Alexei Deriglazov

Classical Mechanics

Hamiltonian and Lagrangian Formalism

Springer

Classical Mechanics

Alexei Deriglazov

Classical Mechanics

Hamiltonian and Lagrangian Formalism

Springer

Alexei Deriglazov
Depto. de Matemática
Universidade Federal de Juiz de Fora
36036-330 Juiz de Fora
MG, Brazil
alexei.deriglazov@ufjf.edu.br

ISBN 978-3-642-14036-5 e-ISBN 978-3-642-14037-2
DOI 10.1007/978-3-642-14037-2
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010932503

Cover design: eStudio Calamar S.L., Heidelberg

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Formalism of classical mechanics underlies a number of powerful mathematical methods, widely used in theoretical and mathematical physics [1–11]. In these lectures we present some selected topics of classical mechanics, which may be useful for graduate level students intending to work in one of the branches of a vast field of theoretical physics. Except for the last chapter, which is devoted to the discussion of singular theories and their local symmetries, the topics selected correspond to the standard course of classical mechanics.

For the convenience of the reader, we have tried to make the material of different chapters as independent as possible. So, the reader who is familiar with Lagrangian mechanics can proceed to any one of Chaps. 3, 4, 5, 6, 7, 8 after reading the second chapter.

In our presentation of the material we have tried, where possible, to replace intuitive motivations and "scientific folklore" by exact proofs or direct computations. To illustrate how classical-mechanics formalism works in other branches of theoretical physics, we have presented examples related to electrodynamics, as well as to relativistic and quantum mechanics. Most of the suggested exercises are directly related to the main body of the text.

While in some cases the formalism is developed beyond the traditional level adopted in the standard textbooks on classical mechanics [12–14], the only mathematical prerequisites are some knowledge of calculus and linear algebra.

In the frameworks of classical and quantum theories, the Hamiltonian and Lagrangian formulations each have advantages and disadvantages. Since our focus here is Hamiltonian mechanics, let us mention some of the arguments for using this type of formalism.

- There is a remarkable democracy between variables of position and velocity in Nature: being *independent* one from another, they contain complete information on the properties of a classical system at a given instance. Besides, just the positions and velocities at the initial instant of time are necessary and sufficient to predict an evolution of the system. In Lagrangian formalism this democracy, while reflected in the initial conditions, is not manifest in the course of evolution, since only variables of position are treated as independent in Lagrangian equations. Hamiltonian formalism restores this democracy, treating positions and velocities on equal footing, as independent coordinates that parameterize a phase space.

- According to the canonical quantization paradigm, the construction of the Hamiltonian formulation for a given classical system is the first necessary step in the passage from classical to quantum theory. It is sufficient to point out that quantum evolution in the Heisenberg picture is obtained from the Hamiltonian equations through replacement of the phase-space variables by corresponding operators. As to the operators, their commutators are required to resemble the Poisson brackets of the phase-space variables.
- The conventional way to describe a relativistic theory is to formulate it in terms of a singular Lagrangian (the singularity is the price we pay for the manifest relativistic invariance of the formulation). It implies a rather complicated structure of Lagrangian equations, which may consist of both second and first-order differential equations as well as algebraic ones. Besides, there may be identities present among the equations, which implies functional arbitrariness in the corresponding solutions. It should be mentioned that, in the modern formulation, physically interesting theories (electrodynamics, gauge field theories, the standard model, string theory, etc.) are of this type. In this case, Hamiltonian formulation gives a somewhat clearer geometric picture of classical dynamics [8]: all the solutions are restricted to lying on some surface in the phase space, while the above-mentioned arbitrariness is avoided by postulating classes of equivalent trajectories. Physical quantities are then represented by functions defined in these classes. The procedure for investigation of this picture is based entirely on the use of special coordinates adopted to the surface, which in turn require a rather detailed development of the theory of canonical transformations. Altogether Hamiltonian formulation leads to a self-consistent physical interpretation of a general singular theory, forming the basis for numerous particular prescriptions and approaches to quantization of concrete theories [10].

Juiz de Fora, July 2010
Alexei Deriglazov

Notation and conventions

The terminology of classical mechanics is not universal. To avoid any confusion, the quantities of the configuration (phase) space are conventionally called Lagrangian (Hamiltonian) quantities.

Generalized coordinates of the configuration space are denoted by q^a. Latin indices from the beginning of the alphabet a, b, c, and so on generally range from 1 to n, $a = 1, 2, \ldots, n$.

Phase space coordinates are often denoted by one letter $z^i = (q^a, p_b)$. Latin indices from the middle of the alphabet i, j, k, and so on generally range from 1 to $2n$, $i = 1, 2, \ldots, 2n$.

Greek indices from the beginning of the alphabet α, β, γ are used to denote some subgroup of the group of variables, for example $q^a = (q^1, q^\alpha)$, $\alpha = 2, 3, \ldots, n$.

Repeated indices are generally summed, unless otherwise indicated. The "up" and "down" position of the index of any quantity is fixed. For example, we write q^a, p_b and never any other way.

Time variable is denoted either by τ or by t. A dot over any quantity denotes the time-derivative of that quantity

$$\dot{q}^a = \frac{dq^a}{d\tau},$$

while partial derivatives are denoted by

$$\frac{\partial L(q)}{\partial q^a} = \partial_a L, \qquad \frac{\partial H(z)}{\partial z^i} = \partial_i H,$$

The same symbol is generally used to denote a variable and a function. For example, we write $z'^i = z'^i(z^j)$, instead of the expression $z'^i = f^i(z^j)$ for the change of coordinates.

The notation

$$F(q, v)|_{v(z)} \equiv F(q, v)|_{v=v(z)} \equiv F(q, v)|,$$

implies the substitution of the function $v^a(z)$ in place of the variable v^a.

Contents

Chapter 1
Sketch of Lagrangian Formalism

1.1 Newton's Equation

System of particles. To start with, we recall how a system of particles is described in classical mechanics. Analytic description is achieved by introducing three-dimensional Euclidean space equipped with a Cartesian coordinate system. Then its points are labeled by position vectors $M \leftrightarrow \vec{r} = (x^1, x^2, x^3) \equiv (x, y, z)$. The time evolution of a particle is presented by a curve $\vec{r} = \vec{r}(t)$. The evolution is governed by *Newton's equation*

$$m\ddot{\vec{r}} = \vec{F}(\vec{r}, \dot{\vec{r}}, t) \quad \Leftrightarrow \quad m\ddot{x}^a = F^a(x^b, \dot{x}^b, t), \quad a = 1, 2, 3. \tag{1.1}$$

For a system of particles with the position vectors $\vec{r}_i$, $i = 1, 2, \ldots, N$ we write

$$m_i\ddot{\vec{r}}_i = \vec{F}_i(\vec{r}_j, \dot{\vec{r}}_j, t). \tag{1.2}$$

We assume that the force $\vec{F}$ in classical mechanics is a known function of indicated arguments (or derivable from a potential). So Eq. (1.1) relates accelerations, velocities and coordinates, that is, it represents a system of three ordinary differential second-order equations for determining three functions $x^a(t)$.

Example Electric charges in movement produce electromagnetic force in the space around them. This can be described by vectors of electric $\vec{E}(t, x^a)$ and magnetic $\vec{B}(t, x^a)$ fields given at each space-time point. Then Newton's equation of a particle with electric charge e on this external field is

$$m\ddot{\vec{r}} = e\vec{E}(t, \vec{r}) + \frac{e}{c}[\dot{\vec{r}}, \vec{B}(t, \vec{r}\,)]. \tag{1.3}$$

Here $[\dot{\vec{r}}, \vec{B}\,]^a = \epsilon^{abc}\dot{x}^b B^c$ is a *vector product*, and c is a universal constant (see Sect. 1.3).

A. Deriglazov, *Classical Mechanics*, DOI 10.1007/978-3-642-14037-2_1,

Particular and general solutions. Mathematically, Eq. (1.1) belong to the class of *normal* systems, that is *all* higher derivatives $\ddot{x}^a$ are separated on the left-hand side of the equations. According to the theory of differential equations, a normal system has well-established properties. In particular, under known restrictions on the right-hand side, the theorem of the existence and uniqueness of a solution holds: given the numbers x_0^a, v_0^a, there is (at least locally) a unique solution $x^a(\tau)$ of the system (1.1) that obeys the initial conditions $x^a(0) = x_0^a$, $\dot{x}^a(0) = v_0^a$.

When equations can be supplemented by initial conditions that guarantee a unique solution, we say that the equations admit formulation of the *Cauchy problem*. Using this terminology, a normal system admits formulation of the Cauchy problem.

This theorem implies that ordinary differential equation admits an infinite number of solutions. They can be described simultaneously using the notion of a general solution. It is not difficult to forecast that the family of solutions can be parameterized by six parameters. Roughly speaking, to kill two derivatives acting, for example, on x^1, we need to carry out two integrations. This implies the appearance of two integration constants, say c^1, d^1, in the resulting expression. In this way we arrive at the notion of a general solution defined as follows. A function of (1+6) variables $\vec{r}(t, c^a, d^a)$ is called a *general solution* to the system (1.1) if (a) it satisfies the system for any values of c^a, d^a; (b) given the initial conditions $\vec{r}_0$, $\vec{v}_0$, there are numbers $\tilde{c}^a$, $\tilde{d}^a$ such that $\vec{r}(0, \tilde{c}^a, \tilde{d}^a) = \vec{r}_0$, $\dot{\vec{r}}(0, \tilde{c}^a, \tilde{d}^a) = \vec{v}_0$.

Exercise
Confirm that any particular solution to the normal system is contained in its general solution.

The physical content of these mathematical facts can be summarized as follows. First, only positions and velocities at a given instance are necessary to predict the future of a system (we need not know, for example, accelerations). It is said that $\vec{r}_0$, $\vec{v}_0$ unambiguously determine the instantaneous state of a system. Second, a system evolves in time in a unique way. In contrast to quantum mechanics, evolution in classical mechanics has a causal character. These properties of Newton's universe are known as *Newton's principle of determinism*.

We recall the definition of some quantities that will play a fundamental role in the discussion of Lagrangian formalism.

Kinetic energy. Let $\vec{r}(t)$ be a solution to Eq. (1.1). The work done by the force $\vec{F}$ is equal to the value of the line integral along the curve, and using Newton's equation can be computed as follows:

$$A = \int_{M_1}^{M_2} \vec{F} d\vec{r} = \int_{t_1}^{t_1} m \frac{d\vec{v}}{dt} \vec{v} dt = \int_{t_1}^{t_1} \frac{d}{dt} \left(\frac{1}{2} m \vec{v}^{\,2} \right) dt = \frac{1}{2} m \vec{v}^{\,2} \Big|_{t_1}^{t_2}. \quad (1.4)$$

The quantity $T = \frac{1}{2}m\vec{v}^{\,2}$ is called *kinetic energy*. It is said that the work produces a change in the kinetic energy of a particle

$$A = T(t_2) - T(t_1). \tag{1.5}$$

Potential energy. Properties of conservative force. To proceed further we restrict ourselves to the case of a *conservative* system. A field of force is *conservative* (or *potential*), if it can be derived as the gradient of a function $U(x^a)$

$$F^a = -\frac{\partial U}{\partial x^a}. \tag{1.6}$$

This function is called the *potential energy* of a system. Note that $U + const$ leads to the same force as U. So, potential energy is defined with only an additive constant. It is often used to choose a zero value for the potential energy at a desired point. For the work done by the potential force we write

$$A = -\int_{M_1}^{M_2} \vec{\nabla} U d\vec{r}, \quad \text{here} \quad \vec{\nabla} \equiv \left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right). \tag{1.7}$$

Let us enumerate the properties of a conservative force

$$[\vec{\nabla}, \vec{F}] = 0, \tag{1.8}$$

$$\oint_{\gamma} \vec{F} d\vec{r} = 0, \tag{1.9}$$

$$\int_{\gamma(1,2)} \vec{F} d\vec{r} = \int_{\beta(1,2)} \vec{F} d\vec{r}, \tag{1.10}$$

$$A = -\left[U(\vec{r}_2) - U(\vec{r}_1)\right]. \tag{1.11}$$

Equation (1.8) states that a potential field is curl-free. This follows from direct computation: $[\vec{\nabla}, \vec{F}]^a = \epsilon^{abc}\partial_b\partial_c U = 0$.

Equation (1.9) states that a potential field does not produce work along any closed line γ. This follows from (1.8) applying Stock's theorem, $\oint_\gamma \vec{F} d\vec{r} = \int_{S_\gamma} [\vec{\nabla}, \vec{F}] d\vec{s} = 0$.

Equation (1.10) states that the work done by a potential field does not depend on a choice of line (either $\gamma(1,2)$ or $\beta(1,2)$) that connects points 1 and 2. To confirm this, apply (1.9) to the closed line $\gamma(1,2) \cup \beta(2,1)$.

Finally, Eq. (1.11) states that the work done by a potential field is equal to the difference of potential energies at the initial and final points. To see this, notice that according to Eq. (1.10) we can compute the work (1.7) using any line connecting M_1 and M_2. Let us take the line composed of intervals parallel to the coordinate axes, $(x_1, y_1, z_1) \to (x_2, y_1, z_1) \to (x_2, y_2, z_1) \to (x_2, y_2, z_2)$, see Fig. 1.1 on page 4. Then the integral can be computed directly

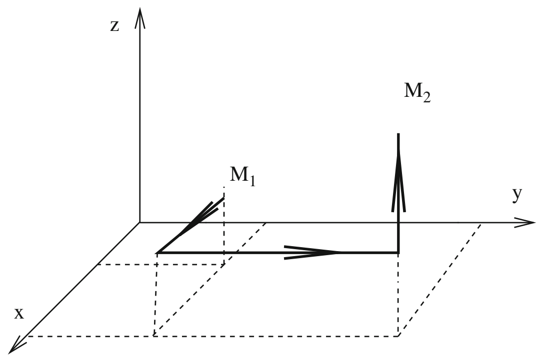

Fig. 1.1 Integration contour used in the proof of Eq. (1.11)

$$\begin{aligned} A &= -\int_{M_1}^{M_2} \vec{F} d\vec{r} = -\int_{M_1}^{M_2} \vec{\nabla} U d\vec{r} \\ &= -\int_{x_1}^{x_2} dx \partial_x U(x, y_1, z_1) - \int_{y_1}^{y_2} dy \partial_y U(x_2, y, z_1) \\ &\quad - \int_{z_1}^{z_2} dz \partial_z U(x_2, y_2, z) = -\left[U(\vec{r}_2) - U(\vec{r}_1) \right]. \end{aligned} \tag{1.12}$$

It is known that on a plane the conditions (1.6), (1.8), (1.9), (1.10), and (1.11) are mutually equivalent; any one of them can be taken as a definition of a (two-dimensional) potential field.

Law of conservation of total energy. Let $x^a(t)$ be a solution to Newton's equation with a conservative force

$$m\ddot{x}^a + \frac{\partial U(x)}{\partial x^a} = 0. \tag{1.13}$$

Comparing (1.5) with (1.11) we conclude that in a movement from one point to another, the change in kinetic energy is always balanced by a change in potential energy. This can be written in the form of a law of conservation

$$[T + U]|_{M_1} = [T + U]|_{M_2}, \tag{1.14}$$

The quantity $E \equiv T + U$ is called the (*total*) *energy* of the system. Eq. (1.14) represents the *law of conservation of total energy* stating that E of a conservative system takes the same value throughout a solution.

Exercise
Total energy is not preserved when the potential depends explicitly on time, $U(x, t)$. Explain why the reasoning presented above does not work for this case.

It is instructive to obtain this result once again, this time as a direct consequence of the equations of motion. Multiplying Eq. (1.13) by $\dot{x}^a(t)$ we obtain

$$m\ddot{x}^a\dot{x}^a + \frac{\partial U}{\partial x^a}\dot{x}^a = \left[\frac{1}{2}m\vec{v}^{\,2} + U\right]^{\cdot} = 0. \tag{1.15}$$

that is $E = \frac{1}{2}m\vec{v}^{\,2} + U = \text{const}$ on the solutions. Energy is an example of a conserved quantity, which can be defined as a function of positions and velocities that is preserved along true trajectories of the system. Let us recall two more examples.

