi}{c}}}$$

and, therefore,

$$\pm(\phi - \phi_0) = \beta,$$

where ϕ_0 is an integration constant. Inserting this expression into (S.4) we obtain

$$\rho^2 + \frac{2cp_\phi}{eB_0} = \frac{2c}{eB_0}\sqrt{2mT - p_z{}^2 + \frac{2eB_0 p_\phi}{c}}\;\rho\cos(\phi - \phi_0),$$

which is the equation of a circular cylinder of radius

$$\left|\frac{c}{eB_0}\right|\sqrt{2mT - p_z{}^2},$$

whose axis is at a distance

$$\left|\frac{c}{eB_0}\right|\sqrt{2mT - p_z{}^2 + \frac{2eB_0 p_\phi}{c}}$$

from the z-axis.

The coordinate z of the particle is given by $z = z_0 + p_z t/m$, where z_0 is a constant. Hence, depending on the value of the constant p_z, the orbit may be a circle (if $p_z = 0$) or a helix (if $p_z \neq 0$).

Exercises of Chapter 2

2.1 From Equation (2.18) it follows that r has its minimum value when $\cos(\theta - \theta_0)$ has its maximum value, which is equal to 1; hence, with $e = 1$, we have $1/r_{\min} = 2mk/l^2$, that is,

$$r_{\min} = \frac{l^2}{2mk}.$$

On the other hand, when $E = 0$, Equation (2.14), with $V(r) = -k/r$, reduces to

$$\pm(t - t_0) = \int \frac{m\,\mathrm{d}r}{\sqrt{\frac{2mk}{r} - \frac{l^2}{r^2}}} = \sqrt{\frac{m}{2k}} \int \frac{r\,\mathrm{d}r}{\sqrt{r - r_{\min}}}$$

and with the change of variable $w = r - r_{\min}$ we obtain

$$\begin{aligned}\pm(t - t_0) &= \sqrt{\frac{m}{2k}} \int \frac{(w + r_{\min})\,\mathrm{d}w}{\sqrt{w}} = \sqrt{\frac{m}{2k}} \left(\frac{2}{3} w^{3/2} + 2r_{\min} w^{1/2} \right) \\ &= \frac{2}{3} \sqrt{\frac{m}{2k}} \, (r + 2r_{\min}) \sqrt{r - r_{\min}}.\end{aligned}$$

2.2 In the case of a hyperbolic orbit, the minimum value of r is $(e - 1)a$ and from Equation (2.18) we obtain

$$e^2 - 1 = \frac{l^2}{mka}.$$

Combining this relation with (2.19) we find that $2E = k/a$. Then, from Equation (2.14), eliminating E and l with the aid of the foregoing equations, we find that

$$\pm(t - t_0) = \sqrt{\frac{ma}{k}} \int \frac{r\,\mathrm{d}r}{\sqrt{(r + a)^2 - a^2 e^2}},$$

where t_0 is an integration constant. With the change of variable $r + a = ae \cosh \psi$, we obtain the analog of Kepler's equation (2.22)

$$t - t_0 = \sqrt{\frac{ma^3}{k}} (e \sinh \psi - \psi).$$

2.5 Making use of the definitions given in Example 2.4, the position vector of the freely falling particle can be expressed in the form

$$\mathbf{r} = R\hat{r} + x\hat{\theta} + y\hat{\phi} + z\hat{r}$$

[cf. Equation (2.34)] and with the aid of Equations (2.35) we find that the velocity of the particle with respect to the inertial frame is

$$\dot{\mathbf{r}} = (\dot{x} - \omega y \cos\theta_0)\hat{\theta} + \big[\dot{y} + \omega x \cos\theta_0 + \omega(R + z)\sin\theta_0\big]\hat{\phi} + (\dot{z} - \omega y \sin\theta_0)\hat{r}.$$

Thus, neglecting the quadratic terms in ω, the Lagrangian of the particle is approximately given by

$$L \simeq \frac{m}{2}\left\{(\dot{x}-\omega y\cos\theta_0)^2+\left[\dot{y}+\omega x\cos\theta_0+\omega(R+z)\sin\theta_0\right]^2+(\dot{z}-\omega y\sin\theta_0)^2\right\} - mgz$$

and the Lagrange equations give

$$\ddot{x} = 2\omega\cos\theta_0\dot{y}, \qquad \ddot{y} = -2\omega\cos\theta_0\dot{x} - 2\omega\sin\theta_0\dot{z}, \qquad \ddot{z} = 2\omega\sin\theta_0\dot{y} - g. \tag{S.5}$$

These equations can be partially integrated to give

$$\dot{x} = 2\omega\cos\theta_0 y, \qquad \dot{y} = -2\omega\cos\theta_0 x - 2\omega\sin\theta_0(z-h), \qquad \dot{z} = 2\omega\sin\theta_0 y - gt, \tag{S.6}$$

where we have taken into account the fact that, at $t = 0$, the particle was at rest at the point $x = 0$, $y = 0$, $z = h$.

In order to solve the set of first-order equations (S.6) we decouple them by calculating the third derivative of, e.g., z with respect to the time. By combining Equations (S.5)–(S.6) we obtain the inhomogeneous, linear, third-order ODE

$$\begin{aligned}\dddot{z} &= 2\omega\sin\theta_0\ddot{y} = 2\omega\sin\theta_0(-2\omega\cos\theta_0\dot{x} - 2\omega\sin\theta_0\dot{z}) \\ &= -4\omega^2\sin\theta_0\cos\theta_0(2\omega\cos\theta_0 y) - 4\omega^2\sin^2\theta_0\dot{z} \\ &= -4\omega^2\cos^2\theta_0(\dot{z}+gt) - 4\omega^2\sin^2\theta_0\dot{z} \\ &= -4\omega^2\dot{z} - 4\omega^2\cos^2\theta_0 gt.\end{aligned}$$

One readily finds that the general solution of this equation has the form $z = -\frac{1}{2}\cos^2\theta_0 gt^2 + c_1\cos 2\omega t + c_2\sin 2\omega t + c_3$, where c_1, c_2, and c_3 are constants. Since $z = h$ at $t = 0$, we have $c_1 + c_3 = h$, which implies that

$$z = h - \tfrac{1}{2}\cos^2\theta_0 gt^2 + c_1(\cos 2\omega t - 1) + c_2\sin 2\omega t.$$

Then, the third equation in (S.6) gives

$$y = \frac{\dot{z}+gt}{2\omega\sin\theta_0} = \frac{\sin^2\theta_0 gt - 2\omega c_1\sin 2\omega t + 2\omega c_2\cos 2\omega t}{2\omega\sin\theta_0}$$

and from the conditions $y = 0$ and $\dot{y} = 0$ at $t = 0$ we find that $c_2 = 0$ and $c_1 = g\sin^2\theta_0/4\omega^2$, respectively. Thus,

$$z = h - \frac{1}{2}\cos^2\theta_0 gt^2 + \frac{g\sin^2\theta_0}{4\omega^2}(\cos 2\omega t - 1), \qquad y = \frac{\sin\theta_0}{2\omega}gt - \frac{g\sin\theta_0}{4\omega^2}\sin 2\omega t.$$

Finally, from the second equation in (S.6) we obtain

$$x = \frac{1}{2}\sin\theta_0\cos\theta_0 g t^2 + \frac{g\sin\theta_0\cos\theta_0}{4\omega^2}(\cos 2\omega t - 1).$$

In the limit where $\omega t \ll 1$, the expressions obtained above reduce to

$$z \simeq h - \frac{gt^2}{2}, \qquad y \simeq \frac{\omega g\sin\theta_0 t^3}{3}, \qquad x \simeq \frac{g\sin\theta_0\cos\theta_0\omega^2 t^4}{6}.$$

2.7 Substituting $Q(q,\dot{q},t) = -2m\gamma\dot{q} - m\omega^2 q$ into the second equation in (1.62) one obtains $\gamma = 0$.

2.9 In the case of the Poisson–Boltzmann equation, $\ddot{q} = -k\dot{q}/t - a\mathrm{e}^q$, Equation (2.44) takes the form

$$\frac{\partial M}{\partial t} + \dot{q}\frac{\partial M}{\partial q} + \left(-\frac{k}{t}\dot{q} - a\mathrm{e}^q\right)\frac{\partial M}{\partial \dot{q}} = M\frac{k}{t}.$$

A particular solution of this equation is $M = t^k$, which, substituted into (2.43), gives

$$L = \tfrac{1}{2}t^k\dot{q}^2 + g(q,t)\dot{q} + h(q,t),$$

where $g(q,t)$ and $h(q,t)$ are functions of q and t only. Substituting this last expression into (2.42), with $f(q,\dot{q},t) = -k\dot{q}/t - a\mathrm{e}^q$, we find

$$\frac{\partial g}{\partial t} - a\mathrm{e}^q t^k - \frac{\partial h}{\partial q} = 0,$$

which, rewritten in the form

$$\frac{\partial g}{\partial t} = \frac{\partial(h + a\mathrm{e}^q t^k)}{\partial q},$$

amounts to the existence of a function $F(q,t)$ such that $g = \partial F/\partial q$ and $h + a\mathrm{e}^q t^k = \partial F/\partial q$ and, therefore,

$$L = \frac{1}{2}t^k\dot{q}^2 - at^k\mathrm{e}^q + \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial t}.$$

The function F is arbitrary and can be taken equal to zero.

2.10 Substituting the function $f(q,\dot{q},t) = -\frac{9}{4}t\dot{q}^{5/3}$ into Equation (2.44) we obtain the equation

$$\frac{\partial M}{\partial t} + \dot{q}\frac{\partial M}{\partial q} - \frac{9}{4}t\dot{q}^{5/3}\frac{\partial M}{\partial \dot{q}} = M\frac{15}{4}t\dot{q}^{2/3},$$

for the function M defined in (2.43). By inspection, we find that a solution of this equation is

$$M = -\dot{q}^{-5/3}$$

(the minus sign is introduced for convenience). Hence

$$L = \frac{9}{2}\dot{q}^{1/3} + g(q,t)\dot{q} + h(q,t),$$

where $g(q,t)$ and $h(q,t)$ are functions of q and t only. Substituting this expression into the Lagrange equation (2.42) we have

$$\frac{9}{4}t + \frac{\partial g}{\partial t} - \frac{\partial h}{\partial q} = 0$$

or, equivalently,

$$\frac{\partial g}{\partial t} = \frac{\partial}{\partial q}\left(h - \frac{9}{4}tq\right),$$

which amounts to the existence of a function $F(q,t)$ such that $g = \partial F/\partial q$ and $h - \frac{9}{4}tq = \partial F/\partial t$. Thus,

$$L = \frac{9}{2}\dot{q}^{1/3} + \frac{9}{4}tq + \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial t}.$$

Since the function F is arbitrary, we can take it equal to zero.

2.12 Starting from $M = m\mathrm{e}^{2\gamma t}$ we obtain $L = \frac{1}{2}m\mathrm{e}^{2\gamma t}\dot{q}^2 + g(q,t)\dot{q} + h(q,t)$, where g and h are functions of q and t only. Substituting this expression into the Lagrange equation, with $\ddot{q} = -2\gamma\dot{q}$ [see (2.52)], we get

$$\frac{\partial g}{\partial t} = \frac{\partial h}{\partial q}$$

which implies the existence of a function $F(q,t)$ such that $g = \partial F/\partial q$, $h = \partial F/\partial t$ and, therefore,

$$L = \frac{1}{2}m\mathrm{e}^{2\gamma t}\dot{q}^2 + \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial t}.$$

Again, the function F is arbitrary and can be taken equal to zero.

2.15 In order to prove that Equations (2.66) are a consequence of Equations (2.67), we multiply both sides of (2.67) by $\dot{q}_i$ obtaining

$$\begin{aligned} 0 &= g_{ij}\dot{q}_i\ddot{q}_j + \frac{\partial g_{ij}}{\partial q_k}\dot{q}_i\dot{q}_k\dot{q}_j - \frac{1}{2}\frac{\partial g_{jk}}{\partial q_i}\dot{q}_i\dot{q}_j\dot{q}_k \\ &= g_{ij}\dot{q}_i\ddot{q}_j + \frac{1}{2}\frac{\partial g_{ij}}{\partial q_k}\dot{q}_k\dot{q}_i\dot{q}_j \\ &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(g_{ij}\dot{q}_i\dot{q}_j), \end{aligned}$$

which means that $g_{ij}\dot{q}_i\dot{q}_j$ is a constant.

On the other hand, since the Lagrangian $L = \frac{1}{2}(g_{ij}\dot{q}_i\dot{q}_j)^p$ does not contain t, the Jacobi integral (1.93) is a constant of motion, which is given by

$$\dot{q}_i\left[p\,(g_{rs}\dot{q}_r\dot{q}_s)^{p-1}g_{ij}\dot{q}_j\right] - \tfrac{1}{2}(g_{ij}\dot{q}_i\dot{q}_j)^p = (p-\tfrac{1}{2})(g_{ij}\dot{q}_i\dot{q}_j)^p.$$

Thus, for $p \neq \frac{1}{2}$, the conservation of the Jacobi integral is equivalent to Equation (2.66).

2.17 The equations for the geodesics of the sphere are given by the Euler–Lagrange equation for the Lagrangian

$$L = \frac{a^2}{2}(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2), \tag{S.7}$$

where a is the radius of the sphere and θ, ϕ are the standard spherical coordinates. One can take advantage of the conservation of the momentum conjugate to ϕ and of L, which leads to a single first-order ODE for θ as a function of ϕ, following the steps in Example 2.16.

A simpler procedure consists in substituting the Lagrangian (S.7) into the Euler–Lagrange equations, which yields

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{\theta}} - \frac{\partial L}{\partial\theta} = \frac{\mathrm{d}}{\mathrm{d}t}(a^2\dot{\theta}) - a^2\sin\theta\cos\theta\dot{\phi}^2, \\ 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{\phi}} - \frac{\partial L}{\partial\phi} = \frac{\mathrm{d}}{\mathrm{d}t}(a^2\sin^2\theta\dot{\phi}) \end{aligned} \tag{S.8}$$

[cf. Equations (2.3)]. The second equation implies that

$$L_3 \equiv a^2\sin^2\theta\dot{\phi} \tag{S.9}$$

is conserved. If $L_3 = 0$, then ϕ is a constant, which corresponds to some plane passing through the origin and containing the z-axis. When $L_3 \neq 0$, by virtue of (S.9) we can replace the derivatives with respect to the time by derivatives with respect to ϕ. Then, the first equation in (S.8) takes the form

$$0 = \frac{L_3}{a^2 \sin^2\theta} \frac{\mathrm{d}}{\mathrm{d}\phi} \left(a^2 \frac{L_3}{a^2 \sin^2\theta} \frac{\mathrm{d}\theta}{\mathrm{d}\phi} \right) - a^2 \sin\theta \cos\theta \left(\frac{L_3}{a^2 \sin^2\theta} \right)^2,$$

which reduces to

$$\frac{\mathrm{d}^2}{\mathrm{d}\phi^2} \cot\theta = -\cot\theta.$$

The general solution of this equation is

$$\cot\theta = A\cos\phi + B\sin\phi,$$

where A and B are constants. Making use of the relation between the spherical coordinates and the Cartesian ones, the last equation amounts to $z = Ax + By$, which corresponds to a plane passing through the origin.

2.21 Substituting the Lagrangian

$$L = \frac{9}{2}\dot{q}^{1/3} + \frac{9}{4}tq$$

into Equation (2.78) we obtain the PDE

$$\frac{9}{4}t\eta + \frac{3}{2}\dot{q}^{-2/3} \left(\frac{\partial\eta}{\partial t} + \dot{q}\frac{\partial\eta}{\partial q} - \dot{q}\frac{\partial\xi}{\partial t} - \dot{q}^2\frac{\partial\xi}{\partial q} \right) + \frac{9}{4}q\xi + \left(\frac{9}{2}\dot{q}^{1/3} + \frac{9}{4}tq \right) \left(\frac{\partial\xi}{\partial t} + \dot{q}\frac{\partial\xi}{\partial q} \right) = \frac{\partial G}{\partial t} + \dot{q}\frac{\partial G}{\partial q}.$$

Equating the coefficients of the various powers of $\dot{q}$ on both sides of the equation we obtain the system of equations

$$3\frac{\partial\xi}{\partial q} = 0, \tag{S.10}$$

$$\frac{9}{4}tq\frac{\partial\xi}{\partial q} = \frac{\partial G}{\partial q}, \tag{S.11}$$

$$\frac{3}{2}\frac{\partial\eta}{\partial q} + 3\frac{\partial\xi}{\partial t} = 0, \tag{S.12}$$

$$\frac{3}{2}\frac{\partial\eta}{\partial t} = 0, \tag{S.13}$$

$$\frac{9}{4}t\eta + \frac{9}{4}q\xi + \frac{9}{4}tq\frac{\partial\xi}{\partial t} = \frac{\partial G}{\partial t}. \tag{S.14}$$

From Equations (S.10) and (S.11) we find that $\xi = A(t)$, where A is a function of one variable, and $G = G(t)$. Then, Equation (S.12) gives

$$\eta = -2q\frac{\mathrm{d}A}{\mathrm{d}t} + B(t),$$

where B is another function of one variable. Substituting this expression into (S.13) we get $A = c_1 t$ and $B = c_2$, where c_1 and c_2 are two arbitrary constants. Thus,

$$\eta = -2c_1 q + c_2$$

and from (S.14) we obtain

$$\frac{\partial G}{\partial t} = \frac{9}{4}c_2 t,$$

i.e., $G = \frac{9}{8}c_2 t^2$, up to an irrelevant additive constant.

Hence, from Equation (2.77), it follows that

$$3c_1\big(-q\dot{q}^{-2/3} + t\dot{q}^{1/3} + \tfrac{3}{4}t^2 q\big) + \tfrac{3}{8}c_2(4\dot{q}^{-2/3} - 3t^2)$$

is a constant of motion. Since c_1 and c_2 are arbitrary, each of the expressions inside the parentheses is a constant of motion. Noting that the first one depends on q, while the second one does not, one concludes that they are functionally independent. In fact, eliminating $\dot{q}$ from these two constants of motion one can find an expression for q as a function of t and the two constants of motion, which is the general solution of the original equation.

2.22 Substituting the Lagrangian (1.54) into Equation (2.78) one obtains the basic equation

$$\begin{aligned}\big(m\omega^2 r\sin^2\theta_0 - mg\cos\theta_0\big)\eta + m\dot{r}\left(\frac{\partial\eta}{\partial t} + \dot{r}\frac{\partial\eta}{\partial r} - \dot{r}\frac{\partial\xi}{\partial t} - \dot{r}^2\frac{\partial\xi}{\partial r}\right)\\ + \left(\frac{m}{2}\dot{r}^2 + \frac{m}{2}\omega^2 r^2\sin^2\theta_0 - mgr\cos\theta_0\right)\left(\frac{\partial\xi}{\partial t} + \dot{r}\frac{\partial\xi}{\partial r}\right) = \frac{\partial G}{\partial t} + \dot{r}\frac{\partial G}{\partial r},\end{aligned}$$

which has to be satisfied for all values of r, $\dot{r}$, and t. Equating the coefficients of the various powers of $\dot{r}$ on each side of this equation we obtain the four equations

$$-\frac{m}{2}\frac{\partial\xi}{\partial r} = 0, \qquad \text{(S.15)}$$

$$m\frac{\partial\eta}{\partial r} - \frac{m}{2}\frac{\partial\xi}{\partial t} = 0, \qquad \text{(S.16)}$$

$$m\frac{\partial\eta}{\partial t} + \left(\frac{m}{2}\omega^2 r^2\sin^2\theta_0 - mgr\cos\theta_0\right)\frac{\partial\xi}{\partial r} = \frac{\partial G}{\partial r}, \qquad \text{(S.17)}$$

$$\big(m\omega^2 r\sin^2\theta_0 - mg\cos\theta_0\big)\eta + \left(\frac{m}{2}\omega^2 r^2\sin^2\theta_0 - mgr\cos\theta_0\right)\frac{\partial\xi}{\partial t} = \frac{\partial G}{\partial t}. \qquad \text{(S.18)}$$

Equations (S.15) and (S.16) imply that $\xi = A(t)$, where $A(t)$ is a function of t only and

$$\eta = \frac{r}{2}\frac{\mathrm{d}A}{\mathrm{d}t} + B(t),$$

where $B(t)$ is a function of t only. Then, from Equations (S.17) and (S.18) we obtain

$$\begin{aligned}\frac{\partial G}{\partial r} &= m\left(\frac{r}{2}\frac{\mathrm{d}^2A}{\mathrm{d}t^2} + \frac{\mathrm{d}B}{\mathrm{d}t}\right),\\ \frac{\partial G}{\partial t} &= \left(m\omega^2 r\sin^2\theta_0 - mg\cos\theta_0\right)\left(\frac{r}{2}\frac{\mathrm{d}A}{\mathrm{d}t} + B\right)\\ &\quad + \left(\frac{m}{2}\omega^2 r^2\sin^2\theta_0 - mgr\cos\theta_0\right)\frac{\mathrm{d}A}{\mathrm{d}t}.\end{aligned} \tag{S.19}$$

The equality of the mixed partial derivatives of G leads to the ODEs

$$\frac{\mathrm{d}^3A}{\mathrm{d}t^3} = 4\omega^2\sin^2\theta_0\frac{\mathrm{d}A}{\mathrm{d}t}, \qquad \frac{\mathrm{d}^2B}{\mathrm{d}t^2} - \omega^2\sin^2\theta_0 B = -\frac{3}{2}g\cos\theta_0\frac{\mathrm{d}A}{\mathrm{d}t}.$$

The general solution of these equations contains five arbitrary constants and therefore the Lagrangian under consideration possesses five one-parameter families of variational symmetries.

Restricting ourselves to the variational symmetries with $\xi = 0$, we have $A = 0$ and $B(t) = a_1\mathrm{e}^{\omega t\sin\theta_0} + a_2\mathrm{e}^{-\omega t\sin\theta_0}$, where a_1 and a_2 are two arbitrary constants. Thus, from Equations (S.19) we find, up to an irrelevant constant term,

$$G = \left(m\omega r\sin\theta_0 - \frac{mg\cos\theta_0}{\omega\sin\theta_0}\right)(a_1\mathrm{e}^{\omega t\sin\theta_0} - a_2\mathrm{e}^{-\omega t\sin\theta_0}).$$

Substituting this expression and $\eta = a_1\mathrm{e}^{\omega t\sin\theta_0} + a_2\mathrm{e}^{-\omega t\sin\theta_0}$ into (2.77) we obtain the constant of motion

$$\begin{aligned}m\dot{r}\eta - G &= a_1\mathrm{e}^{\omega t\sin\theta_0}\left(m\dot{r} - m\omega r\sin\theta_0 + \frac{mg\cos\theta_0}{\omega\sin\theta_0}\right)\\ &\quad + a_2\mathrm{e}^{-\omega t\sin\theta_0}\left(m\dot{r} + m\omega r\sin\theta_0 - \frac{mg\cos\theta_0}{\omega\sin\theta_0}\right).\end{aligned}$$

Since a_1 and a_2 are arbitrary, we actually have two constants of motion which are functionally independent and, except for constant factors, coincide with the constants of motion given by Equations (1.20).

Finally, we can find one-parameter families of variational symmetries proceeding as in Examples 2.19 and 2.20. If $\xi = 0$ and $\eta = a_1 e^{\omega t \sin\theta_0} + a_2 e^{-\omega t \sin\theta_0}$, from (2.58) we propose

$$\frac{\partial t^{(s)}}{\partial s} = 0, \qquad \frac{\partial q^{(s)}}{\partial s} = a_1 e^{\omega t^{(s)} \sin\theta_0} + a_2 e^{-\omega t^{(s)} \sin\theta_0}.$$

The solution of these equations is readily found and is given by

$$t^{(s)} = t^{(0)}, \qquad q^{(s)} = q^{(0)} + a_1 s e^{\omega t^{(0)} \sin\theta_0} + a_2 s e^{-\omega t^{(0)} \sin\theta_0}.$$

Exercises of Chapter 3

3.2 Making use of the definition (3.15) of the components of the inertia tensor, we have

$$I_{1'1'} = m \sum_{\alpha=1}^{4} (x_{2'}^{(\alpha)} x_{2'}^{(\alpha)} + x_{3'}^{(\alpha)} x_{3'}^{(\alpha)}) = 2ma^2,$$

$$I_{1'2'} = -m \sum_{\alpha=1}^{4} x_{1'}^{(\alpha)} x_{2'}^{(\alpha)} = -ma^2,$$

$$I_{1'3'} = -m \sum_{\alpha=1}^{4} x_{1'}^{(\alpha)} x_{3'}^{(\alpha)} = 0,$$

$$I_{2'2'} = m \sum_{\alpha=1}^{4} (x_{1'}^{(\alpha)} x_{1'}^{(\alpha)} + x_{3'}^{(\alpha)} x_{3'}^{(\alpha)}) = 2ma^2,$$

$$I_{2'3'} = -m \sum_{\alpha=1}^{4} x_{2'}^{(\alpha)} x_{3'}^{(\alpha)} = 0,$$

$$I_{3'3'} = m \sum_{\alpha=1}^{4} (x_{1'}^{(\alpha)} x_{1'}^{(\alpha)} + x_{2'}^{(\alpha)} x_{2'}^{(\alpha)}) = 4ma^2,$$

that is,

$$(I_{i'j'}) = ma^2 \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

The principal moments of inertia are the eigenvalues of $(I_{i'j'})$ and, therefore, they are the roots of the polynomial

$$0 = \begin{vmatrix} 2ma^2 - \lambda & -ma^2 & 0 \\ -ma^2 & 2ma^2 - \lambda & 0 \\ 0 & 0 & 4ma^2 - \lambda \end{vmatrix} = (4ma^2 - \lambda)(\lambda^2 - 4ma^2\lambda + 3m^2a^4).$$

Thus, the principal moments of inertia are

$$ma^2, \qquad 3ma^2, \qquad 4ma^2.$$

Substituting these eigenvalues into Equation (3.21) one obtains the unit eigenvectors

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

respectively, which define the principal axes.

3.3 Making use of Equations (3.14) and (3.24) we have

$$a_{is'}\dot{a}_{js'} = a_{is'}\varepsilon_{s'k'r'}\omega_{r'}a_{jk'} = \varepsilon_{ijm}a_{mr'}\omega_{r'} = \varepsilon_{ijm}\omega_m.$$

The last equality follows from the fact that the components of any vector transform as in (3.5).

3.4 According to the hypothesis, the vector $(R_{1'}, R_{2'}, R_{3'})$, which joins O and the center of mass, lies along one principal axis at the center of mass, that is,

$$I^{\text{CM}}_{i'j'}R_{j'} = \lambda R_{i'},$$

where λ is one of the principal moments of inertia at the center of mass. Then, making use of (3.27) we have

$$\begin{aligned} I_{i'j'}R_{j'} &= [M(\mathbf{R}^2\delta_{i'j'} - R_{i'}R_{j'}) + I^{\text{CM}}_{i'j'}]R_{j'} \\ &= M(\mathbf{R}^2R_{i'} - \mathbf{R}^2R_{i'}) + \lambda R_{i'} \\ &= \lambda R_{i'}. \end{aligned}$$

Thus, not only $(I_{i'j'})$ and $(I^{\text{CM}}_{i'j'})$ share one principal axis, they also share the principal moment of inertia in this direction.

If $v_{i'}$ corresponds to a principal axis at the center of mass, orthogonal to that defined by $R_{i'}$, we have $I^{\text{CM}}_{i'j'}v_{j'} = \mu v_{i'}$ and $R_{i'}v_{i'} = 0$ and, therefore,

$$
\begin{aligned}
I_{i'j'}v_{j'} &= [M(\mathbf{R}^2\delta_{i'j'} - R_{i'}R_{j'}) + I^{\rm CM}_{i'j'}]v_{j'} \\
&= M\mathbf{R}^2 v_{i'} + \mu v_{i'} \\
&= (M\mathbf{R}^2 + \mu)v_{i'},
\end{aligned}
$$

showing that $v_{i'}$ also defines a principal axis at O.

3.5 According to the hypothesis, two of the principal moments of inertia coincide. As it is customary, we shall assume that I_1 is equal to I_2. Then, from the last equation in (3.35) it follows that $\omega_{3'}$ is constant. Denoting by a this constant value, the first two equations in (3.35) take the form

$$
\begin{aligned}
\frac{\mathrm{d}\omega_{1'}}{\mathrm{d}t} &= \frac{(I_1 - I_3)a}{I_1}\omega_{2'}, \\
\frac{\mathrm{d}\omega_{2'}}{\mathrm{d}t} &= -\frac{(I_1 - I_3)a}{I_1}\omega_{1'}.
\end{aligned}
$$

The solution of this system is given by $\omega_{1'} = (\omega_{1'})_0 \cos \Omega t + (\omega_{2'})_0 \sin \Omega t$, $\omega_{2'} = (\omega_{2'})_0 \cos \Omega t - (\omega_{1'})_0 \sin \Omega t$, where $\Omega \equiv (I_1 - I_3)a/I_1$, and $(\omega_{1'})_0$, $(\omega_{2'})_0$ are the values of $\omega_{1'}$, $\omega_{2'}$ at $t = 0$. This means that the angular velocity rotates about $\mathbf{e}_{3'}$ with angular velocity Ω.

3.6 Multiplying Equations (3.35) by $\omega_{1'}, \omega_{2'}, \omega_{3'}$, respectively and adding the results one obtains

$$
\omega_{1'} I_1 \frac{\mathrm{d}\omega_{1'}}{\mathrm{d}t} + \omega_{2'} I_2 \frac{\mathrm{d}\omega_{2'}}{\mathrm{d}t} + \omega_{3'} I_3 \frac{\mathrm{d}\omega_{3'}}{\mathrm{d}t} = 0,
$$

which means that $\frac{1}{2}(I_1{\omega_{1'}}^2 + I_2{\omega_{2'}}^2 + I_3{\omega_{3'}}^2)$ is a constant of motion. In a similar manner, multiplying Equations (3.35) by $I_1\omega_{1'}$, $I_2\omega_{2'}$, $I_3\omega_{3'}$, respectively and adding the results one gets

$$
{I_1}^2\omega_{1'} \frac{\mathrm{d}\omega_{1'}}{\mathrm{d}t} + {I_2}^2\omega_{2'} \frac{\mathrm{d}\omega_{2'}}{\mathrm{d}t} + {I_3}^2\omega_{3'} \frac{\mathrm{d}\omega_{3'}}{\mathrm{d}t} = 0,
$$

which means that ${I_1}^2{\omega_{1'}}^2 + {I_2}^2{\omega_{2'}}^2 + {I_3}^2{\omega_{3'}}^2$ is also conserved.

Exercises of Chapter 4

4.4 According to the definition (4.5), the momentum conjugate to q is $p = m \ln \dot{q}$ and, therefore, the Hamiltonian is given by

$$
H = (m \ln \dot{q})\dot{q} - m(\dot{q} \ln \dot{q} - \dot{q}) + 2m\gamma q = m\dot{q} + 2m\gamma q = m\mathrm{e}^{p/m} + 2m\gamma q.
$$

The Hamilton equations give

$$\dot{q} = \frac{\partial H}{\partial p} = \mathrm{e}^{p/m}, \qquad \dot{p} = -\frac{\partial H}{\partial q} = -2m\gamma.$$

The first of these equations is equivalent to the expression for the generalized momentum obtained above and the second one can be immediately integrated giving $p = p_0 - 2m\gamma t$, where p_0 is an integration constant that represents the value of p at $t = 0$. With this expression at hand, the equation for q can be readily integrated, and we obtain

$$q = \frac{\mathrm{e}^{p_0/m}}{2\gamma}(1 - \mathrm{e}^{-2\gamma t}) + q_0,$$

where the constant q_0 is the value of q at $t = 0$.

4.8 Taking into account Equation (4.23), we compute $\partial\Phi/\partial t + \{\Phi, H\}$, using that $H(q_i, p_i, t) = f(t)\Phi(q_i, p_i)$, we have

$$\frac{\partial \Phi}{\partial t} + \{\Phi, H\} = 0 + \{\Phi, f(t)\Phi\} = f(t)\{\Phi, \Phi\} = 0,$$

which shows that, indeed, Φ is a constant of motion.

4.10 Making use of the bilinearity of the Poisson bracket and (4.29) we have, e.g.,

$$\begin{aligned}\{\pi_1, \pi_2\} &= \{p_1 - \frac{e}{c}A_1, p_2 - \frac{e}{c}A_2\} = -\frac{e}{c}\{p_1, A_2\} - \frac{e}{c}\{A_1, p_2\} = \frac{e}{c}\left(\frac{\partial A_2}{\partial x} - \frac{\partial A_1}{\partial y}\right)\\ &= \frac{e}{c}B_3.\end{aligned}$$

The other relations can be derived from this one by cyclic permutations of the indices 1, 2, 3.

4.12 Making use of the properties of the Poisson bracket and Equations (4.36) we have

$$\begin{aligned}\{L_i, L_j\} &= \{a_{ik'}L_{k'}, a_{jm'}L_{m'}\}\\ &= a_{ik'}a_{jm'}\{L_{k'}, L_{m'}\} + a_{ik'}L_{m'}\{L_{k'}, a_{jm'}\} + a_{jm'}L_{k'}\{a_{ik'}, L_{m'}\}\\ &= -\varepsilon_{k'm'r'}a_{ik'}a_{jm'}L_{r'} - a_{ik'}L_{m'}\{a_{jm'}, L_{k'}\} + a_{jm'}L_{k'}\{a_{ik'}L_{m'}\}. \qquad \text{(S.20)}\end{aligned}$$

On the other hand, according to (3.31) and (4.35), using the fact that the $a_{jm'}$ are functions of the q_i only,

$$\{a_{jm'}, L_{k'}\} = \frac{\partial a_{jm'}}{\partial q_s}\frac{\partial L_{k'}}{\partial p_s}$$

$$= a_{jr'}\varepsilon_{m'r'i'}M_{i's}(M^{-1})_{sk'}$$
$$= a_{jr'}\varepsilon_{m'r'i'}\delta_{i'k'}$$
$$= a_{jr'}\varepsilon_{m'r'k'}. \tag{S.21}$$

Substituting this result, together with (3.24), into (S.20) we obtain

$$\begin{aligned}\{L_i, L_j\} &= -\varepsilon_{k'm'r'}a_{ik'}a_{jm'}L_{r'} - a_{ik'}L_{m'}a_{jr'}\varepsilon_{m'r'k'} + a_{jm'}L_{k'}a_{is'}\varepsilon_{k's'm'}\\ &= -\varepsilon_{r'k'm'}a_{ik'}a_{jm'}L_{r'} - \varepsilon_{m'r'k'}a_{jr'}a_{ik'}L_{m'} + \varepsilon_{k's'm'}a_{is'}a_{jm'}L_{k'}\\ &= -a_{sr'}\varepsilon_{ijs}L_{r'} - a_{sm'}\varepsilon_{jis}L_{m'} + a_{rk'}\varepsilon_{ijr}L_{k'}\\ &= \varepsilon_{ijk}L_k.\end{aligned}$$

Similarly, making use of the properties of the Poisson bracket, (4.36) and (S.21),

$$\begin{aligned}\{L_i, L_{j'}\} &= \{a_{ik'}L_{k'}, L_{j'}\}\\ &= a_{ik'}\{L_{k'}, L_{j'}\} + \{a_{ik'}, L_{j'}\}L_{k'}\\ &= -a_{ik'}\varepsilon_{k'j'm'}L_{m'} + a_{is'}\varepsilon_{k's'j'}L_{k'}\\ &= -a_{ik'}\varepsilon_{k'j'm'}L_{m'} + a_{is'}\varepsilon_{s'j'k'}L_{k'}\\ &= 0.\end{aligned}$$

4.14 The differential form $B_2\mathrm{d}z - B_3\mathrm{d}y$, appearing in the definition of $\mathscr{P}_1$, is (locally) exact. In fact, computing the crossed partial derivatives (taking into account that the coefficient of dx is equal to zero) we have the conditions

$$0 = \frac{\partial B_3}{\partial x}, \qquad 0 = \frac{\partial B_2}{\partial x}, \qquad \frac{\partial B_2}{\partial y} = -\frac{\partial B_3}{\partial z},$$

which, according to the hypothesis, reduces to $\partial B_3/\partial z = 0$, but this is precisely what one obtains from the fact that the divergence of **B** must be equal to zero. In a similar manner one concludes that $B_3\mathrm{d}x - B_1\mathrm{d}z$ is locally exact, thus showing that $\mathscr{P}_1$ and $\mathscr{P}_2$ are well-defined functions (even though each of them is defined up to a constant term, arising from the integration).

In order to prove that $\mathscr{P}_1$ and $\mathscr{P}_2$ are constants of motion, taking into account that they do not depend explicitly on the time, we only have to show that their Poisson brackets with the Hamiltonian are equal to zero. The Hamiltonian is given by (see Example 4.3)

$$H = \frac{1}{2m}\left({\pi_1}^2 + {\pi_2}^2 + {\pi_3}^2\right)$$

and, therefore, for example,

$$\{\mathscr{P}_1, H\} = \frac{1}{m}\{\mathscr{P}_1, \pi_1\}\pi_1 + \frac{1}{m}\{\mathscr{P}_1, \pi_2\}\pi_2 + \frac{1}{m}\{\mathscr{P}_1, \pi_3\}\pi_3.$$

Then, making use of (4.32) and (4.29), we find that each term of the last expression separately vanishes,

$$\{\mathscr{P}_1, \pi_1\} = \{\pi_1 + \frac{e}{c}\int (B_2 \mathrm{d}z - B_3 \mathrm{d}y), \pi_1\} = \frac{e}{c}\frac{\partial}{\partial x}\int (B_2 \mathrm{d}z - B_3 \mathrm{d}y) = 0,$$

$$\{\mathscr{P}_1, \pi_2\} = \{\pi_1 + \frac{e}{c}\int (B_2 \mathrm{d}z - B_3 \mathrm{d}y), \pi_2\} = \frac{e}{c}B_3 + \frac{e}{c}\frac{\partial}{\partial y}\int (B_2 \mathrm{d}z - B_3 \mathrm{d}y) = 0,$$

$$\{\mathscr{P}_1, \pi_3\} = \{\pi_1 + \frac{e}{c}\int (B_2 \mathrm{d}z - B_3 \mathrm{d}y), \pi_3\} = -\frac{e}{c}B_2 + \frac{e}{c}\frac{\partial}{\partial z}\int (B_2 \mathrm{d}z - B_3 \mathrm{d}y) = 0.$$

Finally, the Poisson bracket between $\mathscr{P}_1$ and $\mathscr{P}_2$ is

$$\begin{aligned}\{\mathscr{P}_1, \mathscr{P}_2\} &= \{\pi_1, \pi_2\} + \{\pi_1, \frac{e}{c}\int (B_3 \mathrm{d}x - B_1 \mathrm{d}z)\} + \{\frac{e}{c}\int (B_2 \mathrm{d}z - B_3 \mathrm{d}y), \pi_2\}\\ &= \frac{e}{c}B_3 - \frac{e}{c}\frac{\partial}{\partial x}\int (B_3 \mathrm{d}x - B_1 \mathrm{d}z) + \frac{e}{c}\frac{\partial}{\partial y}\int (B_2 \mathrm{d}z - B_3 \mathrm{d}y)\\ &= -\frac{e}{c}B_3,\end{aligned}$$

which is a trivial constant since, as pointed out above, B_3 is a constant. It may be remarked that H, $\mathscr{P}_1$, and $\mathscr{P}_2$ are gauge-independent.

4.15 Making use of the expression of D in terms of L_3 and the Cartesian coordinates and momenta, with the aid of (4.29) and the fact that $\{L_3, f(r)\} = 0$ for any function of r, we have

$$\begin{aligned}\{L_3, D\} &= \{L_3, \frac{1}{m}L_3 p_y - k\frac{x}{r}\}\\ &= \frac{1}{m}L_3\{L_3, p_y\} - \frac{k}{r}\{L_3, x\} - kx\{L_3, \frac{1}{r}\}\\ &= \frac{1}{m}L_3\frac{\partial L_3}{\partial y} + \frac{k}{r}\frac{\partial L_3}{\partial p_x}\\ &= \frac{1}{m}L_3(-p_x) + \frac{k}{r}(-y)\\ &= -\frac{1}{m}L_3 p_x - k\frac{y}{r},\end{aligned}$$

which is the y-component of the Laplace–Runge–Lenz vector (4.52) and, by virtue of the Poisson theorem, is also a constant of motion.

4.18 Expression (4.57) suggests the definition

$$h_E \equiv \frac{{p_u}^2 + {p_v}^2}{2m} - c(k_1 + k_2)\cosh u - c(k_1 - k_2)\cos v - Ec^2(\cosh^2 u - \cos^2 v)$$

so that we have the identity

$$H - E = \frac{1}{c^2(\cosh^2 u - \cos^2 v)}(h_E - 0),$$

which is of the form (4.53), with

$$f = \frac{1}{c^2(\cosh^2 u - \cos^2 v)}, \qquad g(x) = x.$$

Therefore, on the hypersurface $H = E$ (which amounts to $h_E = 0$), the equations of motion can be expressed as the Hamilton equations for h_E with τ as the evolution parameter, and

$$\mathrm{d}t = c^2(\cosh^2 u - \cos^2 v)\,\mathrm{d}\tau.$$

Noting that h_E is the sum of a function of u and p_u only, and another function of v and p_v only, we conclude that each of these functions is a constant of motion (see Proposition 4.7), that is,

$$\frac{{p_u}^2}{2m} - c(k_1 + k_2)\cosh u - Ec^2\cosh^2 u = -D, \tag{S.22}$$

$$\frac{{p_v}^2}{2m} - c(k_1 - k_2)\cos v + Ec^2\cos^2 v = D, \tag{S.23}$$

where D is a constant of motion. Since

$$\frac{\mathrm{d}u}{\mathrm{d}\tau} = \frac{\partial h_E}{\partial p_u} = \frac{p_u}{m},$$

eliminating p_u from (S.22) we obtain

$$\frac{m}{2}\left(\frac{\mathrm{d}u}{\mathrm{d}\tau}\right)^2 - c(k_1 + k_2)\cosh u - Ec^2\cosh^2 u = -D, \tag{S.24}$$

hence

$$\int \mathrm{d}\tau = \pm \int \sqrt{\frac{m}{2}} \frac{\mathrm{d}u}{\sqrt{Ec^2 \cosh^2 u + c(k_1 + k_2)\cosh u - D}}$$

and, in a similar manner, from (S.23) it follows that

$$\int \mathrm{d}\tau = \pm \int \sqrt{\frac{m}{2}} \frac{\mathrm{d}v}{\sqrt{-Ec^2 \cos^2 v + c(k_1 - k_2)\cos v + D}}.$$

In order to find an expression for D in terms of u, v, p_u, p_v only, we multiply Equations (S.22) and (S.23) by $\cos^2 v$ and $\cosh^2 u$, respectively, and add the resulting equations. In this way we obtain

$$D = \frac{\cos^2 v\, {p_u}^2 + \cosh^2 u\, {p_v}^2}{2m(\cosh^2 u - \cos^2 v)} - \frac{k_1 c\, \cosh u \cos v}{\cosh u - \cos v} + \frac{k_2 c\, \cosh u \cos v}{\cosh u + \cos v}.$$

Making use of the relations [see (4.17)]

$$\begin{aligned} p_u &= p_x\, c\, \sinh u \cos v + p_y\, c\, \cosh u \sin v, \\ p_v &= -p_x\, c\, \cosh u \sin v + p_y\, c\, \sinh u \cos v, \end{aligned} \tag{S.25}$$

we find the equivalent expression

$$D = \frac{c^2 {p_x}^2 + (x p_y - y p_x)^2}{2m} + cx\left(\frac{k_2}{r_2} - \frac{k_1}{r_1}\right).$$

A fairly complete discussion of this problem can be found in Pars [11, Sects. 17.10–17.13].

4.19 Making use of the expressions given in Exercise 4.18 one finds that the standard Lagrangian for a two-dimensional isotropic harmonic oscillator in the confocal coordinates introduced in Exercise 4.18 is [see (2.29)]

$$L = \frac{mc^2}{2}(\cosh^2 u - \cos^2 v)(\dot{u}^2 + \dot{v}^2) - \frac{m\omega^2 c^2}{2}(\cosh^2 u + \cos^2 v - 1).$$

Hence, the Hamiltonian is given by (see Proposition 4.5)

$$H = \frac{{p_u}^2 + {p_v}^2}{2mc^2(\cosh^2 u - \cos^2 v)} + \frac{m\omega^2 c^2}{2}(\cosh^2 u + \cos^2 v - 1),$$

and we find that $H = E$ is equivalent to $h_E = 0$ where

$$h_E \equiv \frac{{p_u}^2 + {p_v}^2}{2m} + \frac{m\omega^2 c^4}{2}(\cosh^4 u - \cos^4 v - \cosh^2 u + \cos^2 v) - c^2 E(\cosh^2 u - \cos^2 v).$$

Since h_E is the sum of a function of u and p_u only plus a function of v and p_v only, it follows that each of these functions is a constant of motion, that is,

$$\frac{{p_u}^2}{2m} + \frac{m\omega^2 c^4}{2}(\cosh^4 u - \cosh^2 u) - c^2 E \cosh^2 u = -D, \tag{S.26}$$

$$\frac{{p_v}^2}{2m} - \frac{m\omega^2 c^4}{2}(\cos^4 v - \cos^2 v) + c^2 E \cos^2 v = D, \tag{S.27}$$

where D is a constant of motion. Combining (S.26) and (S.27) to eliminate E we find

$$D = \frac{\cos^2 v\, {p_u}^2 + \cosh^2 u\, {p_v}^2}{2m(\cosh^2 u - \cos^2 v)} + \frac{1}{2} m\omega^2 c^4 \cosh^2 u \cos^2 v.$$

With the aid of (S.25) we obtain the expression of D in terms of the Cartesian coordinates and their conjugate momenta

$$D = \frac{c^2}{2m}({p_x}^2 + m^2\omega^2 x^2) + \frac{1}{2m}(xp_y - yp_x)^2.$$

Since c is an arbitrary constant, it follows that ${p_x}^2 + m^2\omega^2 x^2$ and $(xp_y - yp_x)^2$ are separately conserved.