Law of conservation of angular momentum. Consider a particle in a *central field*. The central field is defined by a potential depending only on a distance to a given point. Choosing a coordinate system with its origin at that point, the potential reads $U = U(r)$, where $r = |\vec{r}| = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}$. Then Newton's equation is

$$m\ddot{\vec{r}} + \frac{dU}{dr}\frac{\vec{r}}{r} = 0. \tag{1.16}$$

Computing the vector product with $\vec{r}$ we obtain

$$\begin{aligned}\left[\vec{r}, m\frac{d}{dt}\vec{v}\right] + \left[\vec{r}, \frac{dU}{dr}\frac{\vec{r}}{r}\right] &= \frac{d}{dt}[\vec{r}, m\vec{v}] - m[\vec{v}, \vec{v}] + \frac{1}{r}\frac{dU}{dr}[\vec{r}, \vec{r}] \\ &= \frac{d}{dt}[\vec{r}, m\vec{v}] = 0. \end{aligned} \tag{1.17}$$

This implies the *law of conservation of angular momentum*

$$\vec{L} \equiv [\vec{r}, m\vec{v}] = \text{const}. \tag{1.18}$$

One remarkable consequence of Eq. (1.18) is that the particle orbit in the central field is a planar one, see Fig. 1.2 on page 6. To see this, take a solution $\vec{r}(t)$ of Newton's equation (1.16). Note that its scalar product with $\vec{L}$ vanishes, $(\vec{r}(t), \vec{L}) = (\vec{r}, [\vec{r}, m\vec{v}]) = 0$. That is, the position vector is orthogonal to the *fixed* vector $\vec{L}$ at any instant of motion. So, $\vec{r}(t)$ at any t lies on a plane containing the center of force and perpendicular to $\vec{L}$. The vector $\vec{L}$ can be determined from the initial conditions: $\vec{L} = [\vec{r}_0, m\vec{v}_0]$. So the plane of motion is the one that contains the vectors of the initial position $\vec{r}_0$ and the initial velocity $\vec{v}_0$.

Conservation law of total momentum. Consider now a two-particle system. Supposing that forces acting upon the particles obey the *third law of Newton*, $\vec{F}_2 = -\vec{F}_1$, Newton equations read

$$m_1\ddot{\vec{r}}_1 = \vec{F}_1, \qquad m_2\ddot{\vec{r}}_2 = -\vec{F}_1. \tag{1.19}$$

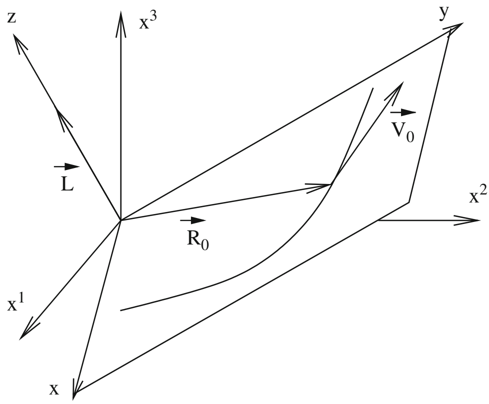

Fig. 1.2 Due to conservation of angular momentum, the trajectory lies on a plane perpendicular to $\vec{L}$. So, in coordinates x, y, z, the three-dimensional Kepler's problem reduces to a two-dimensional one

Taking their sum, one immediately obtains the *law of conservation of total momentum*

$$m_1\vec{v}_1 + m_2\vec{v}_2 = \text{const.} \tag{1.20}$$

The law of conservation, being in fact a first-order differential equation, can be used for simplification of equations of motion. Below we present examples of how this works.

Examples

1. **A one-particle conservative system on a straight line** can be solved by quadrature for an arbitrary potential. In this case, Newton's equation

$$m\ddot{x} + \frac{dU}{dx} = 0, \tag{1.21}$$

is *equivalent*[1] to the law of conservation of energy

$$\frac{1}{2}m\left(\frac{dx}{dt}\right)^2 + U(x) = E. \tag{1.22}$$

[1] We have seen that (1.21) implies (1.22). Conversely, the derivative of (1.22) with respect to t implies (1.21).

So we can study the latter. It represents a first-order differential equation that admits separation of variables, t and x, and can then be immediately integrated out

$$dt = \frac{dx}{\sqrt{\frac{2}{m}(E - U(x))}}, \qquad t - t_0 = \int \frac{dx}{\sqrt{\frac{2}{m}(E - U(x))}}. \quad (1.23)$$

Computing the integral on r.h.s. we obtain $t - t_0 = f(x, E)$. The inverse function, $x = g(t, t_0, E)$ represents a general solution with two integration constants t_0, E.

2. **Reduction of the three-dimensional Kepler's problem to a two-dimensional one.** The *Kepler's problem* consists of the description of a particle in a central field with a potential that is proportional to the inverse degree of r

$$U = -\frac{\alpha}{r}, \quad r = |\vec{r}| = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}, \quad \alpha = \text{const.} \quad (1.24)$$

Newtonian gravitational attraction and Coulomb interaction belong to this class of central fields. Newton's equation (1.16) acquires the form

$$m\ddot{x}^a + \frac{\alpha x^a}{r^3} = 0, \qquad a = 1, 2, 3. \quad (1.25)$$

As we have seen above, the particle trajectory lies on a plane that passes through the center of the field and is perpendicular to the constant vector $\vec{L}$. Let us introduce the coordinate system (x, y, z) with its z-axis along $\vec{L}$ and with the x and y axis on the plane of motion; see Fig. 1.2 on page 6.

Equations of motion for the new variables have the same form (1.25) (this will be discussed in some detail in the next section). In this system we have $\vec{r}(t) = (x(t), y(t), 0)$, that is, the z-coordinate has trivial dynamics. The third equation of the system (1.25) is satisfied and can be omitted. So, conservation of the angular momentum allows us to simplify the problem: the three-dimensional problem reduces to a two-dimensional one

$$m\ddot{x} + \frac{\alpha x}{\sqrt{x^2 + y^2}} = 0, \qquad m\ddot{y} + \frac{\alpha y}{\sqrt{x^2 + y^2}} = 0. \quad (1.26)$$

This two-dimensional problem will be solved in Sect. 1.6.

Exercise
As we have seen, Eqs. (1.25) imply conservation of energy and angular momentum

$$\frac{1}{2}m\dot{\vec{r}}^{\,2} - \frac{\alpha}{r} = E, \qquad [\vec{r}, m\dot{\vec{r}}\,] = \vec{L}. \tag{1.27}$$

Show that, in turn, Eqs. (1.27) imply (1.25). Hence the systems (1.25) and (1.27) are equivalent.

1.2 Galilean Transformations: Principle of Galilean Relativity

Cartesian coordinates and the evolution parameter which appear in Newton equations represent an idealization of the data of the measurement devices (rulers and clocks) used by an observer in his laboratory. The laboratory is called an *observer* or (*reference*) *frame*.[2]

The first law of Newton states the existence of *inertial frames* with the following property: motion of a free particle in any one of them *looks* as rectilinear motion along a straight line, $\frac{d^2x^a}{dt^2} = 0$. The principle of Galilean relativity *postulates* the set of transformations relating space-time coordinates of inertial frames. It also states that equations of motion of any mechanical system, like the free one, retain their *form* unchanged under the transformations, $S'_a = D_{ab}S_b$ (here $S_a = 0$ ($S'_a = 0$) stands for equations of motion written by observer O (O'), and D is an invertible matrix). This property is called *covariance* of the equations. Below we present the mathematical formulation of the principle of Galilean relativity and discuss its physical content.

Formulation of the Galilean principle. Consider an n-particle system with a potential that depends on the relative distances between the particles. That is $U(r_{jk})$ represents a function of the variables r_{jk}, $j < k$

$$r_{jk} = |\vec{r}_j - \vec{r}_k| = \sqrt{\sum_{a=1}^{3}(x^a_j - x^a_k)^2}. \tag{1.28}$$

Newton equations read

$$m_i\frac{d^2\vec{r}_i}{dt^2} + \sum_{k=1}^{n}\frac{\partial U}{\partial r_{ik}}\frac{\vec{r}_i - \vec{r}_k}{r_{ik}} = 0, \qquad i = 1, 2, \ldots, n. \tag{1.29}$$

[2] Due to the presence of the clocks, the term “reference frame” is used instead of “coordinate system”.

We point out that the equations describe a rather general class of interacting systems, including Newtonian gravity and Coulomb forces.

A remarkable algebraic property of the equations is that they take the same *form* if we make the substitution (called below the *Galilean transformation*)

$$t = t' + a, \tag{1.30}$$

$$\vec{r}_i = R\vec{r}_i{}' + \vec{V}t' + \vec{C}, \tag{1.31}$$

where R^{ab} is an orthogonal matrix, $RR^T = 1$, and $\vec{V}$, $\vec{C}$, a are arbitrary constants. Denoting the l.h.s. of Eq. (1.29) as $\vec{S}_i(t, \vec{r})$, one verifies that the equality

$$\vec{S}_i(t, \vec{r}) = R\,\vec{S}_i(t', \vec{r}'), \tag{1.32}$$

holds. This property is known as *covariance* of equations. Note that there are a lot of transformations that do not leave the equations covariant (take, for example, $\vec{r}_i = \vec{r}_i{}' + \vec{V}t^3$). The covariance (1.32) of basic equations of classical mechanics under Galilean transformations represents a mathematical formulation for the *principle of Galilean relativity*.

Physical content of the Galilean principle. Its physical content is two-fold, due to the possibility of two different geometric interpretations for the substitution (1.30) and (1.31). In short, it can be treated either as the passage from one reference frame to another (the so-called *passive point of view*), or dislocation of the system under investigation from one region of space-time to another (*the active point of view*). Let us discuss these in further detail.

Passive point of view. In this case $(\vec{r}, t)$ and $(\vec{r}\,', t')$ are regarded as coordinates of the *same* space-time point in two reference frames, O and O'. Suppose they differ by displacement, by orientation of axis and are in relative motion with constant velocity. In non-relativistic mechanics it is *postulated* that formulas (1.30) and (1.31) represent the law of transformation from one frame to the other. It consists of the following transformations.

1. The transformation $x^a = x'^a + C^a$ represents the space displacement: O sees the origin of O' in the point with the position vector $\vec{C}$.
2. The transformation $x^a = R^{ab}x'^b$, where $RR^T = 1$, represents rotation: O sees the coordinate axis of O' rotated through R. To see this, note that O and O', related by this transformation, have the same origin. Besides, it does not change distances between points, in particular, $|\vec{r}\,'| = |\vec{r}|$.
   Any rotation can be parameterized by three numbers. For instance, they can be components of the vector $\vec{\alpha} = (\alpha^1, \alpha^2, \alpha^3)$ directed along the rotation axis, with a length equal to the rotation angle (see the end of this section for details).
3. The transformation $x^a = x'^a + V^a t$ is known as the *Galilean boost*. O sees O' moving with velocity $\vec{V}$ and passing through the origin of O at $t = 0$; see Fig. 1.4 on page 20.
4. The transformation $t = t' + a$ is the time displacement: O sees the O' clock running behind his own by a time a. Notice that the time *intervals* measured by O and O' are the same, $\triangle t = \triangle t'$.

In this interpretation, the covariance (1.32) implies that in all inertial frames the physical system obeys equations of motion of the same *form*. All inertial observers, each using his own coordinates, will discover the same laws of motion studying a given physical system, see Fig. 1.3a on the page 10. From a practical point of view, O, who discovered Eqs. (1.29), need not worry how to write them, if he intends to use another inertial frame – they will be the same. For instance, this has been already used in Example 2 of the previous section; see the discussion after Eq. (1.25).

Active point of view. In this case $(\vec{r}, t)$ and $(\vec{r}\,', t')$ are regarded as coordinates of different space-time points in a given reference frame. Then (1.30) and (1.31) represents the transformation that turns the primed point into an unprimed one. For instance, imagine a physical system located in the vicinity of the observer's origin. Then the transformation (1.31) rotates it by R, displaces it over a distance $\vec{C}$ and makes it move with velocity $\vec{V}$. Besides, Eq. (1.30) means that the system is considered by the observer at a later time.

In this interpretation, covariance means that two copies of a mechanical system, related by the Galilean transformation, obey the same equations of motion, see Fig. 1.3b on page 10. It implies that in this reference frame our space and time look homogeneous as well as the fact that the space looks isotropic. Besides, a mechanical system at rest and its copy in rectilinear motion along a straight line have identical properties.

Let us summarize the discussion. Intuitively, the Galilean relativity principle (1.30), (1.31), and (1.32) can be summarized in two statements. First, different inertial observers studying the same mechanical system will discover laws of motion of the same form. Second, identical experiments made by inertial observers will give identical results. Such properties as homogeneity, isotropy, ..., are implicit in the second statement.

In relativistic mechanics the ideology remains the same. The only thing that changes is the expression for the boost relating two frames in relative motion; see the next section.

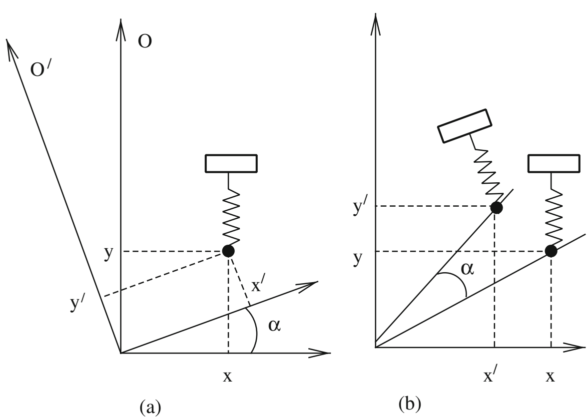

Fig. 1.3 Rotation transformation. **(a)** Passive point of view. **(b)** Active point of view

Structure of a rotation matrix. An arbitrary 3×3 matrix a is determined by its nine matrix elements a_{ab}. So, the set of matrices forms a nine-dimensional space with coordinates a_{ab}. The rotation matrices form a subset defined by the condition

$$R^T R = 1, \quad \text{or} \quad R^T = R^{-1}. \tag{1.33}$$

Since $R^T R$ is a symmetric matrix, the system (1.33) consists of 6 independent equations. So the subset (1.33) can be parameterized by $9 - 6 = 3$ coordinates (intuitively, a rotation can be uniquely specified by the pointing of a vector $\vec{\alpha} = (\alpha^1, \alpha^2, \alpha^3)$ directed along the rotation axis, with a length equal to the rotation angle).

To parameterize the set (1.33), we use the standard procedure known from group theory. We will need to use the exponential of a matrix. Given the matrix a, its exponential e^a is a matrix defined by the power series

$$\left(e^a\right)_{ab} = \left(\sum_{n=0}^{\infty} \frac{a^n}{n!}\right)_{ab} = \delta_{ab} + a_{ab} + \frac{1}{2!} a_{ac} a_{cb} + \dots . \tag{1.34}$$

Some properties of the exponential are

$$\text{if} \quad ab = ba, \quad \text{then} \quad e^a e^b = e^{a+b}; \tag{1.35}$$

$$e^a e^b = e^{a+b+\frac{1}{2}[a,b]+O^3(a,b)}, \quad \text{where} \quad [a,b] = ab - ba; \tag{1.36}$$

$$\left(e^a\right)^T = e^{a^T}; \tag{1.37}$$

$$\left(e^a\right)^{-1} = e^{-a}. \tag{1.38}$$

From the definition it follows that $e^{\mathbf{0}} = \mathbf{1}$. Besides, if a is close to the null matrix, then $e^a \approx \mathbf{1} + a$, that is e^a will be close to the unit matrix. It is known that the exponential establishes an isomorphism of a neighborhood of a null matrix onto a neighborhood of a unit matrix.

Let us try to represent a small rotation R in the form of the exponential of another matrix: $R = e^\omega$. The condition of orthogonality of R gives $(e^\omega)^T = (e^\omega)^{-1}$, or $e^{\omega^T} = e^{-\omega}$, or $\omega^T = -\omega$, that is, the matrix ω must be antisymmetric. Conversely, any antisymmetric matrix generates an orthogonal matrix by an exponential map. An antisymmetric matrix is parameterized by three coordinates. Let us represent it as follows

$$\begin{aligned}
\omega &= \begin{pmatrix} 0 & \omega_{12} & -\omega_{31} \\ -\omega_{12} & 0 & \omega_{23} \\ \omega_{31} & -\omega_{23} & 0 \end{pmatrix} \\
&= \omega_{23} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} + \omega_{31} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + \omega_{12} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
&\equiv \alpha^1 T_1 + \alpha^2 T_2 + \alpha^3 T_3,
\end{aligned} \tag{1.39}$$

where

$$\alpha^a = \frac{1}{2}\epsilon^{abc}\omega_{bc}, \quad \text{that is} \quad \alpha^1 = \omega_{23}, \quad \alpha^2 = \omega_{31}, \quad \alpha^3 = \omega_{12}. \tag{1.40}$$

Note the meaning of the exponential trick: resolution to the equation $\omega^T = -\omega$ is a much easier task than $R^T = R^{-1}$!

Returning to the rotation matrix, it can now be presented as

$$R = e^{\alpha^a T_a} \approx \mathbf{1} + \alpha^a T_a, \qquad a = 1, 2, 3. \tag{1.41}$$

The matrices T_a which appear in the formalism are called *generators of rotations*. Since exponential is an isomorphism, the coordinates α^a of the matrix ω can be taken as coordinates of the orthogonal matrix as well, $R = R(\vec{\alpha})$. They have a simple geometric meaning. Note that the vector $\vec{\alpha}$ is invariant under the rotation, $R(\vec{\alpha})\vec{\alpha} = \vec{\alpha}$, so $\vec{\alpha}$ is directed along the rotation axis.