Finally, making use of the relation

$$\frac{du}{d\tau} = \frac{\partial h_E}{\partial p_u} = \frac{p_u}{m}$$

in (S.26), it follows that

$$\frac{m}{2}\left(\frac{du}{d\tau}\right)^2 + \frac{m\omega^2 c^4}{2}(\cosh^4 u - \cosh^2 u) - c^2 E \cosh^2 u = -D,$$

which gives a direct relation between u and τ. In a similar way one obtains the relation

$$\frac{m}{2}\left(\frac{dv}{d\tau}\right)^2 - \frac{m\omega^2 c^4}{2}(\cos^4 v - \cos^2 v) + c^2 E \cos^2 v = D,$$

which relates v and τ.

4.20 The equation

$$\frac{1}{2}\frac{P{p_x}^2 + Q{p_y}^2}{X+Y} + \frac{\xi+\eta}{X+Y} = E \tag{S.28}$$

is equivalent to

$$\tfrac{1}{2}P{p_x}^2 + \xi - EX + \tfrac{1}{2}Q{p_y}^2 + \eta - EY = 0.$$

Hence, letting

$$h_E \equiv \tfrac{1}{2}P{p_x}^2 + \xi - EX + \tfrac{1}{2}Q{p_y}^2 + \eta - EY,$$

we have

$$H - E = \frac{1}{X+Y}(h_E - 0)$$

[cf. Equation (4.53)]. The form of the Hamiltonian h_E implies that

$$\tfrac{1}{2}P{p_x}^2 + \xi - EX = -D, \qquad \tfrac{1}{2}Q{p_y}^2 + \eta - EY = D, \tag{S.29}$$

where D is a constant of motion. Eliminating E from these last two equations (or, equivalently, substituting (S.28) in one of them) one obtains

$$D = \frac{-\frac{1}{2}(YP{p_x}^2 - XQ{p_y}^2) + \eta X - \xi Y}{X+Y}.$$

Since h_E is the sum of a function of x and p_x only, and a function of y and p_y only, from Equations (4.54) it follows that the equations of motion for x and p_x in terms of τ are independent of the equations of motion for y and p_y, and vice versa. According to (4.55), the parameter τ is related to the time by

$$\mathrm{d}t = (X+Y)\,\mathrm{d}\tau.$$

The complete solution of the equations of motion can be obtained making use of [see Equations (4.54)]

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \frac{\partial h_E}{\partial p_x} = Pp_x, \qquad \frac{\mathrm{d}y}{\mathrm{d}\tau} = \frac{\partial h_E}{\partial p_y} = Qp_y,$$

which, together with (S.29), lead to

$$\frac{\mathrm{d}x}{\sqrt{2P(EX-\xi-D)}} = \pm\mathrm{d}\tau, \qquad \frac{\mathrm{d}y}{\sqrt{2Q(EY-\eta+D)}} = \pm\mathrm{d}\tau.$$

4.21 In this case the equation $H = E$ amounts to

$$\tfrac{1}{2}(P_1 {p_1}^2 + P_2 {p_2}^2 + \cdots + P_n {p_n}^2) + \xi_1 + \xi_2 + \cdots + \xi_n - E(X_1 + X_2 + \cdots + X_n) = 0,$$

which suggests the definition of the equivalent Hamiltonian $h_E = h_1 + h_2 + \cdots + h_n$, where, for $i = 1, 2, \ldots, n$,

$$h_i \equiv \tfrac{1}{2} P_i {p_i}^2 + \xi_i - E X_i \qquad \text{(without summation on } i\text{)}$$

is a function of q_i and p_i only.

Each h_i is a constant of motion, but they are related by the condition $h_1 + h_2 + \cdots + h_n = 0$. Therefore, $H, h_1, h_2, \ldots, h_{n-1}$, are n functionally independent constants of motion, and the solution of the equations of motion is obtained using the equations

$$\frac{\mathrm{d}q_i}{\mathrm{d}\tau} = \frac{\partial h_E}{\partial p_i} = P_i p_i \qquad \text{(without summation on } i\text{)}$$

and the fact that h_i is conserved. In this way we obtain

$$\frac{\mathrm{d}q_i}{\sqrt{2P_i(EX_i - \xi_i + h_i)}} = \pm \mathrm{d}\tau.$$

(This solution is obtained making use of the Lagrangian formalism in Pars [11, Sect. 26.9] and Greenwood ([9], Sect. 2–3.)

4.22 It is convenient to define

$$h_E \equiv \frac{1}{2m}\left(r^2 {p_r}^2 + {p_\theta}^2 + \frac{{p_\phi}^2}{\sin^2\theta}\right) + k\cos\theta - Er^2,$$

so that $H = E$ is equivalent to $h_E = 0$. In fact,

$$H - E = \frac{1}{r^2}(h_E - 0)$$

and, therefore, the parameter τ is related to the time by $\mathrm{d}\tau = \mathrm{d}t/r^2$.

The Hamiltonian h_E can be written as the sum $h_1 + h_2$, where

$$h_1 \equiv \frac{r^2 {p_r}^2}{2m} - Er^2, \qquad h_2 \equiv \frac{1}{2m}\left({p_\theta}^2 + \frac{{p_\phi}^2}{\sin^2\theta}\right) + k\cos\theta.$$

Since h_2 does not contain r and p_r, it follows that h_1 and h_2 are separately conserved (see Proposition 4.7), but their values are related by the condition $h_1+h_2 = h_E = 0$. Thus

$$M \equiv {p_\theta}^2 + \frac{{p_\phi}^2}{\sin^2\theta} + 2mk\cos\theta$$

is a constant of motion, and

$$r^2{p_r}^2 - 2mEr^2 = -M. \tag{S.30}$$

Taking into account the relation (4.56), it follows that $M = \mathbf{L}^2 + 2mk\cos\theta$, where $\mathbf{L}^2$ is the square of the angular momentum of the particle about the origin. It may be noticed that h_2 has the form of the Hamiltonian of a spherical pendulum (of unit length) in a uniform gravitational field [cf. (2.31)].

The Hamilton equations for r and p_r, as functions of τ, are

$$\frac{\mathrm{d}r}{\mathrm{d}\tau} = \frac{\partial h_E}{\partial p_r} = \frac{r^2 p_r}{m}, \qquad \frac{\mathrm{d}p_r}{\mathrm{d}\tau} = -\frac{\partial h_E}{\partial r} = -\frac{r{p_r}^2}{m} + 2Er.$$

Eliminating p_r from the first of these equations with the aid of (S.30), we find that

$$\frac{\mathrm{d}r}{\mathrm{d}\tau} = \pm\frac{r}{m}\sqrt{2mEr^2 - M}.$$

These equations show that there exist solutions with $r =$ const. if and only if $E = 0$ and $M = 0$. In this particular case, the particle will behave exactly as a spherical pendulum.

4.23 Letting

$$h_E \equiv \frac{1}{2m}\left[{p_u}^2 + {p_v}^2 + \frac{{p_\phi}^2(u^2+v^2)}{u^2v^2}\right] - \frac{\gamma}{2}(u^4 - v^4) - E(u^2+v^2),$$

we find that $H = E$ amounts to $h_E = 2k$. The Hamiltonian h_E is equal to the sum $h_1 + h_2$, where

$$h_1 \equiv \frac{{p_u}^2}{2m} + \frac{{p_\phi}^2}{2mu^2} - \frac{\gamma}{2}u^4 - Eu^2, \qquad h_2 \equiv \frac{{p_v}^2}{2m} + \frac{{p_\phi}^2}{2mv^2} + \frac{\gamma}{2}v^4 - Ev^2,$$

and h_1 and h_2 are separately conserved (note that p_ϕ is a constant of motion). Thus, taking into account that $h_1 + h_2 = 2k$,

$$\frac{{p_u}^2}{2m} + \frac{{p_\phi}^2}{2mu^2} - \frac{\gamma}{2}u^4 - Eu^2 = k - D, \qquad \frac{{p_v}^2}{2m} + \frac{{p_\phi}^2}{2mv^2} + \frac{\gamma}{2}v^4 - Ev^2 = k + D, \tag{S.31}$$

where D is a constant of motion. In order to identify D, we multiply these equations by v^2 and u^2, respectively, and subtract the resulting equations, which yields

$$D = \frac{u^2 {p_v}^2 - v^2 {p_u}^2}{2m(u^2+v^2)} + \frac{{p_\phi}^2}{2m}\frac{u^2-v^2}{u^2v^2} + \frac{\gamma}{2}u^2v^2 - k\frac{u^2-v^2}{u^2+v^2}.$$

Making use of the relations

$$\begin{aligned} p_u &= p_x v\cos\phi + p_y v\sin\phi + p_z u, \\ p_v &= p_x u\cos\phi + p_y u\sin\phi - p_z v, \\ p_\phi &= -p_x uv\sin\phi + p_y uv\cos\phi, \end{aligned}$$

which follow from the definition of the parabolic coordinates (u, v, ϕ) in terms of the Cartesian ones and (4.17), one finds that, in terms of the Cartesian coordinates and their conjugate momenta,

$$\begin{aligned} D &= \frac{1}{m}\left(z{p_x}^2 + z{p_y}^2 - xp_xp_z - yp_yp_z\right) + \tfrac{1}{2}\gamma(x^2+y^2) - \frac{kz}{r} \\ &= \frac{1}{m}\left[p_x(zp_x - xp_z) - p_y(yp_z - zp_y)\right] + \tfrac{1}{2}\gamma(x^2+y^2) - \frac{kz}{r} \\ &= \frac{1}{m}\left[A_3 + \tfrac{1}{2}m\gamma(x^2+y^2)\right], \end{aligned}$$

where A_3 is the z-component of the Laplace–Runge–Lenz vector (4.52).

Making use of the fact that

$$\frac{\mathrm{d}u}{\mathrm{d}\tau} = \frac{\partial h_E}{\partial p_u} = \frac{p_u}{m},$$

from (S.31) it follows that

$$\frac{m}{2}\left(\frac{\mathrm{d}u}{\mathrm{d}\tau}\right)^2 + \frac{{p_\phi}^2}{2mu^2} - \frac{\gamma}{2}u^4 - Eu^2 = k - D$$

and, similarly,

$$\frac{m}{2}\left(\frac{\mathrm{d}v}{\mathrm{d}\tau}\right)^2 + \frac{{p_\phi}^2}{2mv^2} + \frac{\gamma}{2}v^4 - Ev^2 = k + D.$$

4.25 When $E = V_0$ the last equation in Example 4.24 reduces to

$$\frac{\mathrm{d}q}{\mathrm{d}\tau} = \pm\sqrt{\frac{8V_0}{m\alpha^2}(2-q^2)}.$$

With the change of variable $q = \sqrt{2}\,\cos\theta$, we obtain

$$\pm\int \mathrm{d}\tau = \int \frac{\mathrm{d}q}{\sqrt{\frac{8V_0}{m\alpha^2}(2-q^2)}} = -\int \frac{\sqrt{2}\,\sin\theta\,\mathrm{d}\theta}{\sqrt{\frac{16V_0}{m\alpha^2}\sin^2\theta}}$$

hence, choosing the integration constants conveniently

$$q = \sqrt{2}\cos\left(\sqrt{\frac{8V_0}{m\alpha^2}}\,\tau\right).$$

Then, substituting into the relation $\mathrm{d}\tau = (\alpha^2 q^2/4)\,\mathrm{d}t$, found in Example 4.24, we have

$$\int \mathrm{d}t = \int \frac{4\,\mathrm{d}\tau}{\alpha^2 q^2} = \int \frac{2}{\alpha^2}\sec^2\left(\sqrt{\frac{8V_0}{m\alpha^2}}\,\tau\right)\mathrm{d}\tau = \sqrt{\frac{m}{2V_0\alpha^2}}\tan\left(\sqrt{\frac{8V_0}{m\alpha^2}}\,\tau\right)$$
$$= \sqrt{\frac{m}{2V_0\alpha^2}}\sqrt{\frac{2}{q^2}-1}$$

and, therefore

$$q^2 = \frac{2}{1+\dfrac{2V_0\alpha^2(t-t_0)^2}{m}},$$

where t_0 is an integration constant.

4.27 According to the chain rule, we have

$$\frac{\mathrm{d}q_j}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}q_j}{\mathrm{d}t} = \frac{1}{E-V}\dot{q}_j$$

[see (4.55)] and, similarly,

$$\frac{\mathrm{d}^2 q_j}{\mathrm{d}\tau^2} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{1}{E-V}\dot{q}_j\right) = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{E-V}\dot{q}_j\right) = \frac{1}{E-V}\left(\frac{\ddot{q}_j}{E-V} + \frac{\dot{q}_j\dot{q}_k}{(E-V)^2}\frac{\partial V}{\partial q_k}\right).$$

Substituting these expressions together with $g_{ij} = (E-V)a_{ij}$ into the geodesic equations

$$g_{ij}q''_j + \frac{\partial g_{ij}}{\partial q_k}q'_k q'_j - \frac{1}{2}\frac{\partial g_{jk}}{\partial q_i}q'_j q'_k = 0$$

we obtain

$$a_{ij}\ddot{q}_j + \frac{a_{ij}\dot{q}_j\dot{q}_k}{E-V}\frac{\partial V}{\partial q_k} + \dot{q}_j\dot{q}_k\frac{\partial a_{ij}}{\partial q_k} - \frac{a_{ij}\dot{q}_j\dot{q}_k}{E-V}\frac{\partial V}{\partial q_k} - \frac{1}{2}\dot{q}_j\dot{q}_k\frac{\partial a_{jk}}{\partial q_i} + \frac{1}{2}\frac{1}{E-V}a_{jk}\dot{q}_j\dot{q}_k\frac{\partial V}{\partial q_i} = 0. \quad \text{(S.32)}$$

On the other hand, the condition $h_E = 1$ amounts to $L_E = 1$, that is

$$1 = \frac{1}{2}(E-V)\,a_{ij}q_i'q_j' = \frac{1}{2}\frac{1}{E-V}a_{ij}\dot{q}_i\dot{q}_j,$$

which implies that (S.32) reduces to

$$a_{ij}\ddot{q}_j + \dot{q}_j\dot{q}_k\frac{\partial a_{ij}}{\partial q_k} - \frac{1}{2}\dot{q}_j\dot{q}_k\frac{\partial a_{jk}}{\partial q_i} + \frac{\partial V}{\partial q_i} = 0$$

and it can be readily verified that these are the Lagrange equations for the Lagrangian (4.64).

4.28 The result can be obtained by a straightforward computation. Under a passive rotation by an angle θ in the uv-plane, the coordinates u, v transform according to Equations (4.72), hence,

$$\begin{aligned} x' &= \tfrac{1}{2}(u'^2 - v'^2) \\ &= \tfrac{1}{2}\big[(u\cos\theta + v\sin\theta)^2 - (-u\sin\theta + v\cos\theta)^2\big] \\ &= \tfrac{1}{2}\big[(u^2 - v^2)(\cos^2\theta - \sin^2\theta) + 4uv\sin\theta\cos\theta\big] \\ &= x\ \cos 2\theta + y\ \sin 2\theta \end{aligned}$$

and

$$\begin{aligned} y' &= u'v' \\ &= (u\cos\theta + v\sin\theta)(-u\sin\theta + v\cos\theta) \\ &= -(u^2 - v^2)\sin\theta\cos\theta + uv(\cos^2\theta - \sin^2\theta) \\ &= -x\ \sin 2\theta + y\ \cos 2\theta, \end{aligned}$$

which correspond to a rotation by an angle 2θ in the xy-plane.

A shorter proof is given using the fact that the definition of the parabolic coordinates employed here is equivalent to $x + \mathrm{i}y = \frac{1}{2}(u + \mathrm{i}v)^2$ and that the rotation (4.72) amounts to the complex equation $u' + \mathrm{i}v' = \mathrm{e}^{-\mathrm{i}\theta}(u + \mathrm{i}v)$. Hence, $x' + \mathrm{i}y' = \frac{1}{2}(u' + \mathrm{i}v')^2 = \frac{1}{2}[\mathrm{e}^{-\mathrm{i}\theta}(u + \mathrm{i}v)]^2 = \frac{1}{2}\mathrm{e}^{-2\mathrm{i}\theta}(u + \mathrm{i}v)^2 = \mathrm{e}^{-2\mathrm{i}\theta}(x + \mathrm{i}y)$, which corresponds to a rotation through 2θ.

4.29 If $p_u = 0$, Equations (4.67) give $u = u_0$ and, choosing $v_0 = 0$, for convenience, $v = p_v \tau / m$. These expressions substituted into (4.40) give

$$x = \frac{1}{2}\left(u_0{}^2 - \frac{p_v{}^2 \tau^2}{m^2}\right), \qquad y = \frac{u_0 p_v}{m}\tau.$$

On the other hand, (4.68) shows that $2L_3 = p_v u_0$, and using the fact that $4mk = p_u{}^2 + p_v{}^2 = p_v{}^2$ we finally obtain the parametric equations of the orbit

$$x = \frac{L_3{}^2}{2mk} - \frac{2k\tau^2}{m}, \qquad y = \frac{2L_3 \tau}{m}.$$

4.30 By comparing Equations (4.73) and (4.74) we have

$$u^2 + v^2 = \frac{4k}{m}\tau^2 + 2r_{\min}$$

and Equation (4.41) gives $u^2 + v^2 = 2r$. Hence, the parameter τ is related to the distance from the origin to the particle by

$$\tau = \pm\sqrt{\frac{m}{2k}(r - r_{\min})}$$

and, making use of (4.75),

$$t = \tau\left(\frac{4k}{3m}\tau^2 + 2r_{\min}\right) = \pm\frac{2}{3}\sqrt{\frac{m}{2k}}\,(r + r_{\min})\,\sqrt{r - r_{\min}}.$$

4.31 Making use of Equations (4.51) and (4.40) we can rewrite the components of the constant vector **A** [defined by (4.70)] in terms of the Cartesian coordinates of the particle and of their conjugate momenta

$$\begin{aligned}
-\tfrac{1}{4}(p_u{}^2 - p_v{}^2) &= -\tfrac{1}{4}\left[(up_x + vp_y)^2 - (-vp_x + up_y)^2\right] \\
&= -\tfrac{1}{4}\left[(u^2 - v^2)(p_x{}^2 - p_y{}^2) + 4uvp_x p_y\right] \\
&= -\tfrac{1}{2}x(p_x{}^2 - p_y{}^2) - yp_x p_y \\
&= xp_y{}^2 - \tfrac{1}{2}x(p_x{}^2 + p_y{}^2) - yp_x p_y \\
&= p_y(xp_y - yp_x) - mx\frac{p_x{}^2 + p_y{}^2}{2m} \\
&= p_y L_3 - mx\frac{k}{r},
\end{aligned}$$

where we have taken into account that $E = 0$ [see Equation (4.44)]. In a similar manner, we have

$$\begin{aligned}
-\tfrac{1}{2} p_u p_v &= -\tfrac{1}{2}(u p_x + v p_y)(-v p_x + u p_y) \\
&= -\tfrac{1}{2}\left[uv({p_y}^2 - {p_x}^2) + (u^2 - v^2) p_x p_y\right] \\
&= -\tfrac{1}{2}[y({p_y}^2 - {p_x}^2) + 2x p_x p_y] \\
&= y{p_x}^2 - \tfrac{1}{2} y({p_x}^2 + {p_y}^2) - x p_x p_y \\
&= p_x(y p_x - x p_y) - my\frac{{p_x}^2 + {p_y}^2}{2m} \\
&= -p_x L_3 - my\frac{k}{r}.
\end{aligned}$$

On the other hand, a direct calculation shows that

$$\begin{aligned}
\mathbf{p} \times \mathbf{L} - \frac{mk}{r}\mathbf{r} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ p_x & p_y & 0 \\ 0 & 0 & L_3 \end{vmatrix} - \frac{mk}{r}(x\mathbf{i} + y\mathbf{j}) = p_y L_3 \mathbf{i} - p_x L_3 \mathbf{j} - \frac{mk}{r}(x\mathbf{i} + y\mathbf{j}) \\
&= \left(p_y L_3 - \frac{mk}{r}x\right)\mathbf{i} + \left(-p_x L_3 - \frac{mk}{r}y\right)\mathbf{j},
\end{aligned}$$

which reproduces the expressions for the components of **A** obtained above.

4.32 Substituting Equations (4.67) into (4.51) we have

$$p_u = \frac{p_u p_x + p_v p_y}{m}\tau + u_0 p_x + v_0 p_y, \qquad p_v = \frac{-p_v p_x + p_u p_y}{m}\tau - v_0 p_x + u_0 p_y$$

and eliminating the parameter τ from these equations we obtain

$$(p_u - u_0 p_x - v_0 p_y)(-p_v p_x + p_u p_y) = (p_v + v_0 p_x - u_0 p_y)(p_u p_x + p_v p_y),$$

which amounts to

$$(u_0 p_v - v_0 p_u)({p_x}^2 + {p_y}^2) - 2 p_u p_v p_x + ({p_u}^2 - {p_v}^2) p_y = 0.$$

The coefficients in this polynomial in p_x and p_y can be identified with the aid of Equations (4.68) and (4.70), which leads to

$$L_3({p_x}^2 + {p_y}^2) + 2A_2 p_x - 2A_1 p_y = 0.$$

If the angular momentum about the origin, L_3, is different from zero, this equation represents a circle passing through the origin in the $p_x p_y$-plane and, when $L_3 = 0$, this is the equation of a straight line in the $p_x p_y$-plane passing through the origin.

4.33 The Hamiltonian

$$h_E = \frac{{p_u}^2 + {p_v}^2}{2m} - E(u^2 + v^2)$$

leads to the equations of motion

$$\frac{\mathrm{d}u}{\mathrm{d}\tau} = \frac{p_u}{m}, \qquad \frac{\mathrm{d}p_u}{\mathrm{d}\tau} = 2Eu, \qquad \frac{\mathrm{d}v}{\mathrm{d}\tau} = \frac{p_v}{m}, \qquad \frac{\mathrm{d}p_v}{\mathrm{d}\tau} = 2Ev,$$

from which one obtains the decoupled equations

$$\frac{\mathrm{d}^2u}{\mathrm{d}\tau^2} = \omega^2 u, \qquad \frac{\mathrm{d}^2v}{\mathrm{d}\tau^2} = \omega^2 v,$$

where $\omega^2 \equiv 2E/m$. Hence, u and v are linear combinations of $\cosh\omega\tau$ and $\sinh\omega\tau$. The orbit in the uv-plane is a hyperbola centered at the origin (see the repulsive isotropic harmonic oscillator in Section 2.1) and by rotating the axes in the uv-plane, if necessary, we can make $u = \beta\cosh\omega\tau$, $v = \alpha\sinh\omega\tau$, where α and β are real constants, with $|\alpha| > |\beta|$ (see Equation (S.33) below).