This implies that, for instance, the matrix $R = e^{\alpha^3 T_3}$ represents a rotation around the z-axis. To confirm this by direct computation, note the properties

$$(T_3)^{2n} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}^n = \begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & (-1)^n & 0 \\ 0 & 0 & 0 \end{pmatrix}; \tag{1.42}$$

$$(T_3)^{2n+1} = \begin{pmatrix} 0 & (-1)^n & 0 \\ -(-1)^n & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{1.43}$$

then

$$R = e^{\alpha^3 T_3} = \sum \frac{1}{n!}(\alpha^3)^n (T_3)^n =$$
$$\begin{pmatrix} \frac{(-1)^n(\alpha^3)^{2n}}{(2n)!} & \frac{(-1)^n(\alpha^3)^{2n+1}}{(2n+1)!} & 0 \\ -\frac{(-1)^n(\alpha^3)^{2n+1}}{(2n+1)!} & \frac{(-1)^n(\alpha^3)^{2n}}{(2n)!} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\alpha^3 & \sin\alpha^3 & 0 \\ -\sin\alpha^3 & \cos\alpha^3 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.44}$$

The corresponding transformation

$$\vec{r}\,' = R\vec{r} = \begin{pmatrix} x\cos\alpha^3 + y\sin\alpha^3 \\ -x\sin\alpha^3 + y\cos\alpha^3 \\ z \end{pmatrix}, \tag{1.45}$$

is precisely the rotation by angle α^3 around the z-axis.

Exercises

1. Prove the properties (1.35), (1.36), (1.37), and (1.38).
2. Prove that $R(\vec{\alpha})\vec{\alpha} = \vec{\alpha}$.
3. Compute the rotation matrices $e^{\alpha^1 T_1}$ and $e^{\alpha^2 T_2}$.
4. Verify that the rotation generators satisfy the algebra $[T_a, T_b] = -\epsilon^{abc} T_c$.

1.3 Poincaré and Lorentz Transformations: The Principle of Special Relativity

A very economic and clear presentation of the *special theory of relativity* can be found in [15, 16]. For a detailed discussion, see [17–19]. Here we only discuss the principle that represents the starting point in the formulation of relativistic mechanics.

Formulation of the principle of special relativity. Similarly to the Galilean principle, the special relativity principle postulates the set of transformations (called Poincaré transformations) relating space-time coordinates of inertial frames, and states that laws of motion are covariant under the Poincaré transformations. The only difference between Galilean and Poincaré transformations is in the expression for the boost relating moving observers.

Poincaré transformations. Differentiation of the Galilean boost $x^a = x'^a + V^a t$ gives a simple transformation rule for a particle velocity

$$\vec{v} = \vec{v}' + \vec{V}. \tag{1.46}$$

The velocity of a particle seen by O is equal to the velocity seen by O' plus the relative velocity between the frames. This rule contradicts the Michelson–Morley experiment, which shows that the velocity of propagation of a light front in a vacuum is the same in all inertial frames! The numerical value of this universal constant is $c = 2,998 \times 10^{10}$cm/s. According to the Michelson–Morley experiment, for the case when both the light and the observer O' move along the x-axis of the observer O, we have $c = c'$ instead of Eq. (1.46).

Hence the Galilean boost is only approximately true, when velocity of the particle and the relative velocity between the reference frames are small as compared with this universal constant. We point out that this does not mean that relativistic effects are always negligible at small velocities. For instance, a magnetic field produced by an electric current passing through a wire represents a pure relativistic effect; see [20].

So, we are forced to replace the Galilean boost by another transformation, which must satisfy two requirements. First, it must be consistent with the constancy of the speed of light. Second, it must reduce to the Galilean boost at the limit $c \to \infty$. It is instructive to discuss the problem in a more general setting, looking for the most general transformation consistent with the constancy of the speed of light.

Consider the propagation of a light between two nearby closed points, separated by dt, dx^a in the O-frame. O writes the law of motion $c^2 dt^2 - (dx^a)^2 = 0$. Similarly, the observer O' writes $c^2 dt'^2 - (dx'^a)^2 = 0$, with the same c, in accordance with the Michelson–Morley experiment. To simplify the notation, it is convenient to unify time and space coordinates into a single object $x^\mu = (x^0, x^a) = (ct, x^a)$, which labels points of space-time (the points of the four-dimensional space-time often called *events*). So, any quantity endowed with a Greek index has four components, $\mu = 0, 1, 2, 3$. We also introduce a 4×4-matrix

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{1.47}$$

Then the law of motion for the light front reads

$$\eta_{\mu\nu} dx^\mu dx^\nu = 0. \tag{1.48}$$

The left-hand side of this equation is similar to the expression for a distance in Euclidean space, $dl^2 = \delta_{ab} dx^a dx^b$, where, instead of the Euclidean metric δ_{ab}, the matrix $\eta_{\mu\nu}$ appears. The matrix $\eta_{\mu\nu}$ is called a *Minkowski metric*.

We look for a transformation $x'^\mu = x'^\mu(x^\nu)$ relating two frames. In the special relativity theory the equation of motion (1.48) is *postulated to be invariant*[3]

$$\eta_{\mu\nu} dx'^\mu(x) dx'^\nu(x) = \eta_{\alpha\beta} dx^\alpha dx^\beta. \tag{1.49}$$

Computing differentials on the l.h.s., it reads

$$\eta_{\mu\nu} \partial_\alpha x'^\mu \partial_\beta x'^\nu = \eta_{\alpha\beta}. \tag{1.50}$$

This equation implies linearity of the transformation law. To see this, compute the derivative with respect to x^γ, $\eta_{\mu\nu}\partial_\gamma(\partial_\alpha x'^\mu \partial_\beta x'^\nu) = 0$, then do a cyclic permutation of indexes in this expression and write the following combination $\eta_{\mu\nu}\partial_\gamma(\partial_\alpha x'^\mu \partial_\beta x'^\nu) - \eta_{\mu\nu}\partial_\alpha(\partial_\beta x'^\mu \partial_\gamma x'^\nu) + \eta_{\mu\nu}\partial_\beta(\partial_\gamma x'^\mu \partial_\alpha x'^\nu) = 0$. Computing the derivatives, this reduces to $2\eta_{\mu\nu}\partial_\alpha x'^\mu \partial_\beta \partial_\gamma x'^\nu = 0$. Since both $\eta_{\mu\nu}$ and $\partial_\alpha x'^\mu$ are invertible matrices, this equation implies $\partial_\beta \partial_\gamma x'^\nu = 0$. That is, x'^μ is at most a linear function of x^ν

[3] Requirement of the covariance, $\eta_{\mu\nu} dx'^\mu dx'^\nu = k(x)\eta_{\rho\delta} dx^\rho dx^\delta$, leads to a broader set of transformations known as a conformal group. In the theory with $k \neq 1$, transverse dimensions of a moving body experience contraction effects [17].

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu} + a^{\mu}. \tag{1.51}$$

Inserting this result into Eq. (1.50) we get a restriction on the matrix Λ

$$\eta_{\mu\nu} \Lambda^{\mu}{}_{\alpha} \Lambda^{\nu}{}_{\beta} = \eta_{\alpha\beta}, \quad \text{or} \quad \Lambda^{T} \eta \Lambda = \eta, \tag{1.52}$$

where$(\Lambda^{T})_{\alpha}{}^{\mu} \equiv \Lambda^{\mu}{}_{\alpha}$ (as always, the left index of any matrix element gives the number of column). The transformations (1.51) are known as *Poincaré transformations*, while those with $a^{\mu} = 0$ are called *Lorentz transformations*.

Thus we have obtained all the transformations that keep Eq. (1.49) invariant and hence are consistent with the independence of the speed of light on a choice of a reference frame. These are given by Eq. (1.51), where a^{μ}, $\mu = 0, 1, 2, 3$ are arbitrary constants while $\Lambda^{\mu}{}_{\nu}$ is a 4×4 matrix that obeys the restriction (1.52).

Equation (1.52) implies that $\det \Lambda = \pm 1$, so the Lorentz matrix is invertible. An inverse matrix is denoted as $\tilde{\Lambda}^{\mu}{}_{\nu}$. Multiplying (1.52) by η from the left and by $\tilde{\Lambda}$ from the right, it reads $\tilde{\Lambda} = \eta \Lambda^{T} \eta$. So the inverse matrix is

$$\tilde{\Lambda}^{\gamma}{}_{\mu} = \eta_{\gamma\beta} (\Lambda^{T})_{\beta}{}^{\nu} \eta_{\nu\mu}, \qquad \tilde{\Lambda}^{\gamma}{}_{\mu} \Lambda^{\mu}{}_{\alpha} = \delta^{\gamma}_{\alpha}. \tag{1.53}$$

Consider the quantity

$$x_{\mu} = \eta_{\mu\nu} x^{\nu} = (x^{0}, -x^{a}). \tag{1.54}$$

Using Eqs. (1.51) and (1.53), it transforms with the help of the inverse matrix

$$x'_{\mu} = x_{\nu} \tilde{\Lambda}^{\nu}{}_{\mu} + a_{\mu}. \tag{1.55}$$

Hence in Minkowski geometry the up and down position of the four-dimensional index of any quantity is fixed and indicates the transformation law of the quantity. Any index can be raised or lowered with the help of the Minkowski metric.

Vector (tensor) fields on Minkowski space. Manifestly covariant equations. Consider the set

$$\{v^{\mu}(x^{\nu}), v'^{\mu}(x'^{\nu}), \ldots\}, \tag{1.56}$$

where $v^{\mu}(x^{\nu})$ is a function representing a quantity in the frame $O(x^{\nu})$; $v'^{\mu}(x'^{\nu})$ represents this quantity in $O'(x'^{\nu})$ and so on. The set is called a *contravariant vector field* if

$$v'^{\mu}(x') = \Lambda^{\mu}{}_{\nu} v^{\nu}(x), \tag{1.57}$$

where the coordinates x'^{μ}, x^{μ} are related by the Poincaré transformation (1.51).

Similarly, a *covariant vector field* is defined by the transformation law

$$\omega'_{\mu}(x') = \omega_{\nu}(x) \tilde{\Lambda}^{\nu}{}_{\mu}. \tag{1.58}$$

To every contravariant quantity there corresponds a covariant one, which is obtained by lowering the up-index with help of the metric, and vice-versa. Quantities equipped with more than one index are called tensor fields. For example, the third-rank contravariant tensor is defined by the transformation law

$$s'^{\mu\nu\rho}(x') = \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}\Lambda^{\rho}{}_{\gamma}s^{\alpha\beta\gamma}(x), \tag{1.59}$$

while the second-rank covariant tensor transforms as

$$t'_{\mu\nu}(x') = s_{\alpha\beta}(x)\tilde{\Lambda}^{\alpha}{}_{\mu}\tilde{\Lambda}^{\beta}{}_{\nu}. \tag{1.60}$$

Finally, a quantity without indexes that transforms as

$$\varphi'(x') = \varphi(x), \tag{1.61}$$

is called a *scalar function*.

Equation (1.52) can be written as $\eta_{\mu\nu} = \eta_{\alpha\beta}\tilde{\Lambda}^{\alpha}{}_{\mu}\tilde{\Lambda}^{\beta}{}_{\nu}$, so the Minkowski metric represents a special example of an x-independent tensor field that has the same components in all reference frames.[4]

Examples

1. Note that tensors of a given rank form a linear space: if $v^{\mu\nu}$, $s^{\mu\nu}$ are tensors, then $av^{\mu\nu} + bs^{\mu\nu}$ is a tensor as well.
2. Examples of scalar functions can be obtained contracting tensor indexes

$$\omega_{\mu}v^{\mu}, \qquad \eta_{\mu\nu}v^{\mu}s^{\nu}, \qquad s_{\mu\nu\rho}\eta^{\mu\rho}v^{\nu}. \tag{1.62}$$

3. Contraction of indexes can also be used to construct new tensors. For example, if $t_{\mu\nu\rho}$ and v^{μ} are tensors, then $t_{\mu\nu\rho}v^{\mu}$ is a second-rank covariant tensor, while $t_{\mu\nu\rho}v^{\mu}v^{\rho}$ is a covariant vector.
4. Starting from the scalar function, let us construct the set

$$\left\{\frac{\partial\varphi(x)}{\partial x^{\mu}}, \frac{\partial\varphi'(x')}{\partial x'^{\mu}}, \ldots\right\}. \tag{1.63}$$

[4] Besides the metric, Minkowski space admits one more invariant constant tensor $\epsilon^{\mu\nu\rho\delta}$ called the *Levi-Civita tensor*. It is defined as follows. It is antisymmetric in any pair of indexes, and in the frame $O(x^{\mu})$ is determined by $\epsilon^{0123} = 1$. Let $O'(x'^{\mu})$ be related to $O(x^{\mu})$ by the Lorentz matrix Λ. If $\det\Lambda = 1$, then $\epsilon'^{0123} = 1$ in the frame O'. If $\det\Lambda = -1$, then $\epsilon'^{0123} = -1$ in the frame O'. Then the transformation law $\epsilon'^{\mu\nu\rho\delta} = \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}\Lambda^{\rho}{}_{\gamma}\Lambda^{\delta}{}_{\sigma}\epsilon^{\alpha\beta\gamma\sigma}$ holds as a consequence of the definition of a determinant.

From $\frac{\partial\varphi'(x'(x))}{\partial x^\mu} = \left.\frac{\partial\varphi'(x')}{\partial x'^\nu}\right|_{x'(x)} \frac{\partial x'^\nu}{\partial x^\mu} = \left.\frac{\partial\varphi'(x')}{\partial x'^\nu}\right|_{x'(x)} \Lambda^\nu{}_\mu$ we write the relationship between $\partial'_\mu\varphi'$ and $\partial_\mu\varphi$

$$\partial'_\mu\varphi'(x') = \partial_\nu\varphi(x)\tilde{\Lambda}^\nu{}_\mu, \tag{1.64}$$

that is, a derivative of a scalar function is a covariant vector. This example shows that it is convenient to adopt the vector transformation law for a partial derivative

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial}{\partial x'^\nu}\tilde{\Lambda}^\nu{}_\mu. \tag{1.65}$$

Similarly, a derivative of a tensor is a tensor as well. For example, $\partial_\mu\partial_\nu t_\rho$ represents a third-rank covariant tensor.

Equations (1.57), (1.58), (1.59), (1.60), and (1.61) clearly show that the scalar, vector and tensor quantities can be used for the construction of manifestly covariant equations: if $S^{\mu\nu\rho}$ is a tensor, then $S^{\mu\nu\rho} = 0$ represents a manifestly covariant equation.

Structure of a Poincaré transformation. Similarly to the Galilean transformation, the Poincaré one is specified by 10 parameters. So, let us compare them, to get an idea about the structure of Poincaré transformations.

1. There are four translations presented, $x'^\mu = x^\mu + a^\mu$, just as in the Galilean case, $t' = t + a$, $x'^a = x^a + c^a$. Six more parameters are necessary to label the Lorentz matrix $\Lambda^\mu{}_\nu$.
2. Consider the Lorentz transformation with $\Lambda^0{}_0 = 1$, and $\Lambda^i{}_0 = \Lambda^0{}_i = 0$, that is

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ 0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ 0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{pmatrix}. \tag{1.66}$$

   Then Eq. (1.52) acquires the form $\Lambda^k{}_i\Lambda^k{}_j = \delta_{ij}$, that is, the 3×3-block $\Lambda^i{}_j$ is the orthogonal matrix. The corresponding transformation represents a spacial rotation

$$x'^0 = x^0, \qquad x'^i = \Lambda^i{}_j x^j. \tag{1.67}$$

3. Matrices with $\Lambda^0{}_i, \Lambda^i{}_0 \neq 0$ produce a transformation that can be compared with the Galilean boost (which is written on the r.h.s. below)

$$x'^0 = \Lambda^0{}_0 x^0 + \Lambda^0{}_b x^b, \qquad (t' = t); \tag{1.68}$$
$$x'^a = \Lambda^a{}_b x^b + \Lambda^a{}_0 x^0, \qquad (x'^a = x^a + V^a t). \tag{1.69}$$

The spacial part, Eq. (1.69), is similar to the Galilean boost, but it is accompanied by the time boost, Eq. (1.68). This ultimately implies that both spacial and time *intervals* between two events look different for observers in relative motion. It is this part of the Poincaré group which differ from the Galilean one.