Then, according to (4.40), in Cartesian coordinates the orbit is given by

$$x = \tfrac{1}{2}(\beta^2\cosh^2\omega\tau - \alpha^2\sinh^2\omega\tau) = \tfrac{1}{4}\big[\alpha^2 + \beta^2 - (\alpha^2 - \beta^2)\cosh 2\omega\tau\big]$$

and

$$y = \alpha\beta\sinh\omega\tau\cosh\omega\tau = \tfrac{1}{2}\alpha\beta\sinh 2\omega\tau,$$

which are parametric equations of one branch of a hyperbola with semiaxes $a \equiv \frac{1}{4}(\alpha^2 - \beta^2)$, and $|b|$, with $b \equiv \frac{1}{2}\alpha\beta$, and one of its foci at the origin.

Since $p_u = m\,\mathrm{d}u/\mathrm{d}\tau = m\omega\beta\sinh\omega\tau$, and $p_v = m\,\mathrm{d}v/\mathrm{d}\tau = m\omega\alpha\cosh\omega\tau$, the value of h_E is related to the constants α and β through

$$\begin{aligned} h_E &= \frac{1}{2m}\big[(m\omega\beta\sinh\omega\tau)^2 + (m\omega\alpha\cosh\omega\tau)^2\big] - E\big[(\beta\cosh\omega\tau)^2 + (\alpha\sinh\omega\tau)^2\big] \\ &= E(\alpha^2 - \beta^2) = 4Ea. \end{aligned} \tag{S.33}$$

On the other hand, $h_E = 2k$, hence,

$$E = \frac{k}{2a}$$

[cf. Equation (4.84)]. That is, the energy is a function of the major semiaxis only.

Finally, from (4.47) and the expressions above we have

$$\begin{aligned} dt &= \big[(\beta\cosh\omega\tau)^2 + (\alpha\sinh\omega\tau)^2\big]d\tau \\ &= \tfrac{1}{2}\big[\beta^2(1+\cosh 2\omega\tau) + \alpha^2(\cosh 2\omega\tau - 1)\big]d\tau \\ &= 2a(e\cosh 2\omega\tau - 1)\, d\tau, \end{aligned}$$

where e is the eccentricity of the hyperbola. Thus, we obtain Kepler's equation for the hyperbolic orbits

$$\frac{\omega t}{a} = e\sinh\psi - \psi,$$

where $\psi \equiv 2\omega\tau$ [cf. Equation (4.88)].

4.34 As in the case of any central force, the angular momentum about the center of force is a constant of motion, for

$$\frac{d\mathbf{L}}{dt} = \frac{d}{dt}(\mathbf{r}\times\mathbf{p}) = \mathbf{r}\times\frac{d\mathbf{p}}{dt} + \frac{d\mathbf{r}}{dt}\times\mathbf{p} = \mathbf{r}\times\left(-\frac{k\,\mathbf{r}}{r^3}\right) + \dot{\mathbf{r}}\times m\dot{\mathbf{r}} = \mathbf{0}.$$

The Laplace–Runge–Lenz vector (4.76) is also conserved. In fact,

$$\frac{d}{dt}\left(\mathbf{p}\times\mathbf{L} - \frac{mk}{r}\mathbf{r}\right) = \frac{d\mathbf{p}}{dt}\times\mathbf{L} + \frac{mk}{r^2}\dot{r}\,\mathbf{r} - \frac{mk}{r}\dot{\mathbf{r}} = \left(-\frac{k\,\mathbf{r}}{r^3}\right)\times\mathbf{L} + \frac{mk}{r^2}\dot{r}\,\mathbf{r} - \frac{mk}{r}\dot{\mathbf{r}},$$

but, by virtue of the well-known formula for the triple vector product, $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}$,

$$\left(-\frac{k\,\mathbf{r}}{r^3}\right)\times\mathbf{L} = -\frac{k}{r^3}\mathbf{r}\times(\mathbf{r}\times\mathbf{p}) = -\frac{k}{r^3}[(\mathbf{r}\cdot\mathbf{p})\mathbf{r} - r^2\mathbf{p}] = -\frac{mk}{r^3}[(\mathbf{r}\cdot\dot{\mathbf{r}})\mathbf{r} - r^2\dot{\mathbf{r}}]$$

and

$$\mathbf{r}\cdot\dot{\mathbf{r}} = \frac{1}{2}\frac{d}{dt}(\mathbf{r}\cdot\mathbf{r}) = \frac{1}{2}\frac{d}{dt}(r^2) = r\dot{r}.$$

In order to prove that the hodograph is a circle, or part of a circle, we note that the Hamilton vector can also be written as

$$\begin{aligned} \frac{\mathbf{L}\times\mathbf{A}}{\mathbf{L}^2} &= \frac{1}{\mathbf{L}^2}\mathbf{L}\times\left(\mathbf{p}\times\mathbf{L} - \frac{mk}{r}\mathbf{r}\right) \\ &= \frac{1}{\mathbf{L}^2}\left[\mathbf{L}^2\mathbf{p} - (\mathbf{L}\cdot\mathbf{p})\mathbf{L} - \frac{mk}{r}\mathbf{L}\times\mathbf{r}\right] \\ &= \mathbf{p} - \frac{mk}{\mathbf{L}^2 r}\mathbf{L}\times\mathbf{r}. \end{aligned}$$

Hence, using the fact that $\mathbf{L}$ is orthogonal to $\mathbf{r}$,

$$\left|\mathbf{p} - \frac{\mathbf{L} \times \mathbf{A}}{\mathbf{L}^2}\right| = \frac{mk}{\mathbf{L}^2 r}\left|\mathbf{L} \times \mathbf{r}\right| = \frac{mk}{|\mathbf{L}|},$$

which shows that the hodograph is (part of) a circle of radius $R = mk/|\mathbf{L}|$ centered at $\mathbf{L} \times \mathbf{A}/\mathbf{L}^2$.

Using the property of the triple scalar product, $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = \mathbf{c} \cdot \mathbf{a} \times \mathbf{b} = \mathbf{b} \cdot \mathbf{c} \times \mathbf{a}$, we find that the distance from the center of the hodograph to the origin is

$$\begin{aligned}
\left|\frac{\mathbf{L} \times \mathbf{A}}{\mathbf{L}^2}\right| &= \left|\mathbf{p} - \frac{mk}{\mathbf{L}^2 r}\mathbf{L} \times \mathbf{r}\right| \\
&= \sqrt{\mathbf{p}^2 - \frac{2mk}{\mathbf{L}^2 r}\mathbf{p} \cdot \mathbf{L} \times \mathbf{r} + \frac{m^2k^2}{\mathbf{L}^4 r^2}\mathbf{L}^2 r^2} \\
&= \sqrt{\mathbf{p}^2 - \frac{2mk}{\mathbf{L}^2 r}\mathbf{L} \cdot \mathbf{r} \times \mathbf{p} + \frac{m^2k^2}{\mathbf{L}^2}} \\
&= \sqrt{\mathbf{p}^2 - \frac{2mk}{r} + \frac{m^2k^2}{\mathbf{L}^2}} \\
&= \sqrt{2mE + \frac{m^2k^2}{\mathbf{L}^2}} \\
&= \sqrt{2mE + R^2},
\end{aligned}$$

which shows that for $E < 0$, the hodograph encloses the origin; for $E = 0$, the hodograph passes through the origin; if $E > 0$, the hodograph does not enclose the origin. (Only when $E < 0$, the hodograph is a closed curve.)

Exercises of Chapter 5

5.3 From the definition of the Poisson bracket we see that $\{Q, P\} = mq - pt$, which is a constant of motion since

$$\frac{\partial(mq - pt)}{\partial t} + \{mq - pt, \frac{p^2}{2m}\} = -p + m\frac{p}{m} = 0.$$

Then, Equations (5.3) and (5.5) give

$$\frac{\partial K}{\partial q} = 0, \qquad \frac{\partial K}{\partial p} = -\frac{tp^2}{m},$$

respectively. Hence, we can take

$$K = -\frac{tp^3}{3m} = -\frac{tP^3}{3m}.$$

5.5 Making use of the definition of the Poisson bracket, we immediately see that $\{Q, P\} = 1$. Then, from Equations (5.12) we obtain

$$\frac{\partial(K-H)}{\partial q} = -mg, \qquad \frac{\partial(K-H)}{\partial p} = gt,$$

which implies that $K = H - mgq + gtp + f(t)$, where $f(t)$ is a function of t only.

5.6 Making use of the definition of the Poisson bracket, we have

$$\{Q, P\} = (\cos\omega t + \omega t \sin\omega t)\cos\omega t - \frac{1}{m\omega}(\omega t\cos\omega t - \sin\omega t)\, m\omega\sin\omega t = 1,$$

which shows that the coordinate transformation is canonical. Computing the right-hand sides of Equations (5.12) we have

$$\frac{\partial(K-H)}{\partial q} = -m\omega^2 q\cos^2\omega t + \omega p\sin\omega t\cos\omega t$$

and

$$\frac{\partial(K-H)}{\partial p} = \omega q\sin\omega t\cos\omega t - \frac{p}{m}\sin^2\omega t,$$

therefore, up to an additive function of t only,

$$K - H = -\frac{p^2}{2m}\sin^2\omega t - \frac{m\omega^2}{2}q^2\cos^2\omega t + \omega pq\sin\omega t\cos\omega t$$

and, making use of the given expressions, we obtain

$$\begin{aligned}
K &= H - \frac{p^2}{2m}\sin^2\omega t - \frac{m\omega^2}{2}q^2\cos^2\omega t + \omega pq\sin\omega t\cos\omega t \\
&= \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2 - \frac{p^2}{2m}\sin^2\omega t - \frac{m\omega^2}{2}q^2\cos^2\omega t + \omega pq\sin\omega t\cos\omega t \\
&= \frac{p^2}{2m}\cos^2\omega t + \frac{m\omega^2}{2}q^2\sin^2\omega t + \omega pq\sin\omega t\cos\omega t \\
&= \frac{1}{2m}\big(p\cos\omega t + m\omega q\sin\omega t\big)^2 \\
&= \frac{P^2}{2m}.
\end{aligned}$$

5.11 Since we are looking for a type F_2 generating function, we consider from the start the left-hand side of (5.17), and write the resulting expressions in terms of (q, P, t),

$$\begin{aligned} p\mathrm{d}q + Q\mathrm{d}P + (K - H)\mathrm{d}t &= (P - mgt)\mathrm{d}q + \big(q + \tfrac{1}{2}gt^2\big)\mathrm{d}P + (K - H)\mathrm{d}t \\ &= \mathrm{d}\big(Pq - mgtq + \tfrac{1}{2}gt^2 P\big) + mgq\mathrm{d}t - gtP\mathrm{d}t + (K - H)\mathrm{d}t \\ &= \mathrm{d}\big(Pq - mgtq + \tfrac{1}{2}gt^2 P\big) + \big(K - H + mgq - gtP\big)\mathrm{d}t. \end{aligned}$$

This last equation shows that the coordinate transformation is canonical, that K must be given by

$$K = H - mgq + gtP + f(t),$$

where $f(t)$ is a function of t only, and that

$$F_2 = Pq - mgtq + \tfrac{1}{2}gt^2 P + \int^t f(u)\,\mathrm{d}u.$$

The sets (q, Q, t) and (p, P, t) are not functionally independent (because $Q = q + \frac{1}{2}gt^2$ and $P = p + mgt$) and, therefore, this canonical transformation *cannot* be obtained by means of a type F_1, or a type F_4, generating function. However, the set (p, Q, t) is functionally independent and there exists a type F_3 generating function,

$$\begin{aligned} F_3 &= F_2 - pq - PQ \\ &= \tfrac{1}{2}mg^2 t^3 + \int^t f(u)\,\mathrm{d}u - pQ + \tfrac{1}{2}gt^2 p - mgtQ. \end{aligned}$$

5.12 A straightforward computation gives

$$\frac{\partial^2 F_1}{\partial Q \partial q} = \frac{\partial}{\partial Q}(k \ln t - \ln Q) = -\frac{1}{Q},$$

which is different from zero and, therefore, F_1 defines a canonical transformation which is given implicitly by [see Equations (5.16)]

$$p = \left(\frac{\partial F_1}{\partial q}\right)_{Q,t} = k \ln t - \ln Q, \qquad P = -\left(\frac{\partial F_1}{\partial Q}\right)_{q,t} = \frac{q}{Q}.$$

Hence, the coordinate transformation generated by the function F_1 given in this exercise is

$$Q = t^k \mathrm{e}^{-p}, \qquad P = qt^{-k}\mathrm{e}^{p},$$

and the new Hamiltonian is related to the old one by $K - H = kq/t$.

5.15 The canonical transformation found in Example 5.8 is given by $Q = 2p\sin^2\frac{1}{2}q$, $P = \cot\frac{1}{2}q$; hence, (q, P, t) is functionally dependent and therefore this canonical transformation cannot be generated by F_2. By contrast, apart from (q, Q, t), the sets (p, Q, t) and (p, P, t) are functionally independent and this canonical transformation can also be generated by F_3 and F_4.

5.16 The answer is yes and it suffices to give an explicit example of such a canonical transformation. A simple example is provided by the linear transformation

$$Q = \frac{1}{\sqrt{2}}(q + p), \qquad P = \frac{1}{\sqrt{2}}(p - q).$$

One can readily verify that $\{Q, P\} = \frac{1}{2}\{q+p, p-q\} = 1$ and that the sets (q, Q, t), (q, P, t), (p, Q, t), and (p, P, t) are all functionally independent.

5.19 By means of a straightforward computation we see that

$$\begin{aligned} p\mathrm{d}q - H\mathrm{d}t - (P\mathrm{d}Q - K\mathrm{d}t) &= p\mathrm{d}q - t^{1/2}p\left(t^{-1/2}\mathrm{d}q - \tfrac{1}{2}t^{-3/2}q\,\mathrm{d}t\right) + (K - H)\mathrm{d}t \\ &= \left(K - H + \tfrac{1}{2}t^{-1}pq\right)\mathrm{d}t, \end{aligned}$$

which shows that the transformation is canonical and we can take $K = H - \frac{1}{2}t^{-1}pq$, that is,

$$K = \frac{P^2}{2mt} - \frac{k}{Q^2 t} - \frac{PQ}{2t}$$

and, therefore, tK (or, equivalently, $tH - \frac{1}{2}pq$) is a constant of motion (see Exercise 4.8).

On the other hand, H is also a constant of motion and, therefore, we have two functionally independent constants of motion (provided that the value of H is not zero, see below) that can be used to find the solution of the equations of motion without having to solve differential equations. In fact, from the algebraic expressions

$$\frac{p^2}{2m} - \frac{k}{q^2} = E, \qquad \frac{tp^2}{2m} - \frac{kt}{q^2} - \frac{pq}{2} = C,$$

where E and C are constants (the values of H and $tH - \frac{1}{2}pq$, respectively), one finds

$$q^2 = \frac{2}{mE}(C - Et)^2 - \frac{k}{E}.$$

5.20 Substituting the expressions given in this exercise one finds that

$$p\mathrm{d}q - H\mathrm{d}t - (P\mathrm{d}Q - K\mathrm{d}t) = p\mathrm{d}q - t^{-s}p\left(t^s\mathrm{d}q + st^{s-1}q\mathrm{d}t\right) + (K - H)\mathrm{d}t$$
$$= \left(K - H - st^{-1}pq\right)\mathrm{d}t,$$

hence, the coordinate transformation is canonical and the new Hamiltonian can be chosen as

$$K = H + \frac{spq}{t} = \frac{tp^2}{2} - \frac{q^3}{3t^4} + \frac{spq}{t} = \frac{t^{2s+1}P^2}{2} - \frac{Q^3}{3t^{3s+4}} + \frac{sPQ}{t}.$$

The last expression shows that it is convenient to take $s = -1$, so that,

$$K = \frac{P^2}{2t} - \frac{Q^3}{3t} - \frac{PQ}{t},$$

and we conclude that $\Phi \equiv tK$ (or, equivalently, $tH - pq$) is a constant of motion (see Exercise 4.8).

Making use of the Hamilton equations, and expressing P in terms of Q and the conserved quantity Φ, we have

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \frac{\partial K}{\partial P} = \frac{P}{t} - \frac{Q}{t} = \pm\frac{\sqrt{Q^2 + \frac{2}{3}Q^3 + 2\Phi}}{t}.$$

5.24 Letting

$$P \equiv p + mgt,$$

which is a constant of motion, we find the expression for p as a function of (q, P, t),

$$p = P - mgt,$$

which, substituted into Equation (5.25), leads to the generating function

$$F_2(q, P, t) = \int^q (P - mgt)\,\mathrm{d}q' = Pq - mgtq + f(P, t),$$

where $f(P, t)$ is an arbitrary function of P and t.

Then, the new Hamiltonian is given by

$$K = H + \frac{\partial F_2}{\partial t} = \frac{(P - mgt)^2}{2m} + mgq - mgq + \frac{\partial f}{\partial t} = \frac{P^2}{2m} - gtP + \frac{1}{2}mg^2t^2 + \frac{\partial f}{\partial t}.$$

Thus, choosing

$$f = -\frac{tP^2}{2m} + \frac{1}{2}gt^2 P - \frac{1}{6}mg^2t^3,$$

we have $K = 0$. The new coordinate Q is given by

$$Q = \frac{\partial F_2}{\partial P} = q - \frac{tP}{m} + \frac{1}{2}gt^2,$$

which is a constant of motion because $K = 0$. This last expression gives q as a function of t.

5.27 According to the hypotheses, for each positive value of E_0, the level curve $H(q, p) = p^2/2m + V(q) = E_0$ is a simple closed curve in the qp-plane (the phase space). Since H is an even function of p, these closed curves are symmetric under the reflection on the q-axis and the integral (5.29) along the curve $H = E_0$ can be written as

$$P(E_0) = \frac{1}{\pi}\int_{q_-(E_0)}^{q_+(E_0)} \sqrt{2m[E_0 - V(q)]}\,\mathrm{d}q,$$

where $q_-(E)$ and $q_+(E)$ are the points where the value of $V(q)$ coincides with E (see Figure 5.3). If $q_{\min}$ denotes the value of q where $V(q)$ attains its minimum value (which is equal to zero), we have

$$P(E_0) = \frac{\sqrt{2m}}{\pi}\left[\int_{q_-(E_0)}^{q_{\min}} \sqrt{E_0 - V(q)}\,\mathrm{d}q + \int_{q_{\min}}^{q_+(E_0)} \sqrt{E_0 - V(q)}\,\mathrm{d}q\right].$$

In the intervals $[q_-(E_0), q_{\min}]$ and $[q_{\min}, q_+(E_0)]$, the function $V(q)$ is injective and has a unique inverse function, which we will denote as $q_-(V)$ and $q_+(V)$, respectively (see Figure S.2). Hence, using V as the integration variable, we have

$$P(E_0) = \frac{\sqrt{2m}}{\pi}\left[\int_{E_0}^{0} \sqrt{E_0 - V(q)}\,\frac{\mathrm{d}q_-}{\mathrm{d}V}\mathrm{d}V + \int_{0}^{E_0} \sqrt{E_0 - V(q)}\,\frac{\mathrm{d}q_+}{\mathrm{d}V}\mathrm{d}V\right].$$

Integrating by parts and using the fact that $q_-(0) = q_+(0) = q_{\min}$ we obtain

$$\begin{aligned} P(E_0) = \frac{\sqrt{2m}}{\pi}\Bigg[& q_-(V)\sqrt{E_0 - V}\Big|_{E_0}^{0} + \int_{E_0}^{0} \frac{q_-(V)\,\mathrm{d}V}{2\sqrt{E_0 - V}} \\ & + q_+(V)\sqrt{E_0 - V}\Big|_{0}^{E_0} + \int_{0}^{E_0} \frac{q_+(V)\,\mathrm{d}V}{2\sqrt{E_0 - V}}\Bigg]. \\ = \frac{1}{\pi}\sqrt{\frac{m}{2}}\int_0^{E_0} & \frac{[q_+(V) - q_-(V)]}{\sqrt{E_0 - V}}\mathrm{d}V, \end{aligned}$$

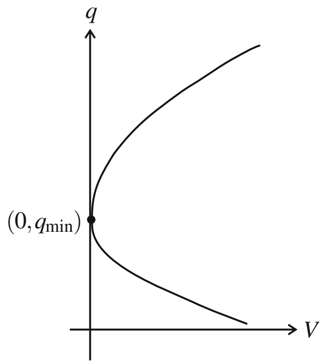


Fig. S.2 The portion of the graph above [resp. below] the point $(0, q_{\min})$ is the graph of the function $q_+(V)$ [resp. $q_-(V)$]

which shows that the function $P(H)$ is determined by the width $w(E) = q_+(E) - q_-(E)$.

5.28 From the definition of the Lagrange brackets (5.37), we have

$$\frac{\partial}{\partial u}[v, w]_{Q,P} + \frac{\partial}{\partial v}[w, u]_{Q,P} + \frac{\partial}{\partial w}[u, v]_{Q,P}$$
$$= \frac{\partial}{\partial u}\left(\frac{\partial Q_j}{\partial v}\frac{\partial P_j}{\partial w} - \frac{\partial P_j}{\partial v}\frac{\partial Q_j}{\partial w}\right) + \frac{\partial}{\partial v}\left(\frac{\partial Q_j}{\partial w}\frac{\partial P_j}{\partial u} - \frac{\partial P_j}{\partial w}\frac{\partial Q_j}{\partial u}\right)$$
$$+ \frac{\partial}{\partial w}\left(\frac{\partial Q_j}{\partial u}\frac{\partial P_j}{\partial v} - \frac{\partial P_j}{\partial u}\frac{\partial Q_j}{\partial v}\right).$$

Making use of the Leibniz rule and assuming that the partial derivatives commute one readily sees that all terms cancel out.