Structure of the Lorentz transformation. To study Lorentz boosts in detail, we parameterize an arbitrary Lorentz transformation

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu, \qquad \eta_{\mu\nu}\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta = \eta_{\alpha\beta}, \tag{1.70}$$

using the same procedure that was used in the previous section for rotation matrices. We try to represent the Lorentz transformation Λ in the form of the exponential of another matrix: $\Lambda = e^\omega$. The condition $\Lambda^T\eta\Lambda = \eta$ can be written as $\eta e^{\omega^T}\eta = e^{-\omega}$ and then as $e^{\eta\omega^T\eta} = e^{-\omega}$. This implies $\eta\omega^T\eta = -\omega$, or, finally

$$(\eta\omega)^T = -\eta\omega. \tag{1.71}$$

So, e^ω will be the Lorentz matrix if and only if $\eta\omega$ is antisymmetric. Any ω that obeys (1.71) is specified by 6 parameters and can be written in the form (confirm this!)

$$\omega = \begin{pmatrix} 0 & \omega^{01} & \omega^{02} & \omega^{03} \\ \omega^{01} & 0 & \omega^{12} & \omega^{13} \\ \omega^{02} & -\omega^{12} & 0 & \omega^{23} \\ \omega^{03} & -\omega^{13} & -\omega^{23} & 0 \end{pmatrix} =$$
$$\omega^{01}M_{01} + \omega^{02}M_{02} + \omega^{03}M_{03} + \omega^{12}M_{12} + \omega^{13}M_{13} + \omega^{23}M_{23}, \tag{1.72}$$

where $(M_{0a})^\mu{}_\nu = \delta^\mu_0\delta_{\nu a} + \delta^\mu_a\delta_{\nu 0}$ are the *Lorentz-boost generators*, and $(M_{ab})^\mu{}_\nu = \delta^\mu_a\delta_{\nu b} - \delta^\mu_b\delta_{\nu a}$ are the rotation generators (it is convenient to label them by a pair of indexes). The manifest form of the generators is as follows:

$$M_{01} = \begin{pmatrix} 0&1&0&0 \\ 1&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}, M_{02} = \begin{pmatrix} 0&0&1&0 \\ 0&0&0&0 \\ 1&0&0&0 \\ 0&0&0&0 \end{pmatrix},$$
$$M_{03} = \begin{pmatrix} 0&0&0&1 \\ 0&0&0&0 \\ 0&0&0&0 \\ 1&0&0&0 \end{pmatrix}; \tag{1.73}$$

$$M_{12}=\begin{pmatrix}0&0&0&0\\0&0&1&0\\0&-1&0&0\\0&0&0&0\end{pmatrix},\ M_{13}=\begin{pmatrix}0&0&0&0\\0&0&0&1\\0&0&0&0\\0&-1&0&0\end{pmatrix},$$

$$M_{23}=\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&1\\0&0&-1&0\end{pmatrix}; \qquad (1.74)$$

Equation (1.72) can be written as $\omega = \sum_{\alpha<\beta} \omega^{\alpha\beta} M_{\alpha\beta}$, where the sum is over all values of α, β such that $\alpha < \beta$. It is convenient to duplicate a number of parameters and generators defining $\omega^{10} \equiv -\omega^{01}$, $M_{10} \equiv -M_{01}$, and so on. Then (1.72) reads $\omega = \frac{1}{2}\omega^{\alpha\beta} M_{\alpha\beta}$, where the sum is now performed over all values of α, β. The parameters $\omega^{\alpha\beta}$ in this expression form an antisymmetric matrix.

Returning to the Lorentz transformation, it can now be presented as

$$\Lambda = e^{\frac{1}{2}\omega^{\alpha\beta} M_{\alpha\beta}} \approx \mathbf{1} + \frac{1}{2}\omega^{\alpha\beta} M_{\alpha\beta}, \qquad \alpha, \beta = 0, 1, 2, 3. \qquad (1.75)$$

According to the previous section, the transformations $e^{\omega^{23} M_{23}}$, $e^{\omega^{31} M_{31}}$ and $e^{\omega^{12} M_{12}}$ produce spacial rotations about the axis x^1, x^2 and x^3.

Interpretation of the Lorentz boost. Finally we are ready to discuss Lorentz boosts. Consider, for example

$$\Lambda = e^{\alpha M_{01}}, \qquad \alpha \equiv \omega^{01}. \qquad (1.76)$$

To find its manifest form, note that $(M_{01})^{2n}$ has only non-zero matrix elements $(M_{01})^0{}_0 = 1$, $(M_{01})^1{}_1 = 1$, while $(M_{01})^{2n+1} = M_{01}$. Then we compute

$$\Lambda^0{}_0 = \Lambda^1{}_1 = 1 + \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} + \ldots = \frac{1}{2}(e^{\alpha} + e^{-\alpha}) = \cosh\alpha,$$
$$\Lambda^0{}_1 = \Lambda^1{}_0 = \alpha + \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} + \ldots = \frac{1}{2}(e^{\alpha} - e^{-\alpha}) = \sinh\alpha, \qquad (1.77)$$

and the Lorentz-boost matrix is

$$\Lambda = e^{\alpha M_{01}} = \begin{pmatrix}\cosh\alpha & \sinh\alpha & 0 & 0\\ \sinh\alpha & \cosh\alpha & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}. \qquad (1.78)$$

This produces the transformation ($x^0 = ct$, $x'^0 = ct'$)

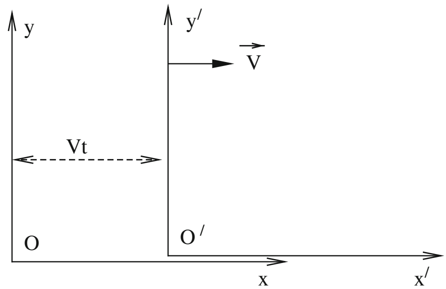

Fig. 1.4 Both Galilean and Lorentz boosts prescribe a law of transformation between observers in relative motion

$$
\begin{aligned}
t' &= t\cosh\alpha + \frac{x^1}{c}\sinh\alpha,\\
x'^1 &= tc\sinh\alpha + x^1\cosh\alpha,\\
x'^2 &= x^2, \qquad x'^3 = x^3,
\end{aligned}
\tag{1.79}
$$

In the special theory of relativity it is *postulated* that this transformation relates the coordinates of observer O', who moves along the x^1 axis of O at speed V, passing through its origin at $t = 0$; see Fig. 1.4 on page 20. Then its origin has the coordinate x^1 at the instant $t = \frac{x^1}{V}$. Coordinates of this event in the frame O' are t', $x'^1 = 0$. Using these values in the second equation from (1.79), we obtain

$$
\tanh\alpha = -\frac{V}{c}, \quad \text{then} \quad \begin{cases} \cosh\alpha = \frac{1}{\sqrt{1-\tanh^2\alpha}} = \frac{1}{\sqrt{1-V^2/c^2}}, \\ \sinh\alpha = \tanh\alpha\cosh\alpha = -\frac{V/c}{\sqrt{1-V^2/c^2}}, \end{cases} \tag{1.80}
$$

So, the final form of the Lorentz boost is

$$
t' = \frac{t - \frac{V}{c^2}x^1}{\sqrt{1-V^2/c^2}}, \quad x'^1 = \frac{x^1 - Vt}{\sqrt{1-V^2/c^2}}, \quad x'^2 = x^2, \;\; x'^3 = x^3. \tag{1.81}
$$

The expressions can be inverted (do this!), with the result being

$$
t = \frac{t' + \frac{V}{c^2}x'^1}{\sqrt{1-V^2/c^2}}, \quad x^1 = \frac{x'^1 + Vt'}{\sqrt{1-V^2/c^2}}, \quad x^2 = x'^2, \;\; x^3 = x'^3. \tag{1.82}
$$

Comments

1. Note that at the limit $c \to \infty$ this reduces to the Galilean boost. Besides, when $V > c$, the transformations have no sense.
2. Equation (1.81) implies the following transformation rule for velocity

$$
v' = \frac{dx'}{dt'} = \frac{dx - V dt}{dt - \frac{V}{c^2}dx} = \frac{v - V}{1 - \frac{vV}{c^2}}. \tag{1.83}
$$

Then $v = c$ implies $v' = c$, as should be the case.

3. The time *interval* between two events looks different for observers in relative motion. Consider two events (for example, two flashes) that happen at the same point x'^1 of O' at the instants t'_1 and t'_2. That is, the O'-clock registers the interval $\triangle t' = t'_2 - t'_1$. From Eq. (1.82) we have

$$t_1 = \frac{t'_1 + \frac{V}{c^2} x'^1}{\sqrt{1 - V^2/c^2}}, \quad t_2 = \frac{t'_2 + \frac{V}{c^2} x'^1}{\sqrt{1 - V^2/c^2}}. \tag{1.84}$$

So the O-clock will register the interval

$$\triangle t = \frac{\triangle t'}{\sqrt{1 - V^2/c^2}} > \triangle t', \tag{1.85}$$

that is, a clock moving with respect to O ticks more slowly as compared with a clock at rest.

4. The spacial *interval* between two events looks different for observers in relative motion. Take a pivot at rest with respect to O' and placed along its x'^1 axis. Its length in O' is $l' = x'^1_2 - x'^1_1$. To find its length in O, one needs to compute the coordinates of the ends of the pivot at the *same instant* t. Use Eq. (1.81)

$$l' = x'^1_2 - x'^1_1 = \frac{l}{\sqrt{1 - V^2/c^2}} > l = x^1_2 - x^1_1, \tag{1.86}$$

that is, a pivot moving with respect to O looks shorter as compared with the pivot at rest. Material bodies experience *contraction* in the direction of motion.

5. General Lorentz transformation can be presented as a product of two rotations and the Lorentz boost (1.81) in the direction of motion of observer O'

$$\Lambda = \Lambda_{R(\beta)} \times \Lambda_V \times \Lambda_{R(\alpha)}, \tag{1.87}$$

see Fig. 1.5 on page 21.

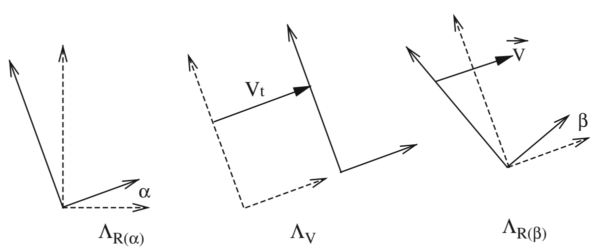

Fig. 1.5 General Lorentz transformation can be decomposed into two rotations and a boost, $\Lambda = \Lambda_{R(\beta)} \times \Lambda_V \times \Lambda_{R(\alpha)}$

Mechanical covariants and invariants. Poincaré transformations of the Minkowski space coordinates are linear, $x'^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu} + a^{\mu}$. They induce a certain transformation law of the functions $x^a(t)$. The functions $x^a(t)$ are necessary to describe the evolution of a system. The problem is that the induced transformations are higher nonlinear ones (see Sect. 7.4). While a relativistic action and equations of motion can be formulated in terms of $x^a(t)$, the Poincaré covariance of such a formulation is not under control. The conventional way to avoid the problem is as follows.

Let us write the Minkowski-space curve $x^a(t)$ in a parametric form

$$x^0 = x^0(\tau), \qquad x^a = x^a(\tau), \tag{1.88}$$

where τ is a parameter along the curve. In contrast to $x^a(t)$, the functions $x^{\mu}(\tau)$ transform as coordinates, $x'^{\mu}(\tau) = \Lambda^{\mu}{}_{\nu}x^{\nu}(\tau) + a^{\mu}$. So the quantity $\dot{x}^{\mu}(\tau) \equiv \frac{dx^{\mu}}{d\tau}$ transforms homogeneously, $\dot{x}'^{\mu} = \Lambda^{\mu}{}_{\nu}\dot{x}^{\nu}$, that is, it represents a contravariant vector.

To construct the *manifestly* covariant quantities we have now the building blocks

$$\eta_{\mu\nu}, \qquad \dot{x}^{\mu}(\tau), \qquad \frac{d}{d\tau}, \qquad \epsilon_{\mu\nu\rho\lambda}. \tag{1.89}$$

So

$$\begin{gathered} \dot{x}^{\mu}(\tau) = 0, \quad \dot{x}_{\mu} \equiv \eta_{\mu\nu}\dot{x}^{\nu} = 0, \quad \ddot{x}^{\mu} = 0, \\ (f(\dot{x}_{\nu}\dot{x}^{\nu})\dot{x}^{\mu})\dot{} = 0, \quad \epsilon_{\mu\nu\rho\lambda}\dot{x}^{\rho}\ddot{x}^{\lambda} = 0, \end{gathered} \tag{1.90}$$

are examples of the manifestly Poincaré-covariant equations, while

$$\dot{x}_{\mu}\dot{x}^{\mu} \equiv \eta_{\mu\nu}\dot{x}^{\mu}\dot{x}^{\nu}, \quad f(\dot{x}_{\mu}\dot{x}^{\mu}), \quad \epsilon_{\mu\nu\rho\lambda}\dot{x}^{\mu}\ddot{x}^{\nu}\overset{(3)}{x}{}^{\rho}\overset{(4)}{x}{}^{\lambda}, \tag{1.91}$$

represent invariants (scalar functions). Note that $f(\dot{x}_{\mu}\dot{x}^{\mu})$ is the only invariant that does not involve higher derivatives.

It is important to note that the functions $x^{\mu}(\tau)$ have no direct physical meaning. The observer O measures space-time coordinates t, x^a, that is, he deals with the functions $x^a(t)$, not with $x^{\mu}(\tau)$. By construction, they are related by $x^a(t) = x^a(\tau)|_{\tau(t)}$, where $\tau(t)$ is the solution to the equation $x^0 = x^0(\tau)$.

Proper time parametrization. Natural parametrization along a curve can be constructed as follows. To put this in concrete terms, suppose that the curve $x^a(t)$ passes through the origin at $t = 0$, $x^a(0) = 0$. The interval between the events $(0, x^a(0))$ and $(t, x^a(t))$ can be used to construct the function

$$\tau(t) = \frac{1}{c}\sqrt{x_{\mu}x^{\mu}}. \tag{1.92}$$

As we know, the interval $x_\mu x^\mu$ is invariant under the Poincaré transformations. Besides, $\tau(t)$ is an increasing function[5] of t. So it can be taken as a parameter along the curve.

This has a simple interpretation for the particle moving with constant velocity. Choosing the frame that moves with the particle, $\frac{dx^a}{dt} = 0$, we obtain from (1.92) $\tau = t$. Hence τ is equal to the *proper time* of the particle, that is the time measured in the frame where the particle is at rest.

1.4 Principle of Least Action

We have seen that the sum of kinetic and potential energy is an important quantity characterizing a physical system. Surprisingly enough, their difference, $T - U$, also plays a very special role, being a basic quantity used in the formulation of the least action principle. In modern mechanics this underlies equations of motion.

Construction. A mechanical system with the position variables q^a can be characterized by the *function of Lagrange* $L(q^a, \dot{q}^a, t)$, which in many cases is given by the difference $L = T - U$. We can compute the mean value[6] of the Lagrangian along any curve $q^a(t)$ at a fixed interval $[t_1, t_2]$

$$S[q] = \int_{t_1}^{t_2} dt L(q, \dot{q}, t). \tag{1.93}$$

This integral is called the *Lagrangian action* of the system. It associates the number $S[q]$ to each curve $q^a(t)$. We can compare the numbers $S[q]$ calculated for the curves that join two given points over the time interval chosen; see Fig. 1.6 on page 25. We restrict ourselves to this class of curves.

According to the *principle of least action* (or Hamilton's principle), with a mechanical system can be associated a Lagrangian function L. Then the system moves between two given points along the curve that provides the minimum of the Lagrangian action.

This statement of the existence of a Lagrangian is rather general. In particular, for a mechanical system that obeys the Hamiltonian equations of motion (see below), the existence of a Lagrangian function can be proved. While in some cases it requires the use of rather sophisticated methods, Lagrangians have been found for most fundamental equations of mathematical physics. Hence they follow the principle of least action.

To provide the minimum, $q^a(t)$ must obey certain second-order differential equations constructed in terms of L. While we discuss this in the next section, it may be instructive to carry out an intuitive computation that illustrates a relationship

---

[5] Its derivative is $\frac{ds}{dt} = 2s^{-1}(c^2 t - x^a \frac{dx^a}{dt})$. For a massive particle $v(t) < c$ for any t. This implies $\frac{dx^a}{dt} < c$ and $x^a(t) < ct$ for any a, so $\frac{ds}{dt} > 0$.