5.32 The set of variables $(q_1, q_2, q_3, P_1, Q_2, P_3)$ is functionally independent since the Jacobian

$$\frac{\partial(q_1, q_2, q_3, P_1, Q_2, P_3)}{\partial(q_1, q_2, q_3, p_1, p_2, p_3)}$$

is different from zero (specifically, its value is $1/\mu$). Alternatively, the functional independence of $(q_1, q_2, q_3, P_1, Q_2, P_3)$ follows from the fact that q_1, q_2, q_3, p_1, p_2, and p_3 can be written in terms of $(q_1, q_2, q_3, P_1, Q_2, P_3)$, viz.,

$$q_1 = q_1, \quad q_2 = q_2, \quad q_3 = q_3, \quad p_1 = P_1 - \frac{\mu}{2}q_2, \quad p_2 = \mu Q_2 - \frac{\mu}{2}q_1, \quad p_3 = P_3.$$

These expressions are also useful in the computation of the left-hand side of the differential form $p_1 \mathrm{d}q_1 + p_2 \mathrm{d}q_2 + p_3 \mathrm{d}q_3 + Q_1 \mathrm{d}P_1 - P_2 \mathrm{d}Q_2 + Q_3 \mathrm{d}P_3 = \mathrm{d}(F_1 + P_1 Q_1 + P_3 Q_3)$, which gives, using again Equations (5.47),

$$\begin{aligned} p_1 \mathrm{d}q_1 + p_2 \mathrm{d}q_2 + p_3 \mathrm{d}q_3 + Q_1 \mathrm{d}P_1 - P_2 \mathrm{d}Q_2 + Q_3 \mathrm{d}P_3 \\ &= \left(P_1 - \frac{\mu}{2} q_2\right) \mathrm{d}q_1 + \left(\mu Q_2 - \frac{\mu}{2} q_1\right) \mathrm{d}q_2 + P_3 \mathrm{d}q_3 + (q_1 - Q_2) \mathrm{d}P_1 \\ &\quad - (P_1 - \mu q_2) \mathrm{d}Q_2 + q_3 \mathrm{d}P_3 \\ &= \mathrm{d}\left(P_1 q_1 - \frac{\mu}{2} q_1 q_2 + \mu q_2 Q_2 - P_1 Q_2 + q_3 P_3\right). \end{aligned}$$

The function inside the last parenthesis is a generating function of the transformation.

On the other hand, a straightforward computation shows that the Jacobians

$$\frac{\partial(q_1, q_2, q_3, Q_1, Q_2, Q_3)}{\partial(q_1, q_2, q_3, p_1, p_2, p_3)}, \qquad \frac{\partial(q_1, q_2, q_3, P_1, P_2, P_3)}{\partial(q_1, q_2, q_3, p_1, p_2, p_3)},$$

$$\frac{\partial(p_1, p_2, p_3, Q_1, Q_2, Q_3)}{\partial(q_1, q_2, q_3, p_1, p_2, p_3)}, \qquad \frac{\partial(p_1, p_2, p_3, P_1, P_2, P_3)}{\partial(q_1, q_2, q_3, p_1, p_2, p_3)},$$

are all equal to zero and, therefore, *there do not exist* generating functions of type F_1, F_2, F_3, or F_4.

5.34 A straightforward computation, making use of the expressions (5.44), yields

$$p_i \mathrm{d}q_i - H \mathrm{d}t - (P_i \mathrm{d}Q_i - K \mathrm{d}t) = [K - H + \omega_r (q_1 p_2 - q_2 p_1)] \, \mathrm{d}t,$$

which shows that the new Hamiltonian has to be given by $K - H = -\omega_r (q_1 p_2 - q_2 p_1) + f(t)$, where $f(t)$ is a function of t only [cf. (5.45)], and $F_1 = \int^t f(u) \, \mathrm{d}u$. The sets (q_i, Q_i, t) and (p_i, P_i, t) are functionally dependent and, therefore, there do not exist generating functions of type F_1 or F_4. However, (q_i, P_i, t) [as well as (p_i, Q_i, t)] is functionally independent, and the type F_2 generating function is

$$\begin{aligned} F_2 &= F_1 + Q_i P_i \\ &= \int^t f(u) \, \mathrm{d}u + (q_1 \cos \omega_r t + q_2 \sin \omega_r t) P_1 + (-q_1 \sin \omega_r t + q_2 \cos \omega_r t) P_2 + q_3 P_3. \end{aligned}$$

5.37 Substituting the expressions for the Q_i in terms of q_i into Equation (5.65) and solving for P_1 and P_2 one readily obtains

$$P_1 = \left(p_1 - \frac{\partial F_1}{\partial q_1}\right) \cos q_2 - \left(p_2 - \frac{\partial F_1}{\partial q_2}\right) \frac{\sin q_2}{q_1},$$

$$P_2 = \left(p_1 - \frac{\partial F_1}{\partial q_1}\right) \sin q_2 + \left(p_2 - \frac{\partial F_1}{\partial q_2}\right) \frac{\cos q_2}{q_1},$$

where F_1 is an arbitrary function of q_1, q_2, and t.

5.46 Making use of the expressions (5.106), (5.107), and the relation $\varepsilon_{\mu\nu}\varepsilon_{\rho\nu} = \delta_{\mu\rho}$ we have

$$\begin{aligned}
[x_\alpha, x_\beta]_{Q,P}\ \{x_\gamma, x_\beta\}_{Q,P} &= \varepsilon_{\mu\nu}\frac{\partial y_\mu}{\partial x_\alpha}\frac{\partial y_\nu}{\partial x_\beta}\varepsilon_{\rho\sigma}\frac{\partial x_\gamma}{\partial y_\rho}\frac{\partial x_\beta}{\partial y_\sigma} \\
&= \varepsilon_{\mu\nu}\varepsilon_{\rho\sigma}\frac{\partial y_\mu}{\partial x_\alpha}\frac{\partial x_\gamma}{\partial y_\rho}\delta_{\nu\sigma} \\
&= \delta_{\mu\rho}\frac{\partial y_\mu}{\partial x_\alpha}\frac{\partial x_\gamma}{\partial y_\rho} \\
&= \delta_{\alpha\gamma}.
\end{aligned}$$

5.47 A simple example is given by the coordinate transformation

$$Q_1 = q_1, \qquad Q_2 = q_2, \qquad P_1 = p_2, \qquad P_2 = -p_1.$$

One can readily verify that the Jacobian of this transformation is equal to 1, but this coordinate transformation is not canonical, e.g., $\{Q_1, P_1\} = \{q_1, p_2\} = 0$.

5.51 Making use of the chain rule and Equations (5.128) we have

$$\begin{aligned}
\frac{\partial}{\partial s}G\big(Q_i(q_j, p_j, t, s), P_i(q_j, p_j, t, s), t\big) &= \frac{\partial G}{\partial Q_i}\frac{\partial Q_i}{\partial s} + \frac{\partial G}{\partial P_i}\frac{\partial P_i}{\partial s} \\
&= \frac{\partial G}{\partial Q_i}\frac{\partial G}{\partial P_i} - \frac{\partial G}{\partial P_i}\frac{\partial G}{\partial Q_i} \\
&= 0,
\end{aligned}$$

which means that $G\big(Q_i(q_j, p_j, t, s), P_i(q_j, p_j, t, s), t\big)$ does not depend on s, and since $Q_i(q_j, p_j, t, 0) = q_i$ and $P_i(q_j, p_j, t, 0) = p_i$, it follows that

$$G\big(Q_i(q_j, p_j, t, s), P_i(q_j, p_j, t, s), t\big) = G(q_i, p_i, t).$$

5.57 Since the transformations (5.143) do not involve the time, $K_s = H + f(t,s)$, where $f(t,s)$ is some function of t and s only. Then, making use of the formulas (5.143), we find that

$$\begin{aligned}
H(Q, P, t) = \frac{P^2}{2m} - F(t)Q = \frac{p^2}{2m} - F(t)(q+s) &= H(q, p, t) - F(t)s \\
= K_s(q, p, t) - f(t, s) - F(t)s.
\end{aligned}$$

Hence, choosing $f(t,s) = -F(t)s$, we see that the given Hamiltonian is invariant under the translations (5.143). Furthermore, from Equations (5.121) and (5.123) we find that, up to an additive trivial constant, the generating function of these transformations is

$$G = p - \int^t F(u)\,\mathrm{d}u.$$

5.58 Since the given transformations do not involve the time, $K_s = H + f(t,s)$, where $f(t,s)$ is some function of t and s only. On the other hand, substituting the expressions for the new coordinates into the Hamiltonian (5.48) we obtain

$$\begin{aligned} H\big(Q_i(q_j,p_j,t),P_i(q_j,p_j,t),t\big) &= \frac{1}{2m}\left\{\left[p_x + \frac{eB_0}{2c}sa_2 + \frac{eB_0}{2c}(y - sa_2)\right]^2 \right. \\ &\quad \left. + \left[p_y - \frac{eB_0}{2c}sa_1 - \frac{eB_0}{2c}(x - sa_1)\right]^2 + {p_z}^2\right\} \\ &= \frac{1}{2m}\left[\left(p_x + \frac{eB_0}{2c}y\right)^2 + \left(p_y - \frac{eB_0}{2c}x\right)^2 + {p_z}^2\right] \\ &= H(q_i,p_i,t). \end{aligned}$$

Thus, choosing $f(t,s) = 0$, we see that the Hamiltonian (5.48) is invariant under the transformations being considered. Making use of Equations (5.121) and (5.123) we find that the generating function of these translations is

$$G = -a_1\left(p_x - \frac{eB_0}{2c}y\right) - a_2\left(p_y + \frac{eB_0}{2c}x\right) - a_3 p_z,$$

which must be a constant of motion. Since the three constants, a_1, a_2, a_3 are arbitrary, the functions accompanying them are separately constants of motion. (The conservation of p_z also follows from the fact that z is an ignorable coordinate in H.)

5.60 Substituting the Cartesian components of the vector potential

$$\mathbf{A} = \frac{gz(y\,\mathbf{i} - x\,\mathbf{j})}{(x^2+y^2)\sqrt{x^2+y^2+z^2}} + \nabla\xi,$$

into the right-hand sides of Equations (5.149) one obtains

$$\begin{aligned} \frac{\partial G}{\partial x} &= -p_y - \frac{e}{c}\frac{\partial}{\partial x}\left(y\frac{\partial \xi}{\partial x} - x\frac{\partial \xi}{\partial y}\right), \\ \frac{\partial G}{\partial y} &= p_x - \frac{e}{c}\frac{\partial}{\partial y}\left(y\frac{\partial \xi}{\partial x} - x\frac{\partial \xi}{\partial y}\right), \end{aligned}$$

$$\frac{\partial G}{\partial z} = -\frac{e}{c}\frac{\partial}{\partial z}\left(y\frac{\partial \xi}{\partial x} - x\frac{\partial \xi}{\partial y}\right).$$

The solution of these equations, together with Equations (5.148) and $\partial G/\partial t = 0$ is, up to an additive trivial constant,

$$G = -xp_y + yp_x - \frac{e}{c}\left(y\frac{\partial \xi}{\partial x} - x\frac{\partial \xi}{\partial y}\right).$$

In the present case, owing to the vector potential chosen, the canonical momentum is related to the kinematical momentum by (see Example 4.3)

$$p_x = m\dot{x} + \frac{e}{c}\left(\frac{gzy}{(x^2+y^2)r} + \frac{\partial \xi}{\partial x}\right),$$
$$p_y = m\dot{y} + \frac{e}{c}\left(\frac{-gzx}{(x^2+y^2)r} + \frac{\partial \xi}{\partial y}\right),$$
$$p_z = m\dot{z} + \frac{e}{c}\frac{\partial \xi}{\partial z},$$

hence,

$$G = -xm\dot{y} + ym\dot{x} + \frac{eg}{c}\frac{z}{r},$$

which does not contain the arbitrary function ξ, and coincides with (5.151).

5.61 Making use of Equations (5.148)–(5.149) we find that the total differential of G is

$$\begin{aligned} \mathrm{d}G = q_2\mathrm{d}p_1 - q_1\mathrm{d}p_2 &+ \left[-p_2 + \frac{e}{c}A_2 - \frac{e}{c}\left(q_2\frac{\partial A_1}{\partial q_1} - q_1\frac{\partial A_1}{\partial q_2}\right)\right]\mathrm{d}q_1 \\ &+ \left[p_1 - \frac{e}{c}A_1 - \frac{e}{c}\left(q_2\frac{\partial A_2}{\partial q_1} - q_1\frac{\partial A_2}{\partial q_2}\right)\right]\mathrm{d}q_2 - \frac{e}{c}\left(q_2\frac{\partial A_3}{\partial q_1} - q_1\frac{\partial A_3}{\partial q_2}\right)\mathrm{d}q_3, \end{aligned}$$

which, by virtue of the relation $\mathbf{B} = \nabla \times \mathbf{A}$, can be rewritten as

$$\begin{aligned} \mathrm{d}G = \mathrm{d}\left[q_2\left(p_1 - \frac{e}{c}A_1\right) - q_1\left(p_2 - \frac{e}{c}A_2\right)\right] &+ \frac{e}{c}\big[(q_1B_1 + q_2B_2)\,\mathrm{d}q_3 \\ &- B_3(q_1\mathrm{d}q_1 + q_2\mathrm{d}q_2)\big], \end{aligned}$$

which is the desired result.

5.62 According to Equations (5.121), if the given transformations are canonical, then the total differential of their generating function must be given by

$$\mathrm{d}G = -\mathrm{d}p_3 + \frac{e}{c}\frac{\partial A_i}{\partial q_3}\mathrm{d}q_i.$$

The right-hand side of this equation is an exact differential if and only if

$$\frac{\partial}{\partial q_j}\frac{\partial A_i}{\partial q_3} = \frac{\partial}{\partial q_i}\frac{\partial A_j}{\partial q_3}$$

$(i, j, = 1, 2, 3)$, which is equivalent to the condition

$$\frac{\partial}{\partial q_3}\left(\frac{\partial A_i}{\partial q_j} - \frac{\partial A_j}{\partial q_i}\right) = 0.$$

Recalling that $\mathbf{B} = \nabla \times \mathbf{A}$, we see that the Cartesian components of $\mathbf{B}$ must be functions of q_1 and q_2 only. Then, considering the gauge-independent combination $p_3 - (e/c)A_3$, we have

$$\begin{aligned}\mathrm{d}G &= -\mathrm{d}\left(p_3 - \frac{e}{c}A_3\right) + \frac{e}{c}\left[\left(\frac{\partial A_1}{\partial q_3} - \frac{\partial A_3}{\partial q_1}\right)\mathrm{d}q_1 + \left(\frac{\partial A_2}{\partial q_3} - \frac{\partial A_3}{\partial q_2}\right)\mathrm{d}q_2\right] \\ &= -\mathrm{d}\left(p_3 - \frac{e}{c}A_3\right) + \frac{e}{c}(B_2\mathrm{d}q_1 - B_1\mathrm{d}q_2).\end{aligned}$$

5.64 The direct computation of $\partial G/\partial t + \{G, H\}$, with $G = mq - pt + \chi(t)$ and $H = p^2/2m + V(q, t)$, gives

$$\begin{aligned}\frac{\partial G}{\partial t} + \{G, H\} &= -p + \frac{\mathrm{d}\chi}{\mathrm{d}t} + \{mq - pt + \chi, p^2/2m + V(q, t)\} \\ &= -p + \frac{\mathrm{d}\chi}{\mathrm{d}t} + p + t\frac{\partial V}{\partial q} \\ &= \frac{\mathrm{d}\chi}{\mathrm{d}t} + t\frac{\partial V}{\partial q},\end{aligned}$$

hence, G is a constant of the motion if and only if

$$\frac{\partial V}{\partial q} = -\frac{1}{t}\frac{\mathrm{d}\chi}{\mathrm{d}t}.$$

The right-hand side of the last equation is a function of t only and, therefore, up to an irrelevant constant term,

$$V(q,t) = -\frac{q}{t}\frac{d\chi}{dt},$$

which is of the form (5.155), with $\chi(t)$ arbitrary.

5.65 We start by calculating $H\big(Q_i(q_j,p_j,t,v), P_i(q_j,p_j,t,v), t\big)$, which gives

$$\begin{aligned} H(Q_i,P_i,t) &= \frac{{P_1}^2}{2m_1} + \frac{{P_2}^2}{2m_2} + V\big(|Q_2 - Q_1|\big) \\ &= \frac{(p_1 - m_1 v)^2}{2m_1} + \frac{(p_2 - m_2 v)^2}{2m_2} + V\big(|(q_2 - vt) - (q_1 - vt)|\big) \\ &= \frac{{p_1}^2}{2m_1} + \frac{{p_2}^2}{2m_2} + V(|q_2 - q_1|) - vp_1 - vp_2 + \frac{m_1}{2}v^2 + \frac{m_2}{2}v^2 \\ &= H(q_i,p_i,t) - vp_1 - vp_2 + \tfrac{1}{2}m_1 v^2 + \tfrac{1}{2}m_2 v^2. \end{aligned}$$

On the other hand, from Equations (5.42) one finds that $K_v - H = -vp_1 - vp_2 + f(t,v)$, where $f(t,v)$ is a function of t and v only. Hence, choosing $f(t,v) = \frac{1}{2}m_1 v^2 + \frac{1}{2}m_2 v^2$, we have $H(Q_i,P_i,t) = K_v(q_i,p_i,t)$ and

$$\left.\frac{\partial K_v}{\partial v}\right|_{v=0} = -p_1 - p_2.$$

Then, from Equations (5.121) and (5.123), we find that, up to an additive trivial constant,

$$G = m_1 q_1 + m_2 q_2 - (p_1 + p_2)t.$$

The value of G is the initial value (at $t = 0$) of the coordinate of the center of mass of the system multiplied by the total mass. The conservation of G means that the center of mass moves at a constant velocity.

5.66 The Hamilton equations corresponding to the Hamiltonian

$$H = \frac{p_1 p_2}{m} + m\omega^2 q_1 q_2$$

are

$$\dot{q}_1 = \frac{p_2}{m}, \qquad \dot{q}_2 = \frac{p_1}{m}, \qquad \dot{p}_1 = -m\omega^2 q_2, \qquad \dot{p}_2 = -m\omega^2 q_1.$$

Eliminating the generalized momenta from these equations we find

$$\ddot{q}_1 = \frac{\dot{p}_2}{m} = -\omega^2 q_1, \qquad \ddot{q}_2 = \frac{\dot{p}_1}{m} = -\omega^2 q_2,$$

which are the equations of motion of a two-dimensional isotropic harmonic oscillator if q_1, q_2 are taken as the Cartesian coordinates of the particle.

Now, we consider the one-parameter family of coordinate transformations given in this exercise:

$$Q_1 = \mathrm{e}^s q_1, \qquad Q_2 = \mathrm{e}^{-s} q_2, \qquad P_1 = \mathrm{e}^{-s} p_1, \qquad P_2 = \mathrm{e}^s p_2.$$

Computing the differential form $p_i \mathrm{d}q_i - H\mathrm{d}t + (P_i \mathrm{d}Q_i - K\mathrm{d}t)$ we find that

$$p_i \mathrm{d}q_i - H\mathrm{d}t + (P_i \mathrm{d}Q_i - K\mathrm{d}t) = (K - H)\mathrm{d}t,$$

which is an exact differential if and only if $K = H + f(t)$, where $f(t)$ is a function of t only. Thus, the coordinate transformations are canonical and the Hamiltonians have to be related as shown. Choosing $f(t) = 0$ we see that the condition (5.137) is satisfied, that is, the Hamiltonian is invariant under the given one-parameter group of canonical transformations. According to Equations (5.121), the generating function of this one-parameter group of canonical transformations is found to be $G = q_1 p_1 - q_2 p_2 = q_1 m\dot{q}_2 - q_2 m\dot{q}_1$, which corresponds to the angular momentum about the origin and is indeed a constant of motion. (Usually, the angular momentum is related to a group of rotations but in this example it generates changes of scale.)

5.68 The equations of motion derived from the Hamiltonian

$$H = \frac{p^2}{2m} - ktq$$

are

$$\dot{q} = \frac{p}{m}, \qquad \dot{p} = kt$$

and, therefore, owing to the definitions $Q = q - ts/m$ and $P = p - s$, we have

$$\dot{Q} = \dot{q} - \frac{s}{m} = \frac{p - s}{m} = \frac{P}{m}, \qquad \dot{P} = \dot{p} = kt,$$

which are of the same form as the equations of motion written in terms of the original variables q, p.

5.70 Making use of Equations (5.128) one finds that G generates the one-parameter group of canonical transformations

$$Q_1 = x - \frac{s}{m\omega_c}\sin\omega_c t, \qquad Q_2 = y - \frac{s}{m\omega_c}\cos\omega_c t,$$
$$P_1 = p_x - \frac{s}{2}\cos\omega_c t, \qquad P_2 = p_y + \frac{s}{2}\sin\omega_c t.$$

With these expressions we can now compute $p_i \mathrm{d}q_i - H\mathrm{d}t - (P_i \mathrm{d}Q_i - K_s \mathrm{d}t)$, and we obtain

$$p_i \mathrm{d}q_i - H\mathrm{d}t - (P_i \mathrm{d}Q_i - K_s \mathrm{d}t) = \mathrm{d}\left[\frac{s}{2}(x\cos\omega_c t - y\sin\omega_c t)\right]$$
$$+\left[K_s - H - \frac{s^2}{2m} + \frac{s\cos\omega_c t}{m}\left(p_x + \frac{eB_0}{2c}y\right) - \frac{s\sin\omega_c t}{m}\left(p_y - \frac{eB_0}{mc}x\right)\right]\mathrm{d}t,$$

which implies that

$$K_s - H = \frac{s^2}{2m} - \frac{s\cos\omega_c t}{m}\left(p_x + \frac{eB_0}{2c}y\right) + \frac{s\sin\omega_c t}{m}\left(p_y - \frac{eB_0}{mc}x\right) + f(t,s),$$

where $f(t,s)$ is some function of t and s only. (Alternatively, the dependence of $K_s - H$ on x, y, p_x, p_y can be obtained from Equations (5.42).)

On the other hand, substituting the expressions for the Q_i and P_i obtained above into (5.48) we obtain

$$H(Q_i(q_j, p_j, t), P_i(q_j, p_j, t), t)$$
$$= H(q_i, p_i, t) + \frac{1}{2m}\left[s^2 - 2s\cos\omega_c t\left(p_x + \frac{eB_0}{2c}y\right) + 2s\sin\omega_c t\left(p_y - \frac{eB_0}{mc}x\right)\right]$$

and, therefore, the invariance condition (5.137) is satisfied if $f(t,s) = 0$.

5.72 Substituting the function

$$G = -\frac{t^2}{2m}({p_x}^2 + {p_y}^2) + t(xp_x + yp_y) - \frac{m}{2}(x^2 + y^2) - \frac{1}{2}gt^3 p_y + \frac{m}{2}gt^2 y - \frac{1}{8}mg^2t^4$$

into Equations (5.128) one obtains the set of ODEs

$$\frac{\mathrm{d}Q_1}{\mathrm{d}s} = -\frac{t^2 P_1}{m} + tQ_1, \qquad \frac{\mathrm{d}P_1}{\mathrm{d}s} = -tP_1 + mQ_1,$$
$$\frac{\mathrm{d}Q_2}{\mathrm{d}s} = -\frac{t^2 P_2}{m} + tQ_2 - \frac{1}{2}gt^3, \quad \frac{\mathrm{d}P_2}{\mathrm{d}s} = -tP_2 + mQ_2 - \frac{1}{2}mgt^2.$$

Combining these equations one finds that the second derivatives with respect to the parameter s of the Q_i and P_i are equal to zero. Hence, they are linear functions of s and imposing the initial conditions we find

$$Q_1 = x - \frac{t^2 p_x}{m}s + txs, \qquad P_1 = p_x - tp_x s + mxs,$$
$$Q_2 = y - \frac{t^2 p_y}{m}s + tys - \frac{1}{2}gt^3 s, \; P_2 = p_y - tp_y s + mys - \frac{1}{2}mgt^2 s.$$

Then, making use of the equations of motion (5.164) we obtain

$$\begin{aligned}
\dot{Q}_1 &= \frac{p_x}{m} - \frac{2tp_x}{m}s + xs + \frac{tp_x}{m}s = \frac{P_1}{m},\\
\dot{Q}_2 &= \frac{p_y}{m} - \frac{2tp_y}{m}s + t^2 gs + \frac{tp_y}{m}s + ys - \frac{3}{2}gt^2 s = \frac{P_2}{m},\\
\dot{P}_1 &= 0 - p_x s + p_x s = 0,\\
\dot{P}_2 &= -mg - p_y s + tmgs + p_y s - mgts = -mg,
\end{aligned}$$

which have the same form as (5.164).

5.73 As shown in Example 5.50, the function

$$G(q, p, t) = qp - \frac{3}{2}gt^2 p - \frac{tp^2}{m} + mgtq - \frac{1}{2}mg^2 t^3$$

generates the one-parameter group of canonical transformations given by

$$Q = -\frac{1}{2}gt^2 + \left(\frac{tp}{m} + gt^2\right)e^{-s} + \left(q - \frac{1}{2}gt^2 - \frac{tp}{m}\right)e^{s}, \quad P = (p+mgt)e^{-s} - mgt,$$

and the new Hamiltonian is

$$\begin{aligned}
K_s &= H + \frac{p^2}{2m}(e^{-2s} - 1) + mgq(e^{s} - 1) + gtp(e^{-2s} - e^{s}) + f(t, s)\\
&= \frac{p^2}{2m} + mgq + \frac{p^2}{2m}(e^{-2s} - 1) + mgq(e^{s} - 1) + gtp(e^{-2s} - e^{s}) + f(t, s)\\
&= \left(\frac{p^2}{2m} + gtp\right)e^{-2s} + (mgq - gtp)e^{s} + f(t, s),
\end{aligned}$$

where $f(t, s)$ is some function of t and s only.

On the other hand,

$$H(Q, P, t) = \frac{P^2}{2m} + mgQ = \frac{(p + mgt)^2 e^{-2s}}{2m} + mg\left(q - \frac{1}{2}gt^2 - \frac{tp}{m}\right)e^{s}.$$

Therefore, the invariance condition $H\big(Q(q, p, t, s), P(q, p, t, s), t\big) = K_s(q, p, t)$ is satisfied if

$$f(t, s) = -\tfrac{1}{2}mg^2t^2(\mathrm{e}^s - \mathrm{e}^{-2s})$$

(cf. Example 5.50). Hence,

$$\begin{aligned}\left.\frac{\partial K_s}{\partial s}\right|_{s=0} &= -2\left(\frac{p^2}{2m} + gtp\right) + mgq - gtp - \frac{3}{2}mg^2t^2 \\ &= -\frac{p^2}{m} + mgq - 3gtp - \frac{3}{2}mg^2t^2,\end{aligned}$$

which coincides with the partial derivative of G with respect to the time, as required by Equation (5.123).

5.75 With the aid of Jacobi's formula

$$\frac{\mathrm{d}\ln|\det A|}{\mathrm{d}t} = \mathrm{tr}\left(A^{-1}\frac{\mathrm{d}A}{\mathrm{d}t}\right),$$

making use of Equation (5.176) and the properties of the trace we have

$$\frac{\mathrm{d}\ln|\det M|}{\mathrm{d}t} = \mathrm{tr}\left[M^{-1}(\Phi M - M\Phi)\right] = \mathrm{tr}\,(M^{-1}\Phi M) - \mathrm{tr}\,\Phi = 0,$$

which shows that $\det M$ is a constant of motion.

5.76 A direct computation shows that if M is given by (5.178), then

$$\det(M - \lambda I) = \Big\{\big([q_1, p_1] - \lambda\big)\big([q_2, p_2] - \lambda\big) - [q_1, p_2][q_2, p_1] - [q_1, q_2][p_1, p_2]\Big\}^2.$$

Therefore, each root of the quadratic polynomial inside the braces is an eigenvalue of M of multiplicity two.

According to the expression above, the eigenvalues of M are the roots of the polynomial

$$\lambda^2 - \big([q_1, p_1] + [q_2, p_2]\big)\lambda + [q_1, p_1][q_2, p_2] - [q_1, p_2][q_2, p_1] - [q_1, q_2][p_1, p_2] = 0.$$

If λ_1 and λ_2 are the roots of this polynomial, then $\mathrm{tr}\,M = 2(\lambda_1 + \lambda_2)$ and $\det M = {\lambda_1}^2{\lambda_2}^2$. Since $\mathrm{tr}\,M$ and $\det M$ are constants of motion, the eigenvalues of M are also constants of motion.

5.77 Making use of (5.114), with the roles of (q_i, p_i, t) and (Q_i, P_i, t) interchanged, we have

$$\{f, g\}_2 = \{x_\alpha, x_\beta\}_2 \frac{\partial f}{\partial x_\alpha} \frac{\partial g}{\partial x_\beta}.$$

On the other hand, using the result of Exercise 5.46 and the definition of the matrix M, we have

$$\delta_{\alpha\gamma} = [x_\alpha, x_\beta]\,\{x_\gamma, x_\beta\}_2 = M_{\mu\alpha}\varepsilon_{\mu\beta}\{x_\gamma, x_\beta\}_2,$$

which means that $\varepsilon_{\mu\beta}\{x_\gamma, x_\beta\}_2 = (M^{-1})_{\gamma\mu}$, where M^{-1} is the inverse of M. Hence, $\{x_\alpha, x_\beta\}_2 = (M^{-1})_{\alpha\gamma}\varepsilon_{\gamma\beta}$ and

$$\{f, g\}_2 = (M^{-1})_{\alpha\gamma}\varepsilon_{\gamma\beta} \frac{\partial f}{\partial x_\alpha} \frac{\partial g}{\partial x_\beta}.$$

5.79 For the Hamiltonian given in this case, the matrix Φ is

$$\Phi = \begin{pmatrix} 0 & 0 & 1/m & 0 \\ 0 & 0 & 0 & 1/m \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and the constant matrices of the form (5.178) that commute with Φ are

$$M = \begin{pmatrix} a & b & 0 & -d \\ b & c & d & 0 \\ 0 & 0 & a & b \\ 0 & 0 & b & c \end{pmatrix}.$$

Since $\det M = (ac - b^2)^2$, the only condition on the constants appearing in M is $ac - b^2 \neq 0$. With $[t, x_\alpha] = 0$, Equations (5.173) imply that, up to a trivial additive constant,

$$K = a\frac{{p_x}^2}{2m} + b\left(\frac{p_x p_y}{m} + mgx\right) + c\left(\frac{{p_y}^2}{2m} + mgy\right) + dmgp_x.$$

In view of the fact that K does not depend on t, it is conserved and since, apart from the condition $ac - b^2 \neq 0$, the constants a, b, c, d are arbitrary, the functions multiplying these constants must be separately conserved. (The example given at the beginning of Section 5.5 corresponds to $b = 1$ and all the other coefficients equal to zero.)

5.80 By inspection one finds that (5.182) has the structure of the matrix (5.178) and by means of a straightforward computation, making use of the equations of motion of the harmonic oscillator, one obtains

$$\frac{\mathrm{d}M}{\mathrm{d}t} = \frac{\lambda}{m}\begin{pmatrix} 0 & p_x p_y - \beta xy & 0 & -xp_y - yp_x \\ -p_x p_y + \beta xy & 0 & xp_y + yp_x & 0 \\ 0 & -\beta(xp_y + yp_x) & 0 & -p_x p_y + \beta xy \\ \beta(xp_y + yp_x) & 0 & p_x p_y - \beta xy & 0 \end{pmatrix},$$

which coincides with the commutator $\Phi M - M\Phi$. In a similar way, one verifies that the entries of (5.182) satisfy the conditions (5.40). For instance,

$$\frac{\partial}{\partial x}[y, p_x] + \frac{\partial}{\partial y}[p_x, x] + \frac{\partial}{\partial p_x}[x, y] = \frac{\partial(\lambda y p_x)}{\partial x} + \frac{\partial(-1)}{\partial y} + \frac{\partial(-\lambda\beta xy)}{\partial p_x} = 0.$$

Finally, with $[t, x_\alpha] = 0$, substituting the expressions for H and M in the right-hand side of (5.173), one finds the terms proportional to λ cancel and, therefore, K can be chosen equal to H.

Exercises of Chapter 6

6.6 The HJ equation corresponding to the Hamiltonian (6.39) is

$$\frac{\left(\frac{\partial S}{\partial u}\right)^2 + \left(\frac{\partial S}{\partial v}\right)^2}{2mc^2(\cosh^2 u - \cos^2 v)} - \frac{(k_1 + k_2)\cosh u + (k_1 - k_2)\cos v}{c\,(\cosh^2 u - \cos^2 v)} + \frac{\partial S}{\partial t} = 0 \qquad \text{(S.34)}$$

and a separable solution of this equation has the form

$$S = A(u) + B(v) - P_1 t, \qquad \text{(S.35)}$$

where A and B are functions of one variable and P_1 is a separation constant. Substituting (S.35) into (S.34) one finds that the functions A and B must satisfy the separated equations

$$\frac{1}{2m}\left(\frac{\mathrm{d}A}{\mathrm{d}u}\right)^2 - c(k_1 + k_2)\cosh u - P_1 c^2 \cosh^2 u = -P_2,$$

$$\frac{1}{2m}\left(\frac{\mathrm{d}B}{\mathrm{d}v}\right)^2 - c(k_1 - k_2)\cos v + P_1 c^2 \cos^2 v = P_2,$$

where P_2 is a second separation constant. Thus,

$$S = \int \sqrt{2m\big[P_1 c^2 \cosh^2 u + c(k_1 + k_2)\cosh u - P_2\big]}\,\mathrm{d}u + \int \sqrt{2m\big[-P_1 c^2 \cos^2 v + c(k_1 - k_2)\cos v + P_2\big]}\,\mathrm{d}v - P_1 t$$

is a complete solution of the HJ equation.

The orbit can be obtained from

$$Q_2 = \frac{\partial S}{\partial P_2} = -\sqrt{\frac{m}{2}}\int \frac{\mathrm{d}u}{\sqrt{P_1 c^2 \cosh^2 u + c(k_1 + k_2)\cosh u - P_2}} + \sqrt{\frac{m}{2}}\int \frac{\mathrm{d}v}{\sqrt{-P_1 c^2 \cos^2 v + c(k_1 - k_2)\cos v + P_2}},$$

using the fact that Q_2 is a constant of motion, while the dependence of the coordinates of the particle on the time is given by

$$Q_1 = \frac{\partial S}{\partial P_1} = \sqrt{\frac{m}{2}}\int \frac{c^2 \cosh^2 u\,\mathrm{d}u}{\sqrt{P_1 c^2 \cosh^2 u + c(k_1 + k_2)\cosh u - P_2}} - \sqrt{\frac{m}{2}}\int \frac{c^2 \cos^2 v\,\mathrm{d}v}{\sqrt{-P_1 c^2 \cos^2 v + c(k_1 - k_2)\cos v + P_2}} - t, \quad \text{(S.36)}$$

using the fact that Q_1 is a constant of motion.

The fact that Q_2 is a constant of motion is equivalent to the ODE

$$\sqrt{\frac{m}{2}}\frac{\mathrm{d}u}{\sqrt{P_1 c^2 \cosh^2 u + c(k_1 + k_2)\cosh u - P_2}} = \sqrt{\frac{m}{2}}\frac{\mathrm{d}v}{\sqrt{-P_1 c^2 \cos^2 v + c(k_1 - k_2)\cos v + P_2}}.$$

Hence, defining the auxiliary variable τ by

$$\mathrm{d}\tau = \sqrt{\frac{m}{2}}\frac{\mathrm{d}u}{\sqrt{P_1 c^2 \cosh^2 u + c(k_1 + k_2)\cosh u - P_2}} = \sqrt{\frac{m}{2}}\frac{\mathrm{d}v}{\sqrt{-P_1 c^2 \cos^2 v + c(k_1 - k_2)\cos v + P_2}},$$

Equation (S.36) amounts to

$$\mathrm{d}t = c^2(\cosh^2 u - \cos^2 v)\,\mathrm{d}\tau$$

(compare with the solution of Exercise 4.18). The last three equations determine u, v, and t as functions of the parameter τ.

6.7 The natural Hamiltonian for a two-dimensional isotropic harmonic oscillator in the elliptic coordinates introduced in Exercise 4.18 is [see (4.58)]

$$H = \frac{{p_u}^2 + {p_v}^2}{2mc^2(\cosh^2 u - \cos^2 v)} + \frac{m\omega^2 c^2}{2}(\cosh^2 u + \cos^2 v - 1),$$

and the corresponding HJ equation is

$$\frac{\left(\dfrac{\partial S}{\partial u}\right)^2 + \left(\dfrac{\partial S}{\partial v}\right)^2}{2mc^2(\cosh^2 u - \cos^2 v)} + \frac{m\omega^2 c^2}{2}(\cosh^2 u + \cos^2 v - 1) + \frac{\partial S}{\partial t} = 0.$$

Taking into account that t does not appear in H, we look for a separable solution $S = A(u) + B(v) - P_1 t$. Substituting this function into the HJ equation we obtain the separated equations

$$\left(\frac{\mathrm{d}A}{\mathrm{d}u}\right)^2 - 2mc^2 P_1 \cosh^2 u + m^2\omega^2 c^4 \cosh^2 u\,(\cosh^2 u - 1) = -P_2,$$

$$\left(\frac{\mathrm{d}B}{\mathrm{d}v}\right)^2 + 2mc^2 P_1 \cos^2 v + m^2\omega^2 c^4 \cos^2 v\,(1 - \cos^2 v) = P_2,$$

where P_2 is a second separation constant. Thus,

$$S = \int \sqrt{2mc^2 P_1 \cosh^2 u - m^2\omega^2 c^4 \cosh^2 u\,(\cosh^2 u - 1) - P_2}\,\mathrm{d}u$$
$$+ \int \sqrt{-2mc^2 P_1 \cos^2 v - m^2\omega^2 c^4 \cos^2 v\,(1 - \cos^2 v) + P_2}\,\mathrm{d}v - P_1 t$$

is a complete solution of the HJ equation.

In order to identify the new momenta we calculate

$$p_u = \frac{\partial S}{\partial u} = \sqrt{2mc^2 P_1 \cosh^2 u - m^2\omega^2 c^4 \cosh^2 u\,(\cosh^2 u - 1) - P_2}$$

and

$$p_v = \frac{\partial S}{\partial v} = \sqrt{-2mc^2 P_1 \cos^2 v - m^2\omega^2 c^4 \cos^2 v\,(1 - \cos^2 v) + P_2}.$$

Combining these expressions we find that $P_1 = H$ and

$$P_2 = \frac{\cos^2 v\, {p_u}^2 + \cosh^2 u\, {p_v}^2}{\cosh^2 u - \cos^2 v} + m^2\omega^2 c^4 \cosh^2 u\, \cos^2 v.$$

With the aid of (S.25) we obtain the equivalent expression in terms of the Cartesian coordinates

$$P_2 = c^2({p_x}^2 + m^2\omega^2 x^2) + (xp_y - yp_x)^2.$$

The equation of the orbit is given by

$$\begin{aligned} Q_2 = \frac{\partial S}{\partial P_2} = &- \int \frac{\mathrm{d}u}{2\sqrt{2mc^2 P_1 \cosh^2 u - m^2\omega^2 c^4 \cosh^2 u\, (\cosh^2 u - 1) - P_2}} \\ &+ \int \frac{\mathrm{d}v}{2\sqrt{-2mc^2 P_1 \cos^2 v - m^2\omega^2 c^4 \cos^2 v\, (1 - \cos^2 v) + P_2}}. \end{aligned}$$

As in the solution of Exercise 6.6 given above, taking into account that Q_2 is a constant of motion, this last equation is equivalent to the existence of an auxiliary parameter, τ, such that

$$\begin{aligned} \mathrm{d}\tau &= \frac{m\,\mathrm{d}u}{\sqrt{2mc^2 P_1 \cosh^2 u - m^2\omega^2 c^4 \cosh^2 u\, (\cosh^2 u - 1) - P_2}} \\ &= \frac{m\,\mathrm{d}v}{\sqrt{-2mc^2 P_1 \cos^2 v - m^2\omega^2 c^4 \cos^2 v\, (1 - \cos^2 v) + P_2}}. \end{aligned} \tag{S.37}$$

On the other hand, the dependence of the coordinates on the time is given by

$$\begin{aligned} Q_1 = \frac{\partial S}{\partial P_1} = &\int \frac{mc^2 \cosh^2 u\, \mathrm{d}u}{\sqrt{2mc^2 P_1 \cosh^2 u - m^2\omega^2 c^4 \cosh^2 u\, (\cosh^2 u - 1) - P_2}} \\ &- \int \frac{mc^2 \cos^2 v\, \mathrm{d}v}{\sqrt{-2mc^2 P_1 \cos^2 v - m^2\omega^2 c^4 \cos^2 v\, (1 - \cos^2 v) + P_2}} - t. \end{aligned}$$

Taking the differential of this equation and using Equations (S.37) one obtains the relation between τ and t, namely

$$\mathrm{d}t = c^2(\cosh^2 u - \cos^2 v)\, \mathrm{d}\tau.$$

6.8 The HJ equation is given by

$$\frac{1}{2t}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{q^6}{6t} + \frac{q}{2t}\frac{\partial S}{\partial q} + \frac{\partial S}{\partial t} = 0.$$

Looking for solutions of the form $S = A(q) + B(t)$, we obtain

$$3\left(\frac{\mathrm{d}A}{\mathrm{d}q}\right)^2 + q^6 + 3q\frac{\mathrm{d}A}{\mathrm{d}q} = -6t\frac{\mathrm{d}B}{\mathrm{d}t},$$

and we conclude that each side of the last equation must be a constant, P, say. Hence,

$$S = -\frac{P}{6}\ln t - \frac{q^2}{4} + \frac{1}{2}\int^q \sqrt{u^2 - \frac{4}{3}u^6 + \frac{4}{3}P}\,\mathrm{d}u$$

and, therefore,

$$p = \frac{\partial S}{\partial q} = -\frac{q}{2} + \frac{1}{2}\sqrt{q^2 - \frac{4}{3}q^6 + \frac{4}{3}P}.$$

From this last expression we find that the constant of motion P is given by

$$P = 3p^2 + q^6 + 3pq$$

[cf. Equation (5.24)]. The dependence of q on the time is determined by

$$Q = \frac{\partial S}{\partial P} = -\frac{1}{6}\ln t + \frac{1}{3}\int^q \frac{\mathrm{d}u}{\sqrt{u^2 - \frac{4}{3}u^6 + \frac{4}{3}P}},$$

using the fact that Q and P are constants of motion (cf. Example 5.17).

6.9 The HJ equation corresponding to the Hamiltonian

$$H = \frac{1}{2}\frac{\mathscr{P}{p_x}^2 + \mathscr{Q}{p_y}^2}{X+Y} + \frac{\xi+\eta}{X+Y},$$

where $\mathscr{P}$, X, ξ are functions of x only, and $\mathscr{Q}$, Y, η are functions of y only, is

$$\frac{1}{2}\frac{\mathscr{P}\left(\frac{\partial S}{\partial x}\right)^2 + \mathscr{Q}\left(\frac{\partial S}{\partial y}\right)^2}{X+Y} + \frac{\xi+\eta}{X+Y} + \frac{\partial S}{\partial t} = 0.$$

A separable solution of this equation must have the form $S = A(x) + B(y) - P_1 t$, where P_1 is a constant and A, B are functions of a single variable. Substituting the proposed expression for S into the HJ equation we obtain

$$\mathscr{P}\left(\frac{\mathrm{d}A}{\mathrm{d}x}\right)^2 + 2\xi - 2P_1 X = -\mathscr{Q}\left(\frac{\mathrm{d}B}{\mathrm{d}y}\right)^2 - 2\eta + 2P_1 Y.$$

Each side of this last equation must be a constant, which we shall denote as $-2P_2$, and in this way we find the separated equations

$$\frac{\mathrm{d}A}{\mathrm{d}x} = \pm\sqrt{\frac{2(P_1X - \xi - P_2)}{\mathscr{P}}}, \qquad \frac{\mathrm{d}B}{\mathrm{d}y} = \pm\sqrt{\frac{2(P_1Y - \eta + P_2)}{\mathscr{Q}}}.$$

Thus

$$S = -P_1 t + \int \sqrt{\frac{2(P_1X - \xi - P_2)}{\mathscr{P}}}\mathrm{d}x + \int \sqrt{\frac{2(P_1Y - \eta + P_2)}{\mathscr{Q}}}\mathrm{d}y$$

is a complete solution of the HJ equation.

The dependence of the coordinates x and y on the time is then given implicitly by the equations

$$Q_1 = -t + \int \frac{X\,\mathrm{d}x}{\sqrt{2\mathscr{P}(P_1X - \xi - P_2)}} + \int \frac{Y\,\mathrm{d}y}{\sqrt{2\mathscr{Q}(P_1Y - \eta + P_2)}} \tag{S.38}$$

and

$$Q_2 = -\int \frac{\mathrm{d}x}{\sqrt{2\mathscr{P}(P_1X - \xi - P_2)}} + \int \frac{\mathrm{d}y}{\sqrt{2\mathscr{Q}(P_1Y - \eta + P_2)}}, \tag{S.39}$$

using the fact that Q_1 and Q_2 must be constants. Hence, Equation (S.39) is equivalent to the ODE

$$\frac{\mathrm{d}x}{\sqrt{2\mathscr{P}(P_1X - \xi - P_2)}} = \frac{\mathrm{d}y}{\sqrt{2\mathscr{Q}(P_1Y - \eta + P_2)}}.$$

Defining the auxiliary parameter τ by

$$\mathrm{d}\tau = \frac{\mathrm{d}x}{\sqrt{2\mathscr{P}(P_1X - \xi - P_2)}},$$

we have

$$\mathrm{d}\tau = \frac{\mathrm{d}y}{\sqrt{2\mathscr{Q}(P_1Y - \eta + P_2)}}.$$

Making use of these last two equations we find that (S.38) is equivalent to

$$\mathrm{d}t = (X + Y)\,\mathrm{d}\tau$$

(compare with the solution of Exercise 4.20).