[6] The mean value would be S divided by the interval $T = t_2 - t_1$. Since we are interested only in curves over the fixed interval, the common factor T^{-1} can be omitted.

between the quantity L and equations of motion. Take the harmonic oscillator with $L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2$. Consider the curve $x(t)$ and another one, close to it, $x(t) + \delta x(t)$. Let us compute the variation of the function L in a linear order on δx (that is, the differential) at a fixed t; see Fig. 1.6 on page 25. We obtain

$$dL = \frac{1}{2}\left[m(\dot{x}+\delta\dot{x})^2 - k(x+\delta x)^2 - m\dot{x}^2 - kx^2\right]\Big|_{O(\delta x)}$$
$$= -(m\ddot{x} + kx)\delta x + (x\delta x)^{\dot{}}. \tag{1.94}$$

That is the differential L consists of equations of motion plus a total derivative term. This can be removed if we integrate this equality on the interval $[t_1, t_2]$. The integral of the total derivative $\int dt(x\delta x)^{\dot{}} = x\delta x|_{t_1}^{t_2}$ vanishes due to the conditions $\delta x(t_1) = \delta x(t_2) = 0$. Then the integral of Eq. (1.94) reads

$$\int dt dL = -\int dt(m\ddot{x} + kx)\delta x. \tag{1.95}$$

Hence the differential of the Lagrangian integrated on the interval vanishes for the curve $x(t)$, which obeys the equations of motion, $\int dt dL(x(t)) = 0$. This equality represents a condition of extremum of the Lagrangian action; see the next section.

1.5 Variational Analysis

Definition of a functional. Consider a set of functions $C = \{q^a(\tau),\ q : \mathbb{R} \to \mathbb{R}^n\}$. The functional S is a rule that associates a real number with any function of the set, that is, S is the map

$$S : C \to \mathbb{R}, \quad \text{or} \quad S : q^a(\tau) \to S[q^a(\tau)] \in \mathbb{R}. \tag{1.96}$$

The straight brackets are used to distinguish functionals from functions.

In Lagrangian mechanics we are interested in a functional of a special form called the Lagrangian action functional. A mechanical system can be characterized by the function of Lagrange $L(q, \dot{q}, \tau)$. Given the trajectory $q^a(\tau)$, $\tau \in [\tau_1, \tau_2]$, the *Lagrangian action functional* is defined by the rule

$$S[q] = \int_{\tau_1}^{\tau_2} d\tau L(q, \dot{q}, \tau). \tag{1.97}$$

Below we systematically omit the integration limits as well as the time variable τ in the Lagrangian.

The variational problem for the Lagrangian action is formulated as follows. Consider the class of curves joining two fixed points q_1, q_2 over a fixed time interval $[\tau_1, \tau_2]$

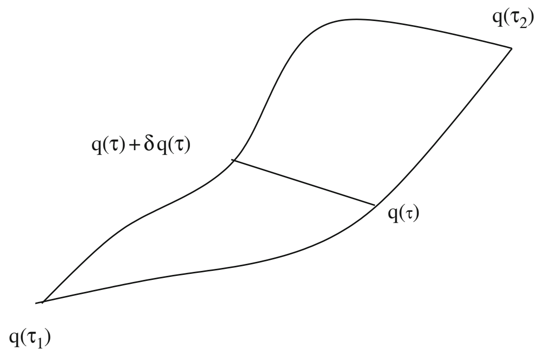

Fig. 1.6 The variational problem is formulated on a class of curves with the same initial and final points

$$q^a(\tau_1) = q_1^a, \qquad q^a(\tau_2) = q_2^a, \tag{1.98}$$

see Fig. 1.6 on page 25. The problem is to find the curve that provides the smallest value of the action functional. According to the least action principle, this is the curve the system with the Lagrangian L chooses as a trajectory of motion.

To analyze the problem, we define the functional analog for the differential of a function called the variation of a functional. Given functions $q^a(\tau)$ and $y^a(\tau)$ of the set C, let us introduce the function $\delta q^a(\tau) \equiv q^a(\tau) - y^a(\tau)$. Then we write $y^a = q^a + \delta q^a$. Let us also introduce the one-parameter family of curves connecting $q^a(\tau)$ and $q^a(\tau) + \delta q^a(\tau)$; see Fig. 1.7 on page 25. For instance, we can take $q^a(\tau, s) \equiv q^a(\tau) + s\delta q^a(\tau)$, $s \in [0, 1]$. It obeys the properties $q^a(\tau, 0) = q^a(\tau)$, $q^a(\tau, 1) = q^a(\tau) + \delta q^a(\tau)$.

With the functional S we associate the usual function[7] of the variable s

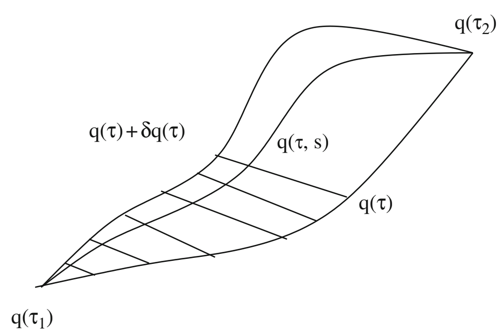

Fig. 1.7 A one-parameter family $q(\tau, s)$ connecting the curves $q(\tau)$ and $q(\tau) + \delta q(\tau)$

[7] Of course $S(s)$ depends also on the choice of $q(\tau)$ and $\delta q(\tau)$.

$$S(s) \equiv S[q^a + s\delta q^a], s \epsilon[-1, 1]. \tag{1.99}$$

Then the *variation δS of the functional S* "at the point $q^a(\tau)$" is defined by the formula

$$\delta S[q] \equiv \left.\frac{dS(s)}{ds}\right|_{s=0} = \left.\frac{d}{ds}S[q+s\delta q]\right|_{s=0}. \tag{1.100}$$

For example, for the velocity-independent functional, $S = \int d\tau L(q)$, we obtain

$$\delta S = \int d\tau \frac{\partial L}{\partial q^a}\delta q^a, \tag{1.101}$$

that is, the integrand represents the usual differential of the function L.

For the Lagrangian action (1.97) the variation (1.100) reads

$$\delta S = \int d\tau \left(\frac{\partial L(q,\dot{q})}{\partial q^a}\delta q^a + \frac{\partial L(q,\dot{q})}{\partial \dot{q}^a}(\delta q^a)^{\cdot}\right). \tag{1.102}$$

Here $\frac{\partial L(q,\dot{q})}{\partial \dot{q}^a}$ stands for a partial derivative with respect to the symbol $\dot{q}^a$. Notice that the integrand looks like the linear term of the power expansion of the function $L(q+\delta q, \dot{q}+(\delta q)^{\cdot})$, if this is treated as a function of *independent* variables q and $\dot{q}$. It is the formal rule used for computation of the variation in practice (omitting the variable s)

$$\delta S = \left(S[q+\delta q] - S[q]\right)|_{\text{linear in}\,\delta\text{q term}} \tag{1.103}$$

This rule has been used in the previous section. Using integration by parts, Eq. (1.102) can be written in the form

$$\delta S = \int d\tau \left(\frac{\partial L}{\partial q^a} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{q}^a}\right)\delta q^a + \left.\frac{\partial L}{\partial \dot{q}^a}\delta q^a\right|_{\tau_1}^{\tau_2}. \tag{1.104}$$

If the variation is computed on the class of functions (1.98), δq^a vanishes at the limiting points, $\delta q^a(\tau_1) = \delta q^a(\tau_2) = 0$. So the last term in (1.104) vanishes as well, and we obtain the *basic formula of variational analysis*

$$\delta S[q] = \int d\tau \left(\frac{\partial L(q,\dot{q})}{\partial q^a} - \frac{d}{d\tau}\frac{\partial L(q,\dot{q})}{\partial \dot{q}^a}\right)\delta q^a. \tag{1.105}$$

We are now ready to discuss the central result that interests us.

Assertion If $q^a(\tau)$ represents an extremum (the maximum or minimum) of the variational problem (1.97) and (1.98), then

$$\delta S[q] = 0 \qquad \text{for any} \quad \delta q^a(\tau). \tag{1.106}$$

Indeed, let q^a be an extremum. Choose some $\delta q^a(\tau)$ and consider the function (1.99). Since $S(0) = S[q]$, the function $S(s)$ has an extremum at $s = 0$. From mathematical analysis, this implies $\left.\frac{dS(s)}{ds}\right|_{s=0} = 0$. Taking into account (1.100), we conclude that $\delta S[q] = 0$.

Eq. (1.105) implies that the extremum $q^a(\tau)$ obeys the condition

$$\int d\tau \left(\frac{\partial L}{\partial q^a} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{q}^a} \right) \delta q^a = 0 \qquad \text{for any} \quad \delta q^a(\tau), \tag{1.107}$$

which implies (this is proved at the end of the section)

$$\frac{\delta S}{\delta q^a} \equiv \frac{d}{d\tau}\frac{\partial L}{\partial \dot{q}^a} - \frac{\partial L}{\partial q^a} = 0. \tag{1.108}$$

This system of ordinary second-order differential equations is known as *Lagrangian equations* or *Euler-Lagrange equations*.

Let us summarize the results. According to the least action principle, the true trajectory of a physical system is an extremum of the variational problem with a properly chosen Lagrangian. In turn, the extremum $q^a(\tau)$ obeys the Lagrangian equations (1.108). Combining these statements we arrive at

Assertion The time evolution of a system described by the Lagrangian $L(q, \dot{q}, \tau)$ is governed by the Lagrangian equations (1.108).

We point out an ambiguity presented in the construction of a Lagrangian. We can modify a given Lagrangian adding the total derivative of any function $N(q, \tau)$. The action

$$S' = \int d\tau \left[L(q, \dot{q}, \tau) + \frac{d}{d\tau} N(q, \tau) \right], \tag{1.109}$$

leads to the same equations of motion as (1.97). To confirm this, note that variation (1.100) of the extra term vanishes due to the boundary conditions $\delta q(\tau_2) = \delta q(\tau_1) = 0$, $\delta \int \dot{N} = \frac{\partial N}{\partial q(\tau_2)}\delta q(\tau_2) - \frac{\partial N}{\partial q(\tau_1)}\delta q(\tau_1) = 0$. Hence $\delta S' = \delta S$.

Examples

1. Let us confirm the validity of the least action principle for a system of particles subject to potential forces. Let $\mathbf{r}_i = (x_i^1, x_i^2, x_i^3)$ stand for the position vector of the i-particle of mass m_i, $i = 1, 2, \ldots, N$, and the function $U(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ describe the potential energy. Take a Lagrangian as kinetic minus the potential energy and write the corresponding action

$$S = \int dt \left[\frac{1}{2} m_i (\dot{\mathbf{r}}_i)^2 - U(\mathbf{r}_i) \right], \tag{1.110}$$

In this case the Lagrangian equations (1.108) reduce to the second law of Newton, as it should be:

$$m_i \ddot{\mathbf{r}}_{\mathbf{i}} = -\frac{\partial U}{\partial \mathbf{r}_i}, \qquad i = 1, 2, \ldots, N. \tag{1.111}$$

2. We specify the previous result for the case of N particles with the same mass m connected by massless springs, see Fig. 1.8 on page 28. All the springs are of the same length $\triangle x$ and rigidity k. Take the displacement $q_i(t) = x_i(t) - (i-1)\triangle x$ of the i-particle from the position of equilibrium as its coordinate. The potential energy of the system is a sum of the energies of the springs, $\sum \frac{1}{2}k(q_{i+1} - q_i)^2$, so the action is given by

$$S = \frac{1}{2}\int dt \left[m\sum_{i=1}^{N}(\dot{q}_i)^2 - k\sum_{i=1}^{N-1}(q_{i+1} - q_i)^2 \right]. \tag{1.112}$$

This implies the equations

$$\begin{gathered} m\ddot{q}_1 = k(q_2 - q_1), \qquad m\ddot{q}_N = k(q_N - q_{N-1}), \\ m\ddot{q}_i = k(q_{i+1} - q_i) - k(q_i - q_{i-1}), \qquad i = 2, 3, \ldots, N-1. \end{gathered} \tag{1.113}$$

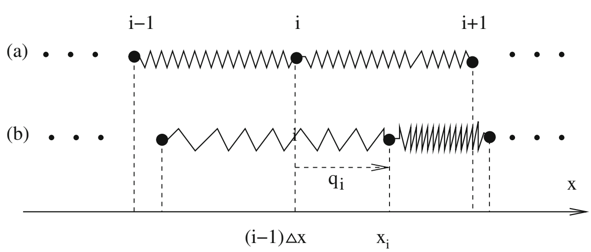

Fig. 1.8 Chain of springs. **(a)**-equilibrium configuration, **(b)**-instantaneous configuration

Non-singular and singular theories. Computing the derivative with respect to τ in the Lagrangian equations (1.108), these can be written as

$$M_{ab}(q, \dot{q})\ddot{q}^b = K_a(q, \dot{q}), \tag{1.114}$$

where

$$M_{ab} \equiv \frac{\partial^2 L(q, \dot{q})}{\partial \dot{q}^a \partial \dot{q}^b}, \qquad K_a \equiv \frac{\partial L}{\partial q^a} - \frac{\partial^2 L}{\partial \dot{q}^a \partial q^b}\dot{q}^b. \tag{1.115}$$

Exercise
Multiplying (1.114) by $\dot{q}^a$, obtain the law of conservation of total energy.

Lagrangian theories are classified according to the properties of the *Hessian matrix* M which appears in front of the second-derivative terms. The theory is called *non-singular* if $\det M \neq 0$ (then there is an inverse matrix denoted $\tilde{M}$). Otherwise, it is called a *singular theory*. Equations of the non-singular theory can be rewritten in the *normal form*

$$\ddot{q}^a = \tilde{M}^{ab}(q, \dot{q}) K_b(q, \dot{q}). \tag{1.116}$$

As has been mentioned above, this implies causal dynamics, as should be the case for a classical-mechanics system. So, in classical mechanics we usually deal with non-singular systems (while a system with holonomic constraints can be formulated in terms of a singular Lagrangian as well, see Sect. 1.10). Singularity is a characteristic property of relativistic theories. The formalism of classical mechanics for singular theories is discussed in some detail in Chap. 8.

To conclude this section, we demonstrate the statement used in the passage from (1.107) to (1.108):

Let

$$\int_{\tau_1}^{\tau_2} d\tau f(q, \dot{q}, \ddot{q}, \ldots)\eta(\tau) = 0, \qquad \text{for any function } \eta(\tau). \tag{1.117}$$

Then $f = 0$.

Indeed, by *reductio ad absurdum* suppose that $f > 0$ on an interval $[\tau', \tau'']$. Take any function $\eta(\tau)$ that has positive values on that interval and vanishes at all other points. Then by construction

$$\int_{\tau_1}^{\tau_2} d\tau f\eta = \int_{\tau'}^{\tau''} d\tau f\eta > 0, \tag{1.118}$$

which contradicts (1.117).

1.6 Generalized Coordinates, Coordinate Transformations and Symmetries of an Action

A change of variables (coordinates) is one of the powerful methods used for studying equations of motion. In this section we show that it can be carried out directly in an action functional instead of in equations of motion.

We recall that an action functional is an operation defined on a given class of functions. Strictly speaking, it does not imply such a thing as "coordinate transformation". For this reason our discussion here will be rather intuitive. A more consistent treatment of the subject will be given in Chap. 7.

Configuration space and generalized coordinates. To put this in concrete terms, consider a system of particles already discussed in Example 1 of the previous section

$$S = \int d\tau \left[\frac{1}{2} m_i (\dot{\mathbf{r}}_i)^2 - U'(\mathbf{r}_i) \right]. \tag{1.119}$$

First, we simplify the notation introducing the quantity $\mathbf{q} = q^a$, $a = 1, 2, \ldots, 3N \equiv n$

$$\mathbf{q} = \begin{pmatrix} m_1 \mathbf{r}_1 \\ m_2 \mathbf{r}_2 \\ \ldots \\ m_N \mathbf{r}_N \end{pmatrix}. \tag{1.120}$$

This can be regarded as the position vector of a point in n dimensional Euclidean space. In this notation the action reads

$$S = \int \left[\frac{1}{2} (\dot{q}^a)^2 - U(q^1, q^2, \ldots, q^n) \right]. \tag{1.121}$$

This implies an elegant mathematical reinterpretation of the initial problem with N particles: Eq. (1.121) looks like an action describing the motion of a *unique* particle in abstract $3N$ dimensional space. It is called the *configuration space* of the system, while q^a are called *generalized coordinates* (more generally, they are any coordinate system of the configuration space).

Remember that Euclidean space has natural metrical properties, that is, we are able to define the distance between points, the length of a curve and a vector, angles between vectors and so on. For two points with coordinates q^a_1, q^a_2 the distance is

$$(\triangle s)^2 = \delta_{ab} \triangle q^a \triangle q^b. \tag{1.122}$$

where δ_{ab} is a unit matrix sometimes called the *Euclidean metric* and $\triangle q^a$ stands for the difference of coordinates, $\triangle q^a = q^a_2 - q^a_1$. The metric determines the length of a vector as well

$$|\vec{w}|^2 = \delta_{ab} w^a w^b. \tag{1.123}$$

So the kinetic term in Eq. (1.121) is just half of the square of the velocity vector. For the later use we also recall the formula for the length of a curve with the parametric equation $q^a = q^a(t)$

$$l = \int_{t_1}^{t_2} dt \sqrt{\delta_{ab} \dot{q}^a \dot{q}^b}. \tag{1.124}$$

Transformation of an action under a change of coordinates. Let q'^a be another parametrization of the configuration space. Consider the change of coordinates, $q^a \to q'^a$, and let

$$q^a = q^a(q'^b), \tag{1.125}$$

stand for the transition function. A change of coordinates is supposed to be an invertible transformation. This implies det $\frac{\partial q^a}{\partial q'^b} \neq 0$, that is, $\frac{\partial q^a}{\partial q'^b}$ is an invertible matrix. Substitution of the transition function into the initial Lagrangian function L gives another function L'

$$L'(q'^a, \dot{q}'^a) \equiv L\left(q^a(q'^b), \frac{\partial q^a}{\partial q'^b}\dot{q}'^b\right). \tag{1.126}$$

We confirm that the corresponding action

$$S'[q'] = \int d\tau L'(q'^a, \dot{q}'^a), \tag{1.127}$$

leads to equations of motion that are equivalent to the initial ones. So it is matter of convenience which one is used for obtaining equations of motion. To prove this, write the Lagrangian equations following from (1.127)

$$\begin{gathered}\frac{d}{d\tau}\frac{\partial L'}{\partial \dot{q}'^a} - \frac{\partial L'}{\partial q'^a} = \\ \frac{d}{d\tau}\left(\left.\frac{\partial L}{\partial \dot{q}^b}\right|_{q(q')}\frac{\partial q^b}{\partial q'^a}\right) - \left.\frac{\partial L}{\partial q^b}\right|_{q(q')}\frac{\partial q^b}{\partial q'^a} - \left.\frac{\partial L}{\partial \dot{q}^c}\right|_{q(q')}\frac{\partial^2 q^c}{\partial q'^a \partial q'^b}\dot{q}'^b = \\ \frac{\partial q^b}{\partial q'^a}\left.\left(\frac{d}{d\tau}\frac{\partial L}{\partial \dot{q}^a} - \frac{\partial L}{\partial q^a}\right)\right|_{q(q')},\end{gathered} \tag{1.128}$$

that is

$$\frac{\delta S'}{\delta q'^a} = \frac{\partial q^b}{\partial q'^a}\left.\frac{\delta S}{\delta q^b}\right|_{q(q')}. \tag{1.129}$$

Since $\frac{\partial q^b}{\partial q'^a}$ is an invertible matrix, $\frac{\delta S'}{\delta q'^a} = 0$ implies $\frac{\delta S}{\delta q^b} = 0$ and vice versa. To avoid a confusion, note that Eq. (1.129) is different from the invariance condition (1.32).

In particular, replacing the Cartesian coordinates q^a in Eq. (1.121) by some q'^a, the action acquires the form

$$S' = \int d\tau \left[\frac{1}{2} g_{ab}(q') \dot{q}'^a \dot{q}'^b - U(q'^a) \right],$$
$$g_{ab} \equiv \sum_c \frac{\partial q^c}{\partial q'^a} \frac{\partial q^c}{\partial q'^b}, \quad U(q') \equiv U(q(q')). \tag{1.130}$$

Equation (1.130) shows that equations of motion in non-Cartesian coordinates generally do not have the Newton form (1.111). In particular, they may contain velocity-dependent terms; see Eq. (1.141) below.

Metric of configuration space. The matrix $g_{ab}(q')$ which appears in the action has a simple geometric interpretation.[8] To see this, let us consider two infinitesimal points of configuration space with coordinates q'^a and $q'^a + dq'^a$. Using the transition function (1.125), their coordinates in the Cartesian system read $q^a(q')$ and $q^a(q' + dq')$. Let us compute the expression for the distance (1.122) in coordinates q'. Expanding the transition function $q^a(q' + dq')$ in a power series up to the first order we obtain

$$(ds)^2 = \delta_{ab} \triangle q^a \triangle q^b \approx \delta_{ab} \frac{\partial q^a}{\partial q'^c} \frac{\partial q^b}{\partial q'^d} dq'^c dq'^d = g_{cd} dq'^c dq'^d. \tag{1.131}$$

Hence, to compute distances (and hence other metrical quantities) in coordinates q', we need to know the matrix $g_{ab}(q')$. So, g_{ab} plays the same role in the coordinates q'^a as δ_{ab} in the Cartesian system: it determines the metrical properties of the space in the coordinates q'. It is called the *metric of configuration space in the coordinates* q'. The length of a vector and of a line can be obtained from (1.123), (1.124) replacing δ_{ab} on $g_{ab}(q')$

$$|\vec{w}|^2 = g_{ab}(q') w'^a w'^b. \tag{1.132}$$
$$l = \int_{t_1}^{t_2} dt \sqrt{g_{ab}(q') \dot{q}'^a \dot{q}'^b}. \tag{1.133}$$

We point out that these expressions are exact, in contrast to the approximate expression for the distance (1.131); see Chap. 6.

Symmetries of an action. This is a good point to discuss the notion of *action symmetry* as a special class of coordinate transformations. Generally, L' as a function of its arguments is different from L, $L'(q', \dot{q}') \neq L(q, \dot{q})|_{q \to q'}$ (Eq. (1.130) clearly shows this). But if it differs on the total derivative term (see (1.109)), $L'(q', \dot{q}') = L(q, \dot{q})|_{q \to q'} + \dot{N}$, or, equivalently

$$L\left(q^a(q'^b), \frac{\partial q^a}{\partial q'^b} \dot{q}'^b \right) = L(q', \dot{q}') + \dot{N}, \tag{1.134}$$

[8] A systematic discussion of the underlying geometry will be given in Chap. 6.

the change of coordinates is called the *global symmetry* of the action. The action is called *invariant* under the transformation.

When a transformation represents the symmetry, it is not necessary to carry out computations to obtain equations of motion in the new coordinates, for they appear from the equations in the initial coordinates by the replacement

$$\left.\frac{\delta S}{\delta q^b}\right|_{q \to q'} = \frac{\delta S'}{\delta q'^a}. \tag{1.135}$$

An important example is given by the Galilean transformation

$$\begin{aligned} t &= t' + a, \\ \vec{r} &= R\vec{r}\,' + \vec{v}t + \vec{c}. \end{aligned} \tag{1.136}$$

This is a symmetry of the action describing a system of particles $\vec{r}_i$, $i = 1, 2, \ldots n$, with a potential that depends only on the relative distance between them

$$S = \int dt \left[\frac{1}{2} m_i \dot{\vec{r}}_i^2 - U(r_{ij})\right], \text{ here } (r_{ij})^2 = \sum_{a=1}^{3} (x_i^a - x_j^a)^2. \tag{1.137}$$

Indeed, substituting (1.136) into (1.137), we obtain $L(\vec{r}(\vec{r}\,')) = L(\vec{r}\,') + \left[2\vec{v}R\vec{r} + \vec{v}^{\,2}t\right]^{\cdot}$.

Examples

1. **Kepler's problem in polar coordinates.** A particle in the central field on a plane is described by the action

$$S = \int dt \left[\frac{1}{2} m \left(\dot{x}^2 + \dot{y}^2\right) + \frac{\alpha}{\sqrt{x^2 + y^2}}\right]. \tag{1.138}$$

This implies the Newton equations (1.26). Introducing the polar coordinates

$$x = r\cos\theta, \qquad x = r\sin\theta, \tag{1.139}$$

we obtain for the derivatives $\dot{x} = \dot{r}\cos\theta - r\dot{\theta}\sin\theta$, $\dot{y} = \dot{r}\sin\theta = r\dot{\theta}\cos\theta$. Then the action acquires the form

$$S = \int dt \left[\frac{1}{2} m \left(\dot{r}^2 + r^2\dot{\theta}^2\right) + \frac{\alpha}{r}\right]. \tag{1.140}$$

This implies the Lagrangian equations

$$\ddot{r} - r\dot{\theta}^2 + \frac{\alpha}{mr^2} = 0, \qquad (r^2\dot{\theta})^{\cdot} = 0. \tag{1.141}$$

From the second equation we obtain $r^2\dot{\theta} = \frac{L}{m}$ (for this choice of integration constant the quantity L is an angular momentum of the particle). Using this expression, θ can be excluded from the first equation, which now reads

$$\ddot{r} - \frac{L^2}{m^2r^3} + \frac{\alpha}{mr^2} = 0. \tag{1.142}$$

This is one advantage of using the polar coordinates for the case: instead of the system of two equations we deal with only one differential equation for one variable. Computing its product with $m\dot{r}$ we obtain $\left[\frac{m}{2}\left(\dot{r}^2 + \frac{L^2}{m^2r^2}\right) - \frac{\alpha}{r}\right]^{\cdot} = 0$, then $\frac{m}{2}\left(\dot{r}^2 + \frac{L^2}{m^2r^2}\right) - \frac{\alpha}{r} = E$. Comparing the l.h.s. of this equality with the Lagrangian (1.138), we conclude that E represents the total energy of the particle with an angular momentum L. Combining the results, the system (1.141) reduces to

$$\dot{r} = \pm\frac{1}{m}\sqrt{2m\left(E + \frac{\alpha}{r}\right) - \frac{L^2}{r^2}}, \qquad \dot{\theta} = \frac{L}{mr^2}. \tag{1.143}$$

These equations can be immediately integrated out giving the law of evolution of the particle, $r(t)$, $\theta(t)$. We discuss only its trajectory, that is, the function $r(\theta)$. Treating $r = r(t), \theta = \theta(t)$ as parametric equations for the trajectory, we write $\frac{dr}{dt} = \frac{dr}{d\theta}\frac{d\theta}{dt}$. Then the system (1.142) implies the following equation for the trajectory

$$\frac{dr}{d\theta} = \pm\frac{r^2}{L}\sqrt{2m\left(E + \frac{\alpha}{r}\right) - \frac{L^2}{r^2}}, \quad \text{or}$$
$$d\theta = \pm\frac{Ldr}{r^2\sqrt{2m\left(E + \frac{\alpha}{r}\right) - \frac{L^2}{r^2}}}. \tag{1.144}$$

Integrating, we obtain

$$\theta + c = \pm\arccos\frac{1}{e}\left(\frac{p}{r} - 1\right), \quad \text{here} \quad p = \frac{L^2}{m\alpha},$$
$$e = \sqrt{1 + \frac{2EL^2}{m\alpha^2}}. \tag{1.145}$$

Choosing the $+$ sign as well as $c = 0$ (it can be verified that any other choice leads to the same *trajectory*), the final expression for the trajectory is

$$r = \frac{p}{1 + e cos\theta}. \tag{1.146}$$

This is an equation of a conic section in polar coordinates (the ellipse for eccentricity $0 < e < 1$, the parabola for $e = 1$ and the hyperbola for $e > 1$). To see this better, let us make the inverse change $\cos\theta = \frac{x}{r}$, $r = \sqrt{x^2 + y^2}$ in Eq. (1.146). For the case $0 < e < 1$, the result can be written as

$$\frac{(x+c)^2}{a^2} + \frac{y^2}{b^2} = 1, \quad \text{here} \quad a = \frac{p}{1-e^2}, \quad b = \frac{p}{\sqrt{1-e^2}},$$
$$c = \frac{ep}{1-e^2}. \tag{1.147}$$

This is an ellipse with the center shifted to the left by a distance c; see Fig. 1.9 on page 37. The shift c coincides with the focal distance, $c^2 = a^2 - b^2$. Hence we have the ellipse with the right focus at the origin of the Cartesian system, which coincides with the center of force.

2. **Two body problem.** One of the important problems of classical mechanics that admits an analytic solution consists of the analysis of motion of two particles subject to a central force. To put this in concrete terms, we consider a system with an inverse degree potential. The corresponding action

$$S = \int dt \left[\frac{1}{2} m_1 \dot{\vec{r}}_1^2 + \frac{1}{2} m_2 \dot{\vec{r}}_2^2 + \frac{\alpha}{|\vec{r}_2 - \vec{r}_1|}\right], \tag{1.148}$$

leads to the Lagrangian equations

$$m_i \ddot{\vec{r}}_i = -\frac{\alpha \vec{r}_i}{|\vec{r}_2 - \vec{r}_1|^3}, \qquad i = 1, 2. \tag{1.149}$$

These have complex right-hand sides that depend on all variables of the problem. To improve this, notice that defining the quantity $r = \vec{r}_2 - \vec{r}_1$, we simplify the potential term in the action. So let us define the relative position vector $\vec{r} = \vec{r}_2 - \vec{r}_1$ and the *center of mass vector* $\vec{R} = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1+m_2}$. These equalities can be inverted

$$\vec{r}_1 = \vec{R} - \frac{m_2\vec{r}}{m_1 + m_2}, \qquad \vec{r}_2 = \vec{R} + \frac{m_1\vec{r}}{m_1 + m_2}, \tag{1.150}$$

so we can perform the change of coordinates, $(\vec{r}_1, \vec{r}_2) \to (\vec{R}, \vec{r})$. In the new variables the action reads

$$S = \int dt \left[\frac{1}{2} M \dot{\vec{R}}^2 + \frac{1}{2} m \dot{\vec{r}}^2 + \frac{\alpha}{|\vec{r}|} \right], \quad \text{here} \quad M = m_1 + m_2,$$
$$m = \frac{m_1 m_2}{m_1 + m_2}. \tag{1.151}$$

This gives equations of motion for two fictitious particles that do not interact with each other

$$\ddot{\vec{R}} = 0, \qquad m\ddot{\vec{r}} = \frac{\alpha \vec{r}}{|\vec{r}|^3}. \tag{1.152}$$

The center-of-mass "particle" moves with constant velocity along the straight line $\vec{R}(t) = \vec{R}_0 + \vec{V} t$. For the $\vec{r}$-particle we have the Kepler problem, so its trajectory is an ellipse with the semi-axis a and b. Given vectors $\vec{R}(t)$, $\vec{r}(t)$, the evolution of true particles is obtained according to Eq. (1.150). A qualitative picture of the motion can easily be obtained in the *center-mass coordinate system* (it is a Cartesian system with its origin at the point $\vec{R}(t)$; see Fig. 1.10 on page 37). The position vectors of the true particles with respect to this system are $\vec{r}\,'_1 = \vec{r}_1 - \vec{R}$, $\vec{r}\,'_2 = \vec{r}_2 - \vec{R}$. Comparing these equalities with (1.150) we conclude

$$\vec{r}\,'_1 = -\frac{m_2}{m_1 + m_2} \vec{r}, \qquad \vec{r}\,'_2 = \frac{m_1}{m_1 + m_2} \vec{r}, \tag{1.153}$$

Hence the trajectory of the $\vec{r}_1$-particle is the ellipse of the $\vec{r}$-particle suppressed by the factor $\frac{m_2}{m_1+m_2}$, which has the semi-axis $\frac{m_2 a}{m_1+m_2}$, $\frac{m_2 b}{m_1+m_2}$. For the $\vec{r}_2$-particle, the ellipse is suppressed by $\frac{m_1}{m_1+m_2}$. Taking into account that $\vec{r}_1$ and $\vec{r}_2$ are antiparallel at any instant, we conclude that the trajectories of the true particles (in the center-mass system) are ellipses with one focus at the center of mass and with the same major semi-axis. Besides, at each instant of motion, a straight line connecting the particles passes through the center of mass; see Fig. 1.11 on page 37.

1.7 Examples of Continuous (Field) Systems

While historically formalism of classical mechanics has been developed for finite-dimensional systems, many of its methods can be generalized for the description of physical systems with an infinite (non-countable) number of degrees of freedom. In this case the configuration of a system is determined by a function of the space-

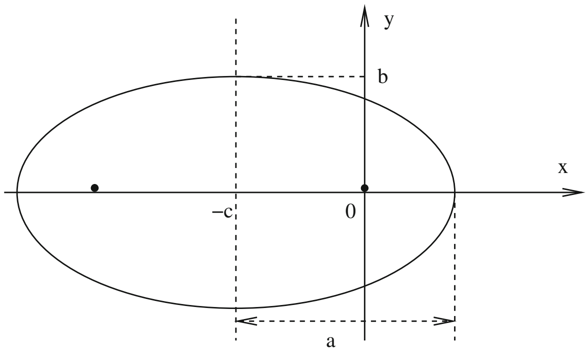

Fig. 1.9 For the case $0 < e < 1$, the trajectory of motion in the central field is an ellipse with one focus at the center of force O

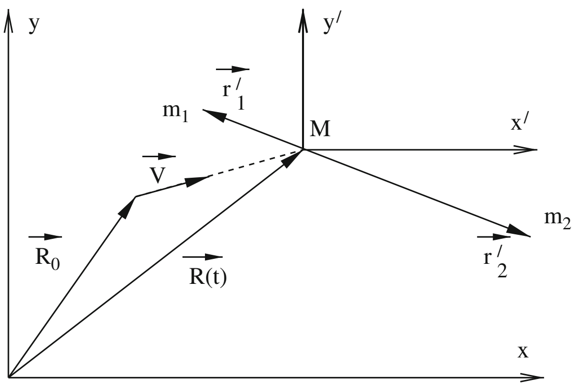

Fig. 1.10 The center-mass coordinates x', y' are the Cartesian system with its origin in the center of mass $\vec{R}(t)$

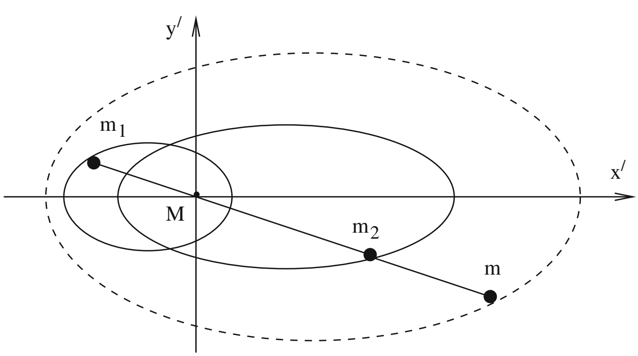

Fig. 1.11 Two-body problem in the center-mass coordinates. Ellipses of m_1 and m_2-particles are obtained from an ellipse of a fictitious m-particle according to (1.153)

time point, say $\varphi^B(\tau, x^a)$, $B = 1, 2, \ldots, n$ called a *field*. Hence to determine an instantaneous configuration we need to specify n numbers φ^A, *at each point* of space.