6.11 Substituting

$$S(x, y, P_1, P_2, t) = P_1 x + P_2 y - \frac{{P_1}^2 + {P_2}^2}{2m} t$$

and

$$\tilde{S}(x, y, \tilde{P}_1, \tilde{P}_2, t) = \frac{m}{2t}\big[(x - \tilde{P}_1)^2 + (y - \tilde{P}_2)^2\big],$$

into Equations (6.44) we obtain

$$0 = \frac{\partial (S - \tilde{S})}{\partial x} = P_1 - \frac{m}{t}(x - \tilde{P}_1), \qquad 0 = \frac{\partial (S - \tilde{S})}{\partial y} = P_2 - \frac{m}{t}(y - \tilde{P}_2),$$

hence,

$$x = \tilde{P}_1 + \frac{t P_1}{m}, \qquad y = \tilde{P}_2 + \frac{t P_2}{m}.$$

Then, eliminating x and y from the difference $S - \tilde{S}$,

$$F = S - \tilde{S} = P_1 \tilde{P}_1 + P_2 \tilde{P}_2.$$

It may be noticed that, as remarked in Section 6.1.1, t automatically disappears when one eliminates the coordinates q_i in $S - \tilde{S}$.

In order to find the Hamiltonian, we compute the partial derivatives of, e.g., $\tilde{S}$ with respect to the coordinates and the time

$$\frac{\partial \tilde{S}}{\partial x} = \frac{m}{t}(x - \tilde{P}_1), \qquad \frac{\partial \tilde{S}}{\partial y} = \frac{m}{t}(y - \tilde{P}_2), \qquad \frac{\partial \tilde{S}}{\partial t} = -\frac{m}{2t^2}\big[(x - \tilde{P}_1)^2 + (y - \tilde{P}_2)^2\big].$$

Taking into account that the HJ equation should not contain $\tilde{S}$ nor the parameters $\tilde{P}_i$, we combine the partial derivatives of $\tilde{S}$ with respect to x, y, and t in such a way that the parameters $\tilde{P}_i$ are eliminated. In this manner, we obtain

$$\frac{1}{2m}\left[\left(\frac{\partial \tilde{S}}{\partial x}\right)^2 + \left(\frac{\partial \tilde{S}}{\partial y}\right)^2\right] + \frac{\partial \tilde{S}}{\partial t} = 0,$$

which is the HJ equation corresponding to the standard Hamiltonian of a free particle.

6.14 The Hamiltonian corresponding to the Lagrangian (1.77) is

$$H = \frac{1}{2m}\left[{p_\rho}^2 + \frac{1}{\rho^2}\left(p_\phi - \frac{eB_0\rho^2}{2c}\right)^2 + {p_z}^2\right]$$

and therefore the HJ equation is given by

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial \rho}\right)^2 + \frac{1}{\rho^2}\left(\frac{\partial S}{\partial \phi} - \frac{eB_0\rho^2}{2c}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2\right] + \frac{\partial S}{\partial t} = 0.$$

Since t, ϕ, and z do not appear explicitly in H, the HJ equation admits separable solutions of the form $S = -P_1 t + P_2\phi + P_3 z + F(\rho)$, where P_1, P_2, and P_3 are separation constants and the function F satisfies the equation

$$\left(\frac{\mathrm{d}F}{\mathrm{d}\rho}\right)^2 + \frac{1}{\rho^2}\left(P_2 - \frac{eB_0\rho^2}{2c}\right)^2 + {P_3}^2 = 2mP_1.$$

Hence,

$$S = -P_1 t + P_2\phi + P_3 z + \int\sqrt{2mP_1 - {P_3}^2 - \left(\frac{P_2}{\rho} - \frac{eB_0\rho}{2c}\right)^2}\,\mathrm{d}\rho,$$

is a complete solution of the HJ equation, which yields

$$p_z = \frac{\partial S}{\partial z} = P_3, \quad p_\phi = \frac{\partial S}{\partial \phi} = P_2, \quad p_\rho = \frac{\partial S}{\partial \rho} = \sqrt{2mP_1 - {P_3}^2 - \left(\frac{P_2}{\rho} - \frac{eB_0\rho}{2c}\right)^2}$$

and, therefore, $P_1 = H$. Thus, the new momenta are the conserved quantities that follow from the fact that the Lagrangian and the Hamiltonian do not depend on t, ϕ, and z.

On the other hand,

$$\begin{aligned}
Q_2 = \frac{\partial S}{\partial P_2} &= \phi + \int \frac{(eB_0\rho^2/2c - P_2)\,\mathrm{d}\rho}{\rho\sqrt{(2mP_1 - {P_3}^2)\rho^2 - (P_2 - eB_0\rho^2/2c)^2}} \\
&= \phi + \int \frac{(eB_0\rho^2/2c - P_2)\,\mathrm{d}\rho}{\rho\sqrt{(2mP_1 - {P_3}^2 + 2eB_0P_2/c)\rho^2 - (P_2 + eB_0\rho^2/2c)^2}}.
\end{aligned}$$

Introducing the variable α by means of

$$\cos\alpha = \frac{P_2 + eB_0\rho^2/2c}{\rho\sqrt{2mP_1 - {P_3}^2 + 2eB_0P_2/c}} \tag{S.40}$$

we have

$$-\sin\alpha\,\mathrm{d}\alpha = \frac{(eB_0\rho^2/2c - P_2)\,\mathrm{d}\rho}{\rho^2\sqrt{2mP_1 - P_3{}^2 + 2eB_0P_2/c}}$$

hence,

$$Q_2 = \phi \mp \alpha.$$

Substituting $\alpha = \pm(\phi - Q_2)$ into (S.40) we obtain

$$\rho^2 + \frac{2cP_2}{eB_0} = \frac{2c}{eB_0}\sqrt{2mP_1 - P_3{}^2 + 2eB_0P_2/c}\;\rho\cos(\phi - Q_2),$$

which is the equation of a circular cylinder of radius

$$R = \left|\frac{c}{eB_0}\right|\sqrt{2mP_1 - P_3{}^2}.$$

6.16 The standard HJ equation corresponding to the Hamiltonian

$$H = \frac{p^2}{2m} - \phi(t)q,$$

is given by

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 - \phi(t)q + \frac{\partial S}{\partial t} = 0.$$

By inspection one finds it convenient to define

$$\tilde{S} \equiv S - q\int^t \phi(u)\,\mathrm{d}u,$$

so that

$$\frac{1}{2m}\left(\frac{\partial\tilde{S}}{\partial q} + \int^t \phi(u)\,\mathrm{d}u\right)^2 + \frac{\partial\tilde{S}}{\partial t} = 0.$$

This last equation does not contain q explicitly and, therefore, admits separable solutions of the form $\tilde{S} = Pq + F(t)$, where P is a constant. Then, one obtains

$$\frac{\mathrm{d}F}{\mathrm{d}t} = -\frac{1}{2m}\left(P + \int^t \phi(u)\,\mathrm{d}u\right)^2$$

and, returning to the original principal function,

$$S = q\int^t \phi(u)\,\mathrm{d}u + Pq - \frac{1}{2m}\int^t \left(P + \int^w \phi(u)\,\mathrm{d}u\right)^2 \mathrm{d}w.$$

6.21 Making use of the given expressions, $P_1 = p_x$ and $P_2 = p_y + mgt$, we have (treating P_1 and P_2 as parameters)

$$\begin{aligned} p_i\mathrm{d}q_i - H\mathrm{d}t &= P_1\mathrm{d}x + (P_2 - mgt)\mathrm{d}y - \left[\frac{P_1{}^2 + (P_2 - mgt)^2}{2m} + mgy\right]\mathrm{d}t \\ &= \mathrm{d}\left[P_1x + P_2y - mgty - \frac{P_1{}^2t}{2m} + \frac{(P_2 - mgt)^3}{6m^2g}\right]. \end{aligned}$$

The expression inside the brackets must be a solution of the HJ equation, which is essentially the R-separable solution found in Example 6.15.

6.22 From the definition of the functions P_1 and P_2 given in Example 5.45 (p. 193) we obtain

$$p_x = P_2 + \frac{eB_0}{2c}y, \qquad p_y = \frac{eB_0}{2c}x + P_1\sec\omega_c t - \left(P_2 + \frac{eB_0}{c}y\right)\tan\omega_c t,$$

hence,

$$\begin{aligned} H = \frac{1}{2m}\Bigg[&\left(P_2 + \frac{eB_0}{c}y\right)^2\sec^2\omega_c t + P_1{}^2\sec^2\omega_c t \\ &- 2P_1\left(P_2 + \frac{eB_0}{c}y\right)\sec\omega_c t\tan\omega_c t\Bigg] \end{aligned}$$

and, with the aid of these expressions, one finds that

$$\begin{aligned} p_x\mathrm{d}x + p_y\mathrm{d}y - H\mathrm{d}t = \mathrm{d}\Bigg[&P_2x + \frac{eB_0}{2c}xy - \frac{P_1{}^2\tan\omega_c t}{2m\omega_c} \\ &- \frac{1}{2m\omega_c}\left(P_2 + \frac{eB_0}{c}y\right)^2\tan\omega_c t + \frac{1}{m\omega_c}P_1\left(P_2 + \frac{eB_0}{c}y\right)\sec\omega_c t\Bigg]. \end{aligned}$$

The expression inside the brackets must be a solution, $S(x, y, P_1, P_2, t)$, of the HJ equation, which is neither separable nor R-separable.

The principal function thus obtained allows us to find a second pair of constants of motion, Q_1, Q_2, given by

$$Q_1 = \frac{\partial S}{\partial P_1} = -\frac{P_1 \tan \omega_c t}{m\omega_c} + \frac{1}{m\omega_c}\left(P_2 + \frac{eB_0}{c}y\right)\sec \omega_c t,$$

$$Q_2 = \frac{\partial S}{\partial P_2} = x - \frac{1}{m\omega_c}\left(P_2 + \frac{eB_0}{c}y\right)\tan \omega_c t + \frac{P_1}{m\omega_c}\sec \omega_c t.$$

Thus, making use of the expressions for P_1 and P_2, we obtain

$$Q_1 = \frac{c}{eB_0}\left[\left(p_x + \frac{eB_0}{2c}y\right)\cos \omega_c t - \left(p_y - \frac{eB_0}{2c}x\right)\sin \omega_c t\right],$$

$$Q_2 = \frac{c}{eB_0}\left(p_y + \frac{eB_0}{2c}x\right),$$

which coincide with the second pair of constants of motion given in Example 5.45.

6.23 From the definitions

$$P_1 \equiv p_1 + p_2, \qquad P_2 \equiv p_1 p_2 - \mathrm{e}^{q_1 - q_2},$$

one verifies that P_1 and P_2 are in involution, and one obtains the local expressions

$$p_1 = \tfrac{1}{2}\left(P_1 + \sqrt{{P_1}^2 - 4P_2 - 4\mathrm{e}^{q_1-q_2}}\right), \qquad p_2 = \tfrac{1}{2}\left(P_1 - \sqrt{{P_1}^2 - 4P_2 - 4\mathrm{e}^{q_1-q_2}}\right)$$

(note that p_1 and p_2 are not defined uniquely by P_1 and P_2) and $H = \frac{1}{2}{P_1}^2 - P_2$. Hence,

$$p_1 \mathrm{d}q_1 + p_2 \mathrm{d}q_2 - H\mathrm{d}t = \mathrm{d}\left[\tfrac{1}{2}P_1(q_1+q_2) - (\tfrac{1}{2}{P_1}^2 - P_2)t + \tfrac{1}{2}\int^{q_1-q_2}\sqrt{{P_1}^2 - 4P_2 - 4\mathrm{e}^u}\,\mathrm{d}u\right].$$

Making use of the generating function thus obtained, we have

$$Q_1 = \frac{\partial S}{\partial P_1} = \frac{1}{2}(q_1+q_2) - P_1 t + \frac{P_1}{2}\int^{q_1-q_2}\frac{\mathrm{d}u}{\sqrt{{P_1}^2 - 4P_2 - 4\mathrm{e}^u}},$$

$$Q_2 = \frac{\partial S}{\partial P_2} = t - \int^{q_1-q_2}\frac{\mathrm{d}u}{\sqrt{{P_1}^2 - 4P_2 - 4\mathrm{e}^u}},$$

and combining these equations we find

$$q_1 + q_2 = P_1(t + Q_2) + 2Q_1, \qquad \int^{q_1-q_2}\frac{\mathrm{d}u}{\sqrt{{P_1}^2 - 4P_2 - 4\mathrm{e}^u}} = t - Q_2.$$

These last expressions implicitly give the remaining part of the solution of the Hamilton equations, taking into account that the Q_i and P_i are constants of motion.

6.25 From

$$P_1 = \frac{{p_x}^2 + {p_y}^2}{2m} + mgy, \qquad P_2 = \frac{p_x p_y}{m} + mgx,$$

we readily obtain

$$x = \frac{P_2}{mg} - \frac{p_x p_y}{m^2 g}, \qquad y = \frac{P_1}{mg} - \frac{{p_x}^2 + {p_y}^2}{2m^2 g}$$

hence,

$$\begin{aligned} -x\,\mathrm{d}p_x - y\,\mathrm{d}p_y - H\mathrm{d}t &= -\left(\frac{P_2}{mg} - \frac{p_x p_y}{m^2 g}\right)\mathrm{d}p_x - \left(\frac{P_1}{mg} - \frac{{p_x}^2 + {p_y}^2}{2m^2 g}\right)\mathrm{d}p_y - P_1\mathrm{d}t \\ &= \mathrm{d}\left(-\frac{P_2 p_x}{mg} + \frac{{p_x}^2 p_y}{2m^2 g} - \frac{P_1 p_y}{mg} + \frac{{p_y}^3}{6m^2 g} - P_1 t\right), \end{aligned}$$

which means that

$$S(p_x, p_y, P_1, P_2, t) = -\frac{P_2 p_x}{mg} + \frac{{p_x}^2 p_y}{2m^2 g} - \frac{P_1 p_y}{mg} + \frac{{p_y}^3}{6m^2 g} - P_1 t$$

is a (type F_4, R-separable) complete solution of the HJ equation.

6.26 Making use of the explicit expressions

$$Q = q - \frac{pt}{m} + \frac{kt^3}{3m}, \qquad H = \frac{p^2}{2m} - ktq,$$

a straightforward computation gives

$$\begin{aligned} \frac{\partial Q}{\partial t} + \{Q, H\} &= \frac{\partial Q}{\partial t} + \frac{\partial Q}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial H}{\partial q} \\ &= -\frac{p}{m} + \frac{kt^2}{m} + \frac{p}{m} - \frac{kt^2}{m} = 0, \end{aligned}$$

which shows that Q is a constant of motion. Since

$$q = Q + \frac{pt}{m} - \frac{kt^3}{3m},$$

we see that, treating Q as a parameter,

$$\begin{aligned}-q\mathrm{d}p - H\mathrm{d}t &= -\left(Q + \frac{pt}{m} - \frac{kt^3}{3m}\right)\mathrm{d}p - \left(\frac{p^2}{2m} - ktQ - \frac{kt^2p}{m} + \frac{k^2t^4}{3m}\right)\mathrm{d}t \\ &= \mathrm{d}\left(-Qp - \frac{tp^2}{2m} + \frac{kt^3p}{3m} + \frac{kt^2Q}{2} - \frac{k^2t^5}{15m}\right).\end{aligned}$$

The expression inside the parenthesis is a (type F_3) generating function, S, that leads to a new Hamiltonian equal to zero. Hence, a second constant of motion is given by

$$P = -\frac{\partial S}{\partial Q} = p - \frac{kt^2}{2}.$$

6.27 From

$$P = p - \int^t \phi(u)\,\mathrm{d}u$$

we have $p = P + \int^t \phi(u)\,\mathrm{d}u$ and, therefore,

$$\begin{aligned}p\mathrm{d}q - H\mathrm{d}t &= \left(P + \int^t \phi(u)\,\mathrm{d}u\right)\mathrm{d}q - \left[\frac{1}{2m}\left(P + \int^t \phi(u)\,\mathrm{d}u\right)^2 - \phi(t)q\right]\mathrm{d}t \\ &= \mathrm{d}\left[Pq + q\int^t \phi(u)\,\mathrm{d}u - \frac{1}{2m}\int^t\left(P + \int^w \phi(u)\,\mathrm{d}u\right)^2\mathrm{d}w\right].\end{aligned}$$

The expression inside the brackets is an R-separable complete solution of the HJ equation.

6.31 Making use of the given expressions, we have

$$\begin{aligned}p\mathrm{d}q - P\mathrm{d}Q + (K-H)\mathrm{d}t &= p\mathrm{d}q - (p + mgt)(\mathrm{d}q + gt\mathrm{d}t) + (K-H)\mathrm{d}t \\ &= -mgt\mathrm{d}q + \left(K - H - gtp - mg^2t^2\right)\mathrm{d}t \\ &= \mathrm{d}(-mgtq) + \left(K - H - gtp - mg^2t^2 + mgq\right)\mathrm{d}t \\ &= \mathrm{d}(-mgtq) - \tfrac{1}{2}mg^2t^2\mathrm{d}t \\ &= \mathrm{d}\left(-mgtq - \tfrac{1}{6}mg^2t^3\right).\end{aligned}$$

Hence, according to (6.78),

$$S = S' - mgtq - \tfrac{1}{6}mg^2t^3$$

$$= \frac{m}{2t}(Q-a)^2 - mgtq - \frac{1}{6}mg^2t^3$$
$$= \frac{m}{2t}\left(q + \frac{1}{2}gt^2 - a\right)^2 - mgtq - \frac{1}{6}mg^2t^3$$
$$= \frac{m}{2t}(q-a)^2 - \frac{m}{2}gt(q+a) - \frac{m}{24}g^2t^3.$$

Note that, in order to identify the function F_1, the Hamiltonians H and K have to be specified in advance.

6.32 From Example 5.7 we have that for the canonical transformation under consideration,

$$F_1 = \frac{m\omega}{\omega t \cos\omega t - \sin\omega t}\big[qQ - \tfrac{1}{2}Q^2\cos\omega t - \tfrac{1}{2}q^2(\cos\omega t + \omega t\sin\omega t)\big] + f(t),$$

where $f(t)$ is a function of t only, which has to be chosen with the aid of the relation

$$K - H = -\frac{m\omega^2}{2}\left(\frac{Q\sin\omega t - q\omega t}{\omega t\cos\omega t - \sin\omega t}\right)^2 + f'(t).$$

Expressing the difference $K - H$ in terms of (q, Q, t) one finds that $f'(t) = 0$ and we take $f = 0$. Then, from Equation (6.78) we have

$$S = S' + F_1 = \frac{m}{2t}(Q-a)^2 + \frac{m\omega\big[2qQ - Q^2\cos\omega t - q^2(\cos\omega t + \omega t\sin\omega t)\big]}{2(\omega t\cos\omega t - \sin\omega t)}.$$

In order to eliminate Q from this last expression we calculate its partial derivative with respect to Q and equate it to zero. This gives

$$Q = \frac{q\omega t - a(\omega t\cos\omega t - \sin\omega t)}{\sin\omega t}.$$

Substituting this expression into that for S, a somewhat lengthy computation gives

$$S = \frac{m\omega[(q^2+a^2)\cos\omega t - 2aq]}{2\sin\omega t}.$$

(Cf. Example 6.40.)

6.33 Making use of the results of Example 5.14, we have [see Equation (5.21)]

$$p\,\mathrm{d}q - H\,\mathrm{d}t - (P\,\mathrm{d}Q - K\,\mathrm{d}t) = \mathrm{d}\big(mvq - \tfrac{1}{2}mv^2t\big),$$

hence, from Equation (6.78) we obtain,

$$S' = S - F_1 = S - mvq + \tfrac{1}{2}mv^2 t$$

that is, eliminating q,

$$S'(Q, t) = S(Q + vt, t) - mvQ - \tfrac{1}{2}mv^2 t.$$

6.34 Substitution of the expressions $Q = q - s$, $P = p$, $K = H$, into the left-hand side of Equation (5.15) gives

$$p\mathrm{d}q - H\mathrm{d}t - (P\mathrm{d}Q - K\mathrm{d}t) = 0$$

hence, we can take $F_1 = 0$ and, therefore, $S' = S$, i.e., $S'(Q, t) = S(Q + s, t)$ [cf. Equation (6.82)].

In a similar manner, substituting $Q = q$, $P = p - s$, $K = H$, into the left-hand side of (5.15) we obtain

$$p\mathrm{d}q - H\mathrm{d}t - (P\mathrm{d}Q - K\mathrm{d}t) = s\mathrm{d}q.$$

Taking $F_1 = sq$, we have $S' = S - sq = S - sQ$, that is, $S'(Q, t) = S(Q, t) - sQ$.

6.35 Substituting the expressions $Q_i = q_i$, $P_i = p_i + \partial(e\xi/c)/\partial q_i$, and $K = H - \partial(e\xi/c)/\partial t$ into $p_i\mathrm{d}q_i - H\mathrm{d}t - (P_i\mathrm{d}Q_i - K\mathrm{d}t)$ we obtain

$$p_i\mathrm{d}q_i - H\mathrm{d}t - (P_i\mathrm{d}Q_i - K\mathrm{d}t) = (p_i - P_i)\mathrm{d}q_i + (K - H)\mathrm{d}t = -\mathrm{d}(e\xi/c),$$

which shows that F_1 can be taken as $-e\xi/c$ and then Equation (6.78) gives $S' = S + e\xi/c$.

6.37 In this case the Hamilton equations give

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial H}{\partial p} = \frac{p}{m}, \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial q} = kt.$$

The solution of the last equation is $p = \frac{1}{2}kt^2 + P$, where P is an integration constant that represents the value of p at $t = 0$. Then, from the first equation we obtain

$$q = \frac{kt^3}{6m} + \frac{Pt}{m} + Q,$$

where Q is the value of q at $t = 0$. Hence,

$$Q = q - \frac{tp}{m} + \frac{kt^3}{3m}, \qquad P = p - \frac{kt^2}{2},$$

and

$$p\mathrm{d}q - H\mathrm{d}t - P\mathrm{d}Q = \mathrm{d}\left(\frac{kt^2q}{2} + \frac{tp^2}{2m} - \frac{kt^3p}{2m} + \frac{k^2t^5}{10m}\right),$$

which allows us to identify the function F_1 appearing in Equation (6.78).

The initial condition is $S'(Q, 0) = aQ$, hence $P = a$, and, eliminating Q and p, we obtain

$$\begin{aligned} S = S' + F_1 &= aQ + \frac{kt^2q}{2} + \frac{tp^2}{2m} - \frac{kt^3p}{2m} + \frac{k^2t^5}{10m} \\ &= a\left[q - \frac{t}{m}\left(a + \frac{kt^2}{2}\right) + \frac{kt^3}{3m}\right] + \frac{kt^2q}{2} + \frac{t}{2m}\left(a + \frac{kt^2}{2}\right)^2 \\ &\quad - \frac{kt^3}{2m}\left(a + \frac{kt^2}{2}\right) + \frac{k^2t^5}{10m} \\ &= aq - \frac{a^2t}{2m} - \frac{akt^3}{6m} + \frac{kt^2q}{2} - \frac{k^2t^5}{40m}, \end{aligned}$$

which coincides with the R-separable solution (6.53).