Here we present a few illustrative examples of such a system, and outline the relevant results of variational analysis that will be used later.

Small oscillations of a non-stretchable string. Consider a non-stretchable string of length L and linear mass density ρ. The string is fixed at one end but can slide without friction at the other end; see Fig. 1.12 on page 38. It is under the constant force T applied to the free end (imagine the mass M hanged on that end, then $T = Mg$). T is called the *string tension*. The tension acts as a return force when the string is replaced out of its position of equilibrium.

The configuration of the string can be described by the displacement function $\varphi(t, x)$; see Fig. 1.12 on page 38. It is an example of a continuous system: instantaneous configuration is determined by the function $\varphi(x)$, $x \in [0, L]$ instead of a set of numbers q^a. Intuitively, it is convenient to imagine that we deal with an infinite number of coordinates $\varphi_x(t) \equiv \varphi(x, t)$ labeled by a "continuous index" x instead of $q_a(t)$ with the discrete index a.

To write an action, we consider the approximation of a small oscillation, $\varphi \ll L$, of the light string, $\rho L \ll M$, and suppose that points of the string can move in a vertical direction only.

The potential energy is equal to the work performed by the force T in the displacement of M from equilibrium to the present instantaneous position, $U = Th$, where h is the difference of lengths of the configurations $h = \int_0^L dx\sqrt{1 + \left(\frac{\partial\varphi}{\partial x}\right)^2} - L \approx \frac{1}{2}\int_0^L dx(\frac{\partial\varphi}{\partial x})^2$ (the integrand has been expanded in a power series up to a linear order). Thus the potential energy is

$$U = T\int_0^L dx\left(\frac{\partial\varphi}{\partial x}\right)^2. \tag{1.154}$$

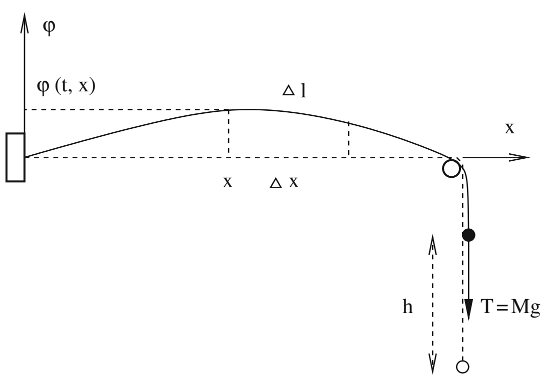

Fig. 1.12 Instantaneous configuration of a string is described by the displacement function $\varphi(t, x)$

To compute the kinetic energy, consider first a small section $\triangle l$ of the string. Its kinetic energy is approximated by (x_0 stands for a point inside $\triangle l$) $\frac{1}{2}\rho\triangle l(\frac{\partial\varphi(t,x_0)}{\partial t})^2 = \frac{1}{2}\rho\sqrt{1+(\frac{\partial\varphi}{\partial x})^2}\triangle x(\frac{\partial\varphi}{\partial t})^2 = \frac{1}{2}\rho(1+(\frac{\partial\varphi}{\partial x})^2)\triangle x(\frac{\partial\varphi}{\partial t})^2 \approx \frac{1}{2}\rho(\frac{\partial\varphi}{\partial t})^2\triangle x$. Integration along the string gives the total kinetic energy

$$T = \frac{1}{2}\rho\int_0^L dx\left(\frac{\partial\varphi}{\partial t}\right)^2. \tag{1.155}$$

The Lagrangian action reads

$$S = \frac{1}{2}\int dtdx\left[\rho\left(\frac{\partial\varphi}{\partial t}\right)^2 - T\left(\frac{\partial\varphi}{\partial x}\right)^2\right]. \tag{1.156}$$

As compared with a finite-dimensional system, see for example Eq. (1.112), the sum on the discrete label i is now replaced by an integral on the continuous label x.

Exercise
Confirm that the same action arises for a stretchable string with fixed ends, supposing that potential energy is proportional to its stretch.

A continuous system as a limit of a discrete one. To study longitudinal vibrations of an elastic pivot, we first approximate it by a chain of N particles connected by springs. The description of the pivot arises at the limit $N \to \infty$. Note that this gives an intuitive explanation why the methods first developed for a system with a finite number of degrees of freedom work for continuous systems as well.

Starting from a spring chain of N particles (see Example 2 of Sect. 1.5) let us try to find the limit when the number of particles on the *fixed* interval $[0, (N-1)\triangle x]$ tends to infinity, $N \to \infty$ (this implies $\triangle x \to 0$). In the process, we also vary the mass of each particle and the rigidity of each spring according the rules $m = \rho\triangle x$, $k = \frac{T}{\triangle x}$, where ρ, T are fixed numbers. Note that ρ has the dimension $\frac{\text{kg}}{\text{m}}$, which is of linear mass density, while T has the dimension of the tension (force) $\frac{\text{kg}\times\text{m}}{\text{sec}^2}$. The action (1.112) reads

$$S = \frac{1}{2}\int dt\left[\rho\sum_{i=1}^{N}(\dot{q}_i)^2\triangle x - T\sum_{i=1}^{N-1}\left(\frac{q_{i+1}-q_i}{\triangle x}\right)^2\triangle x\right]. \tag{1.157}$$

To understand what happens at the limit, it is convenient to plot the displacements q_i (at some fixed instant t) on the plane (x, q); see Fig. 1.13 on page 40. Clearly, when $N \to \infty$, the sequence $\{q_i(t)\}$ approximates to a function $q = \varphi(t, x)$, which is defined by the condition $\varphi(t, x_i) = q_i(t)$. Then sums in the integrand of Eq. (1.157) are just the partial integral sums of the functions $\partial_t\varphi$, $\partial_x\varphi$. Hence at the limit we obtain

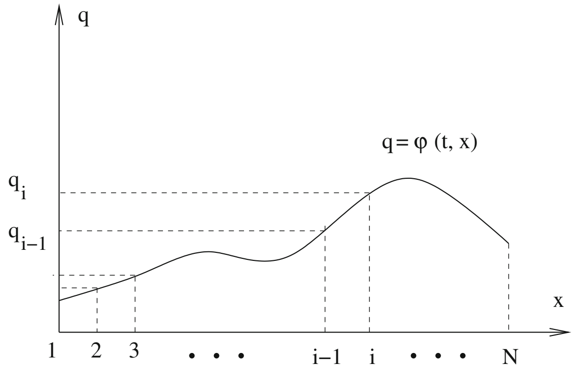

Fig. 1.13 When the number of springs tends to infinity, their displacements q^i define the function $\varphi(t, x)$ according to the rule $\varphi(t, x^i) = q^i(t)$

$$\sum_{i=1}^{N}(\dot{q}_i)^2\triangle x = \sum_{i=1}^{N}\dot{\varphi}^2(t, x_i)\triangle x \overset{N\to\infty}{\longrightarrow} \int dx\left(\frac{\partial\varphi}{\partial t}\right)^2,$$

$$\sum_{i=1}^{N-1}\left(\frac{q_{i+1}-q_i}{\triangle x}\right)^2\triangle x = \sum_{i=1}^{N-1}\left(\frac{\varphi(t, x_{i+1})-\varphi(t, x_i)}{\triangle x}\right)^2\triangle x \overset{N\to\infty}{\longrightarrow}$$
$$\int dx\left(\frac{\partial\varphi}{\partial x}\right)^2. \tag{1.158}$$

That is, the action acquires the form

$$S = \frac{1}{2}\int dtdx\left[\rho\left(\frac{\partial\varphi}{\partial t}\right)^2 - T\left(\frac{\partial\varphi}{\partial x}\right)^2\right]. \tag{1.159}$$

Note that we have arrived at the same final expression as in the previous example; see (1.156). Both the transverse and longitudinal vibrations of a string obey the same equation.

Lagrangian action and equations of field theory. The previous example suggests formal rules for the transition from the finite- dimensional formalism to the field one. We have the table

$$i \to x^a,\ q_i^a \to \varphi^B(x^a),\ q_i^a(t) \to \varphi^A(t, x^a),\ \sum_i \to \int d^3x. \tag{1.160}$$

The least action principle works for field systems as well. For the later use, we outline the resulting formulas for the case of a field $\varphi^B(\tau, x^a)$. To adapt our results to this case, it is sufficient to apply the table presented above to the basic formulas of Sect. 1.5.

A Lagrangian function often has the form $L(\varphi^B, \partial_\tau\varphi^B, \partial_a\varphi^B, \tau, x^a)$, and the Lagrangian action functional is given by the integral

$$S[\varphi] = \int_\Omega d\tau d^3x L(\varphi^B, \partial_\tau\varphi^B, \partial_a\varphi^B, \tau, x^a), \tag{1.161}$$

over a space-time region Ω. A variational problem consists in searching for the function $\varphi^B(t, x)$ that provides an extremum of the functional in an appropriately-chosen class of functions. The choice of the class depends on the particular problem under consideration. We often work with the functions that acquire the prescribed values at the initial τ_1 and the final τ_2 instants of time and vanish at spatial infinity

$$\begin{aligned} &\varphi^B(\tau_1, x) = \varphi_1^B(x), \;\; \varphi^B(\tau_2, x) = \varphi_2^B(x), && \text{initial conditions,} \\ &\lim_{x\to\infty} \varphi^B = 0, && \text{boundary condition.} \end{aligned} \tag{1.162}$$

An extremum is a solution to the following system of partial differential equations

$$\partial_\tau \frac{\partial L}{\partial(\partial_\tau\varphi^B)} + \partial_a \frac{\partial L}{\partial(\partial_a\varphi^B)} - \frac{\partial L}{\partial\varphi^B} = 0, \quad B = 1, 2, \ldots, n. \tag{1.163}$$

which generalizes the Lagrangian equations for the case of a field system. In relativistic theories, temporal and spatial coordinates are unified in a unique quantity $x^\mu = (x^0, x^i)$, where $x^0 = ct$. So the previous formulas acquire a more compact form

$$S[\varphi] = \int_\Omega d^4x L(\varphi^B, \partial_\mu\varphi^B, x^\mu), \tag{1.164}$$

$$\partial_\mu \frac{\partial L}{\partial\partial_\mu\varphi^B} - \frac{\partial L}{\partial\varphi^B} = 0. \tag{1.165}$$

Applying the rules to the string action (1.159) we obtain the equation

$$(\partial_t^2 - c^2\partial_x^2)\varphi = 0, \quad \text{where} \quad c^2 = \frac{T}{\rho}. \tag{1.166}$$

The constant c has the dimension of velocity $\frac{\text{m}}{\text{sec}}$. This is one of the basic equations of mathematical physics, known as the (one-dimensional) *wave equation*. Without going into detail, we present a few examples of its solutions.

Examples of solutions to the wave equation. 1. Infinite string. Solutions to a partial differential equation, in contrast to an ordinary one, generally depend on arbitrary functions. To see this for the present case, write (1.166) in the form

$$(\partial_t - c\partial_x)(\partial_t + c\partial_x)\varphi = 0. \tag{1.167}$$

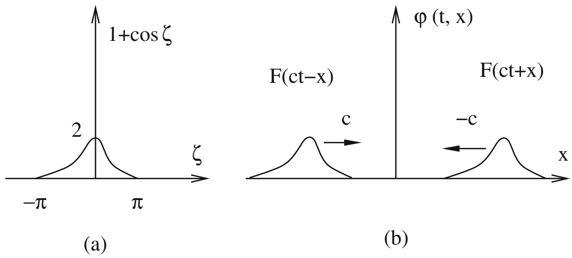

Fig. 1.14 **(a)** Any function $F(\xi)$ can be used to construct a solution to the wave equation: $\varphi = F(ct - x)$. **(b)** Since the wave equation is linear, two wave packets pass right through each other

Then it is clear that $\varphi = F(ct-x)$ and $\varphi = G(ct+x)$, where F and G are *arbitrary* functions of the indicated arguments, obey the wave equation. Take, for example, the following wave packet (see Fig. 1.14a on page 42)

$$F(\xi) = \begin{cases} 0 & \xi < -\pi, \\ 1 + \cos\xi, & -\pi \leq \xi \leq \pi, \\ 0 & \xi > \pi \end{cases} \tag{1.168}$$

The corresponding solution $\varphi = F(ct - x)$ describes the evolution of this perturbation along the string. The packet travels to the right with velocity c, without changing its form, that is, it behaves like a particle.

The wave equation is linear, that is, it has the property that any linear combination of solutions is itself a solution. For instance, the solution $\varphi = F(ct-x)+F(ct+x)$ describes two packets traveling in opposite directions; see Fig. 1.14b on page 42. During the interval $t < -\frac{\pi}{c}$ they approach one another; then they "scatter" near the point $x = 0$ during the interval $-\frac{\pi}{c} < t < \frac{\pi}{c}$; and then diverge, keeping their initial profiles unaltered after the scattering. Linearity of the wave equation will be spoiled when we include in the action interaction terms (like φ^3, $\varphi^4, \ldots$). This would lead to non-trivial scattering effects.

Note that the solutions have been obtained without taking into account any type of boundary conditions. So, they correspond to the case of an infinite string.

Exercises

1. Confirm that $\varphi = \frac{1}{2}g(x-ct)+\frac{1}{2}g(x+ct)+\frac{1}{2c}\int_{x-ct}^{x+ct} v(\xi)d\xi$ is a solution to the wave equation that obeys the initial conditions $\varphi(0,x) = g(x)$, $\partial_t\varphi(0,x) = v(x)$.
2. Consider the case $g(x) = 0$, $v(x) = F(x)$, where F is given by Eq. (1.168). Compute and draw $\varphi(t,x)$ at an instant t_0 such that $2ct_0 > 2\pi$ (Suggestion: consider separately the following intervals of variation for x: $x + ct \in]-\infty, -\pi[\cup]-\pi, \pi[\cup]\pi, -\pi + 2ct[\cup]-\pi + 2ct, \pi + 2ct[\cup]\pi + 2ct, \infty[$.

2. String with fixed ends. Consider the wave Eq. (1.166) with the boundary conditions

$$\varphi(t,0)=\varphi(t,L)=0. \tag{1.169}$$

The linearity of the wave equation allows us to use a powerful tool of the Fourier series expansion to look for solutions. Any function that belong to the class (1.169) can be presented by the Fourier series in terms of $\sin\frac{\pi n}{L}x$

$$\varphi(t,x)=\sum_{n=1}^{\infty}c_n(t)\sin\frac{\pi n}{L}x. \tag{1.170}$$

Substitution into Eq. (1.166) turns the partial differential equation into a system of ordinary equations for the coefficients $c_n(t), \quad n=1,2,\ldots$

$$\ddot{c}_n+\left(\frac{\pi nc}{L}\right)^2 c_n=0, \text{ then } c_n=a_n\sin\frac{\pi nc}{L}t+b_n\cos\frac{\pi nc}{L}t. \tag{1.171}$$

So we have found a set of elementary solutions

$$\begin{aligned}\varphi_n(t,x)&=\left(a_n\sin\frac{\pi nc}{L}t+b_n\cos\frac{\pi nc}{L}t\right)\sin\frac{\pi n}{L}x\\&=d_n\sin\frac{\pi n}{L}x\sin\left(\frac{\pi nc}{L}t+\alpha_n\right),\end{aligned} \tag{1.172}$$

where $d_n=\sqrt{a_n^2+b_n^2}$, $\alpha_n=\arctan\frac{b_n}{a_n}$. These are called *standing waves*. The standing wave with $n=3$ is drawn in Fig 1.15 on page 43. The string points with $x=0, \frac{L}{3}, \frac{2L}{3}, L$ are at rest at any instant of time. Other points accomplish harmonic oscillations with the same frequency $\frac{3\pi c}{L}$ and the amplitude A that depends on x, $A=d_n\sin\frac{\pi n}{L}x$.