6.39 Solving directly the Hamilton equations one obtains

$$Q = q\cos\omega t - \frac{p}{m\omega}\sin\omega t, \qquad P = m\omega q\sin\omega t + p\cos\omega t,$$

where Q and P are the values of q and p at $t = 0$, respectively. Then, a straightforward computation gives

$$p\mathrm{d}q - H\mathrm{d}t - P\mathrm{d}Q = \mathrm{d}\left[pq\sin^2\omega t + \left(\frac{p^2}{2m\omega} - \frac{m\omega q^2}{2}\right)\sin\omega t\cos\omega t\right].$$

The initial condition is $S'(Q, 0) = \alpha Q$, hence $P = \alpha$, and, eliminating Q and p with the aid of the relations

$$Q = -\frac{P}{m\omega}\tan\omega t + q\sec\omega t, \qquad p = -m\omega q\tan\omega t + P\sec\omega t,$$

which follow from the equations above, we obtain

$$\begin{aligned} S = S' + F_1 &= \alpha Q + pq\sin^2\omega t + \left(\frac{p^2}{2m\omega} - \frac{m\omega q^2}{2}\right)\sin\omega t\cos\omega t \\ &= \alpha\left(-\frac{\alpha}{m\omega}\tan\omega t + q\sec\omega t\right) + (-m\omega q\tan\omega t + \alpha\sec\omega t)q\sin^2\omega t \end{aligned}$$

$$+\frac{\sin\omega t\cos\omega t}{2m\omega}\Big[(-m\omega q\tan\omega t+\alpha\sec\omega t)^2-m^2\omega^2q^2\Big]$$
$$=-\left(\frac{\alpha^2}{2m}+\frac{m\omega^2}{2}q^2\right)\frac{\tan\omega t}{\omega}+\alpha q\sec\omega t.$$

6.41 The coordinate transformation given in this exercise amounts to

$$q=q'\mathrm{e}^{-s}+\tfrac{1}{2}gt'^2(\mathrm{e}^{-s}-\mathrm{e}^{-4s}),\qquad t=t'\mathrm{e}^{-2s}.$$

Making use of Equations (6.96) and (6.95) we have

$$p'=\mathrm{e}^{-s}p-\frac{\partial F_1}{\partial q'}$$

and

$$H'=\mathrm{e}^{-2s}H-gt'(\mathrm{e}^{-s}-\mathrm{e}^{-4s})p+\frac{\partial F_1}{\partial t'}.$$

Hence, combining these expressions and the one given for the Hamiltonian, we obtain

$$\begin{aligned}H'&=\mathrm{e}^{-2s}H-gt'(\mathrm{e}^{-s}-\mathrm{e}^{-4s})\,\mathrm{e}^{s}\left(p'+\frac{\partial F_1}{\partial q'}\right)+\frac{\partial F_1}{\partial t'}\\&=\mathrm{e}^{-2s}\left(\frac{p^2}{2m}+mgq\right)-gt'(1-\mathrm{e}^{-3s})\left(p'+\frac{\partial F_1}{\partial q'}\right)+\frac{\partial F_1}{\partial t'}\\&=\frac{1}{2m}\left(p'+\frac{\partial F_1}{\partial q'}\right)^2+mg\mathrm{e}^{-2s}q-gt'(1-\mathrm{e}^{-3s})\left(p'+\frac{\partial F_1}{\partial q'}\right)+\frac{\partial F_1}{\partial t'}\\&=\frac{1}{2m}\left[p'^2+2\frac{\partial F_1}{\partial q'}p'+\left(\frac{\partial F_1}{\partial q'}\right)^2\right]+mg\mathrm{e}^{-2s}\left[q'\mathrm{e}^{-s}+\frac{1}{2}gt'^2(\mathrm{e}^{-s}-\mathrm{e}^{-4s})\right]\\&\quad-gt'(1-\mathrm{e}^{-3s})\left(p'+\frac{\partial F_1}{\partial q'}\right)+\frac{\partial F_1}{\partial t'}.\end{aligned}$$

The terms linear in p' are eliminated choosing F_1 in such a way that $\partial F_1/\partial q'=mgt'(1-\mathrm{e}^{-3s})$, that is,

$$F_1=mgt'q'(1-\mathrm{e}^{-3s})+f(t'),$$

where $f(t')$ is a function of t' only. Thus,

$$H'=\frac{p'^2}{2m}+mgq'-\frac{m}{2}g^2t'^2(3\mathrm{e}^{-3s}-1-2\mathrm{e}^{-6s})+\frac{\mathrm{d}f}{\mathrm{d}t'}$$

and choosing $f = -\frac{1}{6}mg^2t'^3(3\mathrm{e}^{-3s} - 1 - 2\mathrm{e}^{-6s})$, i.e.,

$$F_1 = mgt'q'(1 - \mathrm{e}^{-3s}) + \tfrac{1}{6}mg^2t'^3(1 - 3\mathrm{e}^{-3s} + 2\mathrm{e}^{-6s}),$$

we get

$$H' = \frac{p'^2}{2m} + mgq',$$

which has the same form as the original Hamiltonian.

6.42 Substituting the Lagrangian

$$L(x, y, x', y', z) = n(x, y, z)\sqrt{1 + x'^2 + y'^2}$$

into the Euler–Lagrange equations we obtain, for instance,

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}z}\frac{\partial L}{\partial x'} - \frac{\partial L}{\partial x} \\ &= \frac{\mathrm{d}}{\mathrm{d}z}\frac{nx'}{\sqrt{1 + x'^2 + y'^2}} - \frac{\partial n}{\partial x}\sqrt{1 + x'^2 + y'^2}, \end{aligned}$$

hence, by virtue of the relation $\mathrm{d}s = \sqrt{1 + x'^2 + y'^2}\,\mathrm{d}z$, we have

$$\frac{\partial n}{\partial x} = \frac{1}{\sqrt{1 + x'^2 + y'^2}}\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{n}{\sqrt{1 + x'^2 + y'^2}}\frac{\mathrm{d}x}{\mathrm{d}z}\right) = \frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}x}{\mathrm{d}s}\right),$$

with a similar formula for $\partial n/\partial y$.

Now, with the aid of (1.92), assuming that the Euler–Lagrange equations hold,

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(x'\frac{\partial L}{\partial x'} + y'\frac{\partial L}{\partial y'} - L\right) = -\frac{\partial L}{\partial z},$$

which amounts to

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}z}\left(x'\frac{nx'}{\sqrt{1 + x'^2 + y'^2}} + y'\frac{ny'}{\sqrt{1 + x'^2 + y'^2}} - n\sqrt{1 + x'^2 + y'^2}\right) \\ &\quad = -\frac{\partial n}{\partial z}\sqrt{1 + x'^2 + y'^2}, \end{aligned}$$

that is,

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}z}{\mathrm{d}s}\right)=\frac{\partial n}{\partial z}.$$

6.43 Making use of the Leibniz rule and the ray equation, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\mathbf{r}\times n\,\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)=\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\times n\,\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}+\mathbf{r}\times\frac{\mathrm{d}}{\mathrm{d}s}\left(n\,\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)=\mathbf{r}\times\nabla n,$$

which is equal to zero since, with n being a function of r only, ∇n is proportional to $\mathbf{r}$. Taking into account that $\mathbf{r}$ is orthogonal to the constant vector $\mathbf{L}\equiv\mathbf{r}\times n\,\mathrm{d}\mathbf{r}/\mathrm{d}s$, the vector $\mathbf{r}$ lies on a plane passing through the origin, and therefore each ray lies on a plane passing through the origin.

In order to prove that the vector

$$\mathbf{A}\equiv\mathbf{r}\times\mathbf{L}+\frac{a}{2}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s},$$

where $\mathbf{L}\equiv\mathbf{r}\times n\,\mathrm{d}\mathbf{r}/\mathrm{d}s$ is also constant along each ray we make use of the fact that $\mathbf{L}$ is conserved, the ray equation, the vector identity $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})=(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{a}\cdot\mathbf{b})\mathbf{c}$, the chain rule, the fact that the norm of $\mathrm{d}\mathbf{r}/\mathrm{d}s$ is equal to 1, the fact that the refractive index (6.103) satisfies the equation

$$\frac{\mathrm{d}n}{\mathrm{d}r}=-\frac{2ar}{(b+r^2)^2}=-\frac{2n^2r}{a},$$

and the identity $\mathbf{r}\cdot\mathrm{d}\mathbf{r}/\mathrm{d}\mathbf{s}=\frac{1}{2}\mathrm{d}(\mathbf{r}\cdot\mathbf{r})/\mathrm{d}\mathbf{s}=\frac{1}{2}\mathrm{d}(r^2)/\mathrm{d}s=r\,\mathrm{d}r/\mathrm{d}s$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\left(\mathbf{r}\times\mathbf{L}+\frac{a}{2}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)&=\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\times\mathbf{L}+\frac{a}{2}\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}s^2}\\&=\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\times\left(\mathbf{r}\times n\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)+\frac{a}{2n}\left(\nabla n-\frac{\mathrm{d}n}{\mathrm{d}s}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)\\&=n\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\cdot\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\,\mathbf{r}-\left(\mathbf{r}\cdot\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)n\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}+\frac{a}{2n}\frac{\mathrm{d}n}{\mathrm{d}r}\left(\nabla r-\frac{\mathrm{d}r}{\mathrm{d}s}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)\\&=n\,\mathbf{r}-\left(r\frac{\mathrm{d}r}{\mathrm{d}s}\right)n\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}-n\,r\left(\frac{\mathbf{r}}{r}-\frac{\mathrm{d}r}{\mathrm{d}s}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)\\&=0.\end{aligned}$$

Now we shall prove that the ray is (an arc of) a circle centered at $\mathbf{L}\times\mathbf{A}/\mathbf{L}^2$, where $\mathbf{A}\equiv\mathbf{r}\times\mathbf{L}+(a/2)\,\mathrm{d}\mathbf{r}/\mathrm{d}s$ [cf. Equation (4.78)]. To this end, we calculate the norm of the vector $\mathbf{r}-\mathbf{L}\times\mathbf{A}/\mathbf{L}^2$. As a first step, we note that (cf. the solution of Exercise 4.34)

$$\frac{\mathbf{L}\times\mathbf{A}}{\mathbf{L}^2} = \frac{1}{\mathbf{L}^2}\mathbf{L}\times\left(\mathbf{r}\times\mathbf{L}+\frac{a}{2}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)$$
$$= \frac{1}{\mathbf{L}^2}\left(\mathbf{L}^2\mathbf{r}-(\mathbf{L}\cdot\mathbf{r})\mathbf{L}+\frac{a}{2}\mathbf{L}\times\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)$$
$$= \mathbf{r}+\frac{a}{2\mathbf{L}^2}\mathbf{L}\times\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}.$$

Hence,

$$\left|\mathbf{r}-\frac{\mathbf{L}\times\mathbf{A}}{\mathbf{L}^2}\right| = \frac{a}{2\mathbf{L}^2}\left|\mathbf{L}\times\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right| = \frac{a}{2\mathbf{L}^2}|\mathbf{L}| = \frac{a}{2|\mathbf{L}|},$$

which shows that the ray is (part of) a circle of radius $R = a/2|\mathbf{L}|$.

The distance from the origin to the center of the circle is

$$\left|\frac{\mathbf{L}\times\mathbf{A}}{\mathbf{L}^2}\right| = \left|\mathbf{r}+\frac{a}{2\mathbf{L}^2}\mathbf{L}\times\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right|$$
$$= \sqrt{\mathbf{r}^2+\frac{a}{\mathbf{L}^2}\mathbf{r}\cdot\mathbf{L}\times\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}+\frac{a^2}{4\mathbf{L}^4}\mathbf{L}^2}$$
$$= \sqrt{\mathbf{r}^2+\frac{a}{\mathbf{L}^2}\mathbf{L}\cdot\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\times\mathbf{r}+\frac{a^2}{4\mathbf{L}^2}}$$
$$= \sqrt{\mathbf{r}^2-\frac{a}{\mathbf{L}^2}\frac{\mathbf{L}^2}{n}+\frac{a^2}{4\mathbf{L}^2}}$$
$$= \sqrt{\mathbf{r}^2-(b+r^2)+\frac{a^2}{4\mathbf{L}^2}}$$
$$= \sqrt{-b+R^2},$$

which shows that for $b > 0$, the circle encloses the origin; for $b = 0$, the circle passes through the origin; if $b < 0$, the circle does not enclose the origin.

6.44 Starting from the Lagrangian

$$L(\theta,\phi,\theta',\phi',r) = n(r,\theta,\phi)\sqrt{1+r^2\theta'^2+r^2\sin^2\theta\,\phi'^2},$$

where $\theta' \equiv \mathrm{d}\theta/\mathrm{d}r$, $\phi' \equiv \mathrm{d}\phi/\mathrm{d}r$, we obtain the canonical momenta

$$p_\theta = \frac{\partial L}{\partial\theta'} = \frac{n\,r^2\theta'}{\mathrm{d}s/\mathrm{d}r}, \qquad p_\phi = \frac{\partial L}{\partial\phi'} = \frac{n\,r^2\sin^2\theta\,\phi'}{\mathrm{d}s/\mathrm{d}r},$$

where we have made use of the abbreviation

$$\frac{\mathrm{d}s}{\mathrm{d}r} \equiv \sqrt{1 + r^2\theta'^2 + r^2 \sin^2\theta\, \phi'^2}.$$

Hence,

$$\begin{aligned} H &= p_\theta \theta' + p_\phi \phi' - L \\ &= \frac{n\, r^2\theta'^2 + n\, r^2 \sin^2\theta\, \phi'^2 - n\big(1 + r^2\theta'^2 + r^2 \sin^2\theta\, \phi'^2\big)}{\mathrm{d}s/\mathrm{d}r} \\ &= -\frac{n}{\mathrm{d}s/\mathrm{d}r}. \end{aligned}$$

On the other hand,

$$\left(\frac{\mathrm{d}s}{\mathrm{d}r}\right)^2 = 1 + r^2\theta'^2 + r^2 \sin^2\theta\, \phi'^2 = 1 + \left(\frac{\mathrm{d}s}{\mathrm{d}r}\frac{p_\theta}{n\, r}\right)^2 + \left(\frac{\mathrm{d}s}{\mathrm{d}r}\frac{p_\phi}{n\, r \sin\theta}\right)^2,$$

that is,

$$\left(\frac{\mathrm{d}s}{\mathrm{d}r}\right)^2 \left(1 - \frac{{p_\theta}^2}{n^2\, r^2} - \frac{{p_\phi}^2}{n^2\, r^2 \sin^2\theta}\right) = 1$$

and, therefore,

$$H = \pm\sqrt{n^2 - \frac{{p_\theta}^2}{r^2} - \frac{{p_\phi}^2}{r^2 \sin^2\theta}}.$$

The HJ equation corresponding to this Hamiltonian is

$$\pm\sqrt{n^2 - \frac{1}{r^2}\left(\frac{\partial S}{\partial \theta}\right)^2 - \frac{1}{r^2 \sin^2\theta}\left(\frac{\partial S}{\partial \phi}\right)^2} + \frac{\partial S}{\partial r} = 0,$$

which implies that S satisfies the equation

$$\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \theta}\right)^2 + \frac{1}{r^2 \sin^2\theta}\left(\frac{\partial S}{\partial \phi}\right)^2 = n^2.$$

Recalling the expression for the gradient of a function in spherical coordinates, we see that the last equation amounts to $(\nabla S)^2 = n^2$, which coincides with Equation (6.108).

6.45 Since the refractive index is a function of r only, each light ray lies on a plane passing through the origin. Hence, in order to find the light rays we can consider the two-dimensional eikonal equation

$$\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \theta}\right)^2 = n^2, \tag{S.41}$$

making use of polar coordinates. We look for complete solutions of the form

$$S = A(r) + P\theta, \tag{S.42}$$

where P is a constant. Substituting (S.42) into (S.41) we obtain

$$A(r) = \int \sqrt{n^2 - \frac{P^2}{r^2}}\, \mathrm{d}r = \int \frac{\sqrt{a^2r^2 - P^2(b+r^2)^2}}{r(b+r^2)} \mathrm{d}r,$$

that is,

$$S(r,\theta,P) = \int \frac{\sqrt{a^2r^2 - P^2(b+r^2)^2}}{r(b+r^2)} \mathrm{d}r + P\theta.$$

The light rays are determined by (note that we do not have to specify which variable, r or θ, is taken as the independent one)

$$\begin{aligned} Q = \frac{\partial S}{\partial P} &= \theta - \int \frac{P(b+r^2)\,\mathrm{d}r}{r\sqrt{a^2r^2 - P^2(b+r^2)^2}} \\ &= \theta - \int \frac{P(b+r^2)\,\mathrm{d}r}{r\sqrt{(a^2-4bP^2)r^2 - P^2(b-r^2)^2}}. \end{aligned}$$

The change of variable

$$P(r^2-b) = \sqrt{a^2-4bP^2}\, r \cos\alpha$$

leads to $Q = \theta + \alpha$ and, therefore, the light rays are given by

$$P(r^2-b) = \sqrt{a^2-4bP^2}\, r \cos(\theta - Q),$$

which is the equation of a circle of radius $R = a/2|P|$, with its center at a distance $\sqrt{R^2-b}$ from the origin (cf. the solution of Exercise 6.14).

6.46 If $x = x(s), y = y(s), z = z(s)$ are the Cartesian coordinates of a curve parameterized by its arclength, then the norm of the tangent vector $(\mathrm{d}x/\mathrm{d}s, \mathrm{d}y/\mathrm{d}s, \mathrm{d}z/\mathrm{d}s)$ is equal to 1. On the other hand, if the curve is orthogonal to the level surfaces of $S(x, y, z)$, then its tangent vector must be proportional to the gradient of S, whose norm is equal to the refractive index n, according to the eikonal equation. Hence,

$$\frac{dx_i}{ds} = \pm \frac{1}{n} \frac{\partial S}{\partial x_i}$$

and, making use of the chain rule and the eikonal equation, we see that

$$\frac{d}{ds}\left(n \frac{dx_i}{ds}\right) = \pm \frac{d}{ds}\left(\frac{\partial S}{\partial x_i}\right) = \pm \frac{dx_j}{ds} \frac{\partial^2 S}{\partial x_j \partial x_i} = \frac{1}{n} \frac{\partial S}{\partial x_j} \frac{\partial^2 S}{\partial x_j \partial x_i}$$
$$= \frac{1}{2n} \frac{\partial}{\partial x_i}\left(\frac{\partial S}{\partial x_j} \frac{\partial S}{\partial x_j}\right) = \frac{1}{2n} \frac{\partial}{\partial x_i} n^2 = \frac{\partial n}{\partial x_i},$$

which is the ray equation (see Exercise 6.42), thus showing that the curves orthogonal to the surfaces $S =$ const. are possible light rays.

6.47 The light rays are obtained from

$$Q = \frac{\partial S}{\partial P} = \frac{1}{2} a\big[-(x^2 - y^2) \sin P + 2xy \cos P\big],$$

using the fact that Q and P are constants of motion. The curves given by this equation are hyperbolas. The "wavefronts," given by $S =$ const., are also hyperbolas.

A straightforward computation gives

$$\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 = (ax \cos P + ay \sin P)^2 + (-ay \cos P + ax \sin P)^2 = a^2(x^2 + y^2).$$

Comparing with Equation (6.108) we find that S is a solution of the two-dimensional eikonal equation with a refractive index $n = a\sqrt{x^2 + y^2}$.

References

1. Bowman, F. (2010). *Introduction to Bessel Functions* (Dover, New York).
2. Calkin, M.G. (1996). *Lagrangian and Hamiltonian Mechanics* (World Scientific, Singapore).
3. Colwell, P. (1993). *Solving Kepler's Equation, Over Three Centuries* (Willmann–Bell, Richmond, Virginia).
4. Corben, H.C. and Stehle, P. (1994). *Classical Mechanics*, 2nd ed. (Dover, New York).
5. Crampin, M. and Pirani, F.A.E. (1986). *Applicable Differential Geometry* (Cambridge University Press, Cambridge).
6. Das, A. (1989). *Integrable Models* (World Scientific, Singapore).
7. Eves, H. (1980). *Elementary Matrix Theory* (Dover, New York).
8. Fasano, A. and Marmi, S. (2006). *Analytical Mechanics: An Introduction* (Oxford University Press, Oxford).
9. Greenwood, D.T. (1997). *Classical Dynamics* (Dover, New York).
10. Hermann, R. (1968). *Differential Geometry and the Calculus of Variations* (Academic Press, New York).
11. Pars, L.A. (1965). *A Treatise on Analytical Dynamics* (Heinemann, London).
12. Percival, I. and Richards, D. (1982). *Introduction to Dynamics* (Cambridge University Press, Cambridge).
13. Perelomov, A.M. (1990). *Integrable Systems of Classical Mechanics and Lie Algebras*, Vol. I (Birkhäuser, Basel).
14. Sneddon, I.N. (2006). *Elements of Partial Differential Equations* (Dover, New York).
15. Synge, J.L. (1937). *Geometrical Optics: An Introduction to Hamilton's Method* (Cambridge University Press, Cambridge).
16. Torres del Castillo, G.F. (1999). Symmetry of the Kepler problem in classical mechanics, Rev. Mex. Fís. **45**, 234.
17. Torres del Castillo, G.F., Cruz Domínguez, H.H., de Yta Hernández, A., Herrera Flores, J.E. and Sierra Martínez, A. (2014). Mapping of solutions of the Hamilton–Jacobi equation by an arbitrary canonical transformation, Rev. Mex. Fís. **60**, 301.
18. Torres del Castillo, G.F. and Herrera Flores, J.E. (2016). Symmetries of the Hamiltonian operator and constants of motion, Rev. Mex. Fís. **62**, 135.
19. Torres del Castillo, G.F. and Nájera Salazar, B.C. (2017). Transformation of a wavefunction under changes of reference frame, Rev. Mex. Fís. **63**, 185.

G. F. Torres del Castillo, *An Introduction to Hamiltonian Mechanics*, Birkhäuser Advanced Texts Basler Lehrbücher, https://doi.org/10.1007/978-3-319-95225-3

20. Torres del Castillo, G.F. and Rubalcava-García, I. (2017). Variational symmetries as the existence of ignorable coordinates, Eur. J. Phys. **38**, 025002.
21. Wells, C.G. and Siklos, T.C. (2007). The adiabatic invariance of the action variable in classical mechanics, Eur. J. Phys. **28**, 105.
22. Whittaker, E.T. (1989). *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 4th ed. (Cambridge University Press, Cambridge).

Index

G. F. Torres del Castillo, *An Introduction to Hamiltonian Mechanics*, Birkhäuser Advanced Texts Basler Lehrbücher, https://doi.org/10.1007/978-3-319-95225-3