According to (1.170), any solution to the problem (1.166), (1.169) is given by sum of the standing waves

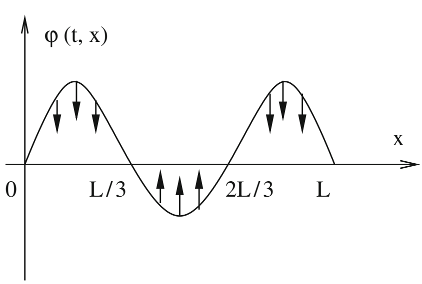

Fig. 1.15 Standing wave with $n=3$

$$\varphi(t,x)=\sum_{n=1}^{\infty}\left(a_n\sin\frac{\pi nc}{L}t+b_n\cos\frac{\pi nc}{L}t\right)\sin\frac{\pi n}{L}x. \tag{1.173}$$

Exercises

1. Confirm that the general solution (1.173) can be written in the form

$$\varphi(t,x)=\sum_{n=-\infty}^{\infty}\alpha_n\left(e^{i\omega_n(ct-x)}-e^{i\omega_n(ct+x)}\right), \tag{1.174}$$

where $\omega_n=\frac{\pi n}{L}$, $\alpha_{-n}=\alpha_n^*$.

2. Find a solution to the wave Eq. (1.166) with the following boundary conditions: $\varphi(t,0)=0$, $\varphi(t,L)=f\sin\beta t$, where f =const, β =const. They correspond to a periodic force applied to the right-hand side of the string. (Suggestion: instead of using the Fourier series, look for a solution to the form $\varphi(t,x)=T(t)X(x)$).

1.8 Action of a Constrained System: The Recipe

An example. While the Lagrangian function can often be written as the difference between kinetic and potential energy, $L=T-U$, this rule is not universal. There are a lot of rather simple systems when it does not work. Take, for example, a mathematical pendulum on a plane, see Fig. 1.16 on page 44. It is clear that the difference

$$T-U=\frac{1}{2}m\left(\dot{x}^2+\dot{y}^2\right)-mgy, \tag{1.175}$$

being considered as the system Lagrangian, leads to wrong equations of motion. These correspond to a free fall in a gravitational field, and do not take into account

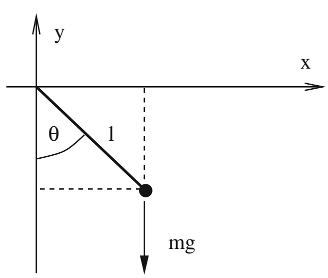

Fig. 1.16 Mathematical pendulum

the *constraint* $x^2 + y^2 = l^2$ that must be satisfied at any instant of motion. Note also that it is impossible to simply add the constraint to the equations of motion, since it would lead to an incompatible system (confirm this!).

The pendulum is an example of a system with a *holonomic* (or *kinematic*) constraint (generally, a holonomic constraint is a restriction on configuration space variables, $G(q^a, \tau) = 0$, that must be satisfied at any instant). In this section we discuss a recipe suitable for the construction of a Lagrangian for this case.

Roughly speaking it can be formulated as follows. First, forget the constraints, and write the Lagrangian of the unconstrained system (it may be $L = T - U$). Second, find a solution to constraints and write L in terms of the independent variables.

For the pendulum, one possibility is $y = -\sqrt{l^2 - x^2}$, then $\dot{y} = \frac{x\dot{x}}{\sqrt{l^2-x^2}}$. Substitution into Eq. (1.175) gives the action

$$S[x] = \int dt \left[\frac{1}{2} m \left(1 - \left(\frac{x}{l} \right)^2 \right)^{-1} \dot{x}^2 + mgl\sqrt{1 - \left(\frac{x}{l} \right)^2} \right]. \tag{1.176}$$

This implies the Lagrangian equation

$$\ddot{x} + \frac{g}{l} x\sqrt{1 - \left(\frac{x}{l} \right)^2} + \frac{x\dot{x}^2}{l^2 \left(1 - \left(\frac{x}{l} \right)^2 \right)} = 0. \tag{1.177}$$

In the approximation of small displacements $\frac{x}{l} \ll 1$ this reduces to the well-known equation of harmonic oscillations

$$\ddot{x} + \frac{g}{l} x = 0. \tag{1.178}$$

Let us discuss the recipe from the geometric point of view. The constraint $x^2 + y^2 = l^2$ represents the equation of a circle in configuration space. The variable x can be taken as a system of coordinates on the line, then $y = -\sqrt{l^2 - x^2}$ is a parametric equation of the line (in the vicinity of the point we are interested in). This solves the constraint. So, geometrically the recipe consists of restricting the unconstrained Lagrangian function on the line.

It is important to notice a freedom implied by the recipe. First, we are free to choose the parametrization of the line. Take, for example, the angle θ as a coordinate. Then the parametric equations of the circle are $x = l \sin\theta$, $y = -l \cos\theta$. They solve the constraint, so can be used in (1.175). This gives a Lagrangian in terms of θ

$$L(\theta) = \frac{1}{2} m l^2 \dot{\theta}^2 + mgl \cos\theta, \tag{1.179}$$

and the equation of motion is

$$\ddot{\theta} + \frac{g}{l} \sin\theta = 0. \tag{1.180}$$

For a small θ one approximates $\sin\theta \approx \theta$ obtaining the equation of harmonic oscillations for θ, $\ddot{\theta} + \frac{g}{l}\theta = 0$.

Second, we are free to choose the generalized coordinates writing an anzatz (1.175) for L. This can be done, for example, in polar coordinates $x = r\sin\theta$, $y = -r\cos\theta$, which gives $L(r,\theta) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + mgr\cos\theta$. Then we use the constraint $r = l$ obtaining Eqs. (1.179) and (1.180) once again.

General recipe. Generalizing, consider a system with the generalized coordinates q^a, $a = 1, 2, \ldots, n$ constrained to move on a k-dimensional surface

$$G_i(q^a) = 0, \qquad i = 1, 2, \ldots, n-k. \tag{1.181}$$

Equations of the surface are assumed to be functionally independent

$$\mathrm{rank}\frac{\partial G_i(q^a)}{\partial q^b} = n - k. \tag{1.182}$$

Suppose also that in the absence of constraints the system can be described by the Lagrangian $L(q^a, \dot{q}^a)$. To write a Lagrangian of the constrained system, choose some coordinates (parametrization) s^α, $\alpha = 1, 2, \ldots, k$ on the surface, and write parametric equations of the surface

$$q^a = q^a(s^\alpha). \tag{1.183}$$

By construction, these solve the constraints, $G_i(q^a(s^\alpha)) \equiv 0$. Substitution of the parametric equations into $L(q^a)$ *gives a Lagrangian of the constrained system*

$$L(s^\alpha, \dot{s}^\alpha) \equiv L\left(q^a(s^\alpha), \dot{q}^a(s^\alpha)\right). \tag{1.184}$$

Using the basic formula (1.108), the Lagrangian equations are

$$\frac{\delta S[s]}{\delta s^\alpha} \equiv \frac{d}{d\tau}\frac{\partial L(s)}{\partial \dot{s}^\alpha} - \frac{\partial L(s)}{\partial s^\alpha} = 0 \quad \alpha = 1, 2, \ldots, k. \tag{1.185}$$

Using Eq. (1.184) as well as the formula for the derivative of a composed function, the equations can be written in terms of the initial Lagrangian

$$\frac{\delta S[s]}{\delta s^\alpha} = \frac{\partial q^a}{\partial s^\alpha}\left.\frac{\delta S[q]}{\delta q^a}\right|_{q^a(s^\alpha)} = 0. \tag{1.186}$$

Exercises

1. Check the equality (1.186).
2. Show that if the unconstrained Lagrangian is non-singular, the same is true for the Lagrangian of the constrained system (1.184).

For later use, we specify the results for a particular case of parametrization of the surface (1.181). Equation (1.182) guarantees that the constraints can be resolved with respect to $(n-k)$ variables among q^a, say q^i. Let the solution be

$$q^i = q^i(q^\alpha). \tag{1.187}$$

Then q^α can be taken as coordinates of the surface. Then the Lagrangian is given by

$$L(q^\alpha, \dot{q}^\alpha) \equiv L\left(q^i(q^\alpha), q^\alpha, \dot{q}^i(q^\alpha), \dot{q}^\alpha\right), \tag{1.188}$$

while the Lagrangian equations acquire the form

$$\left.\frac{\delta S[q^a]}{\delta q^\alpha}\right|_{q^i(q^\alpha)} + \frac{\partial q^i}{\partial q^\alpha}\left.\frac{\delta S[q^a]}{\delta q^i}\right|_{q^i(q^\alpha)} = 0. \tag{1.189}$$

If the constraints depend on time, $G_i(q^a, \tau) = 0$, it is considered to be a fixed parameter of the problem. The recipe remains the same; it is sufficient to replace $q^a(s^\alpha)$ by $q^a(s^\alpha, \tau)$ in the previous formulas.

Example (Thomson-Tait pendulum.) This consists of two equal masses m connected by a massless pivot of length $2b$, the middle of which is attached to the end of a massless pivot of length a, see Fig. 1.17 on page 48. The a-pivot can rotate freely in the (x, y)-plane while the b-pivot can rotate freely in a vertical plane.
The position of the masses can be described by the vectors $\vec{r}_1$, $\vec{r}_2$, that is, the configuration space is six-dimensional. The masses are constrained to move on two-dimensional surface determined by

$$([\vec{r}_1, \vec{r}_2], \vec{n}) = 0, \quad (\vec{r}_1 + \vec{r}_2, \vec{n}) = 0, \quad |\vec{r}_1 - \vec{r}_2| = 2b, \quad |\vec{r}_1 + \vec{r}_2| = 2a, \tag{1.190}$$

where $\vec{n}$ is a unit vector in the direction of the z-axis, $(\ ,\ )$ is a scalar product and $[\ ,\ ]$ is a vector product. Forgetting the constraints, the action is written

$$S = \int dt \frac{1}{2} m \left(\dot{\vec{r}}_1{}^2 + \dot{\vec{r}}_2{}^2\right). \tag{1.191}$$

Exercise
Explain why gravity will not contribute to the potential energy of the constrained system.
Let us take the angles φ and θ as coordinates on the surface. Expressions for $\vec{r}_i$ through the coordinates are

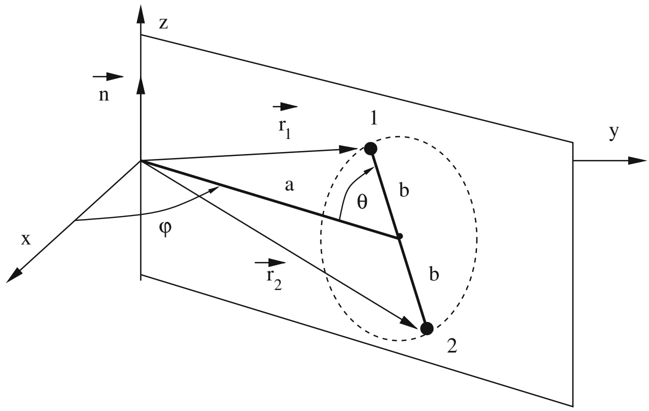

Fig. 1.17 Thomson-Tait pendulum. The a-pivot can rotate freely in the (x, y)-plane while the b-pivot can rotate freely in a vertical plane

$$\begin{aligned} &x_1 = a\sin\varphi - b\cos\theta\sin\varphi, \quad y_1 = a\cos\varphi - b\cos\theta\cos\varphi, \quad z_1 = b\sin\theta; \\ &x_2 = a\sin\varphi + b\cos\theta\sin\varphi, \quad y_2 = a\cos\varphi + b\cos\theta\cos\varphi, \quad z_2 = -b\sin\theta. \end{aligned} \tag{1.192}$$

These equations solve the constraints (1.190). Substituting them into (1.191) we obtain the action of the Thomson–Tait pendulum

$$S = \int dtm\left(b^2\dot{\theta}^2 + (a^2 + b^2\cos^2\theta)\dot{\varphi}^2\right). \tag{1.193}$$

This implies the equations of motion

$$\left[(a^2 + b^2\cos^2\theta)\dot{\varphi}\right]^{\cdot} = 0, \quad \ddot{\theta} + \dot{\varphi}^2\cos\theta\sin\theta = 0. \tag{1.194}$$

Note that it would not be an easy task to find these equations directly, without using variational analysis. The first equation implies $(a^2 + b^2\cos^2\theta)\dot{\varphi} = c =$const. This can be used to exclude $\dot{\varphi}$ from the second equation. Then the system acquires the form

$$\dot{\varphi} = \frac{c}{a^2 + b^2\cos^2\theta}, \quad \ddot{\theta} + \frac{c^2\cos\theta\sin\theta}{(a^2 + b^2\cos^2\theta)^2} = 0. \tag{1.195}$$

Note that the motion of the masses is not a composition of two circular motions with constant angular velocity, as might naively be expected.

Exercises

In the exercises below, masses move without friction on the plane x, y, a spring has an unstretched length l and rigidity k. Springs and pivots are massless.

1. The mass m is attached to a spring; see Fig. 1.18a on page 50. (A) Find the action in the Cartesian coordinates x, y. (B) Find the action and equations of motion in coordinates r, φ, where r is the instantaneous length of the spring.
2. Consider a pendulum with varying length $s = s(t)$, where $s(t)$ is a given function; see Fig. 1.18b on page 50. (A) Find the action in terms of y. (B) Find the action and equations of motion in terms of φ.
3. The masses M and m are attached to the ends of a pivot of length s. The mass M can move freely along the x-axis, while m oscillates in the x, y plane; see Fig. 1.18c on page 50. Find the action and equations of motion in coordinates z, φ.
4. A pivot with masses M and m is attached to one end of a spring. The mass M can move along the y-axis, while m oscillates in the x, y plane; see Fig. 1.19d on page 50. (A) Find the action in coordinates z, y. (B) Find the action and equations of motion in coordinates z, φ.
5. The mass M can slide freely along a pendulum of length s; see Fig. 1.19e on page 50. Find action of the system and equations of motion in coordinates r, φ, where r is the distance of M from the origin.
6. Consider a double pendulum on the plane x, y; see Fig. 1.19f on page 50. Find the action and equations of motion in coordinates φ, θ.
7. A pivot with masses M and m is attached to a spring; see Fig. 1.20g on page 50. M can move along the x-axis, while m oscillates in the x, y-plane. Find the action and equations of motion in coordinates x, φ.
8. Two masses M and m are attached to a rope of length s. The rope can slide freely around a disk of radius R; see Fig. 1.20h on page 50. Find the action and equations of motion in terms of y.
9. Consider a pivot of length s with masses M and m at the ends; see Fig. 1.20k on page 50. M can move along the y-axis, while m can move along the x-axis. Find the action and equations of motion in terms of φ.
10. The masses M and m are attached to the ends of a spring; see Fig. 1.21l on page 51. The mass M can move freely along the x-axis, while m can move along the horizontal string A, at a distance s from the x-axis. Find the action in coordinates x, φ.
11. The masses M and m are attached to the ends of a pivot of length s; see Fig. 1.21m on page 51. The masses can move along the circle of radius R. Find the action and equations of motion of the system.
12. Two pendulums of the same length s are connected by a spring of unstretched length l; see Fig. 1.21n on page 51. Find the action and equations of motion in coordinates φ, θ.

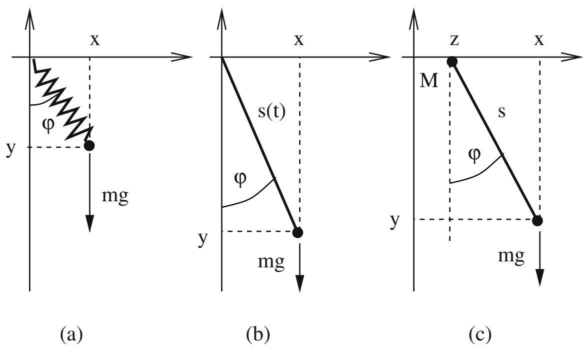

Fig. 1.18 Exercises

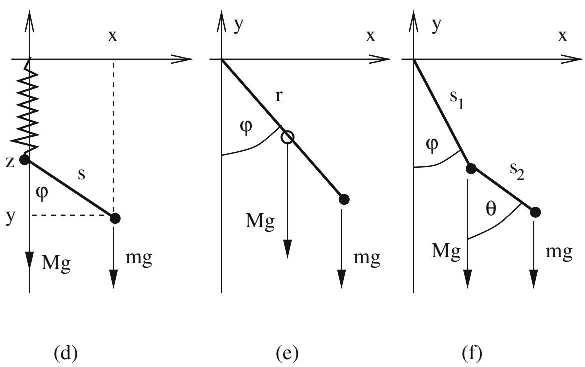

Fig. 1.19 Exercises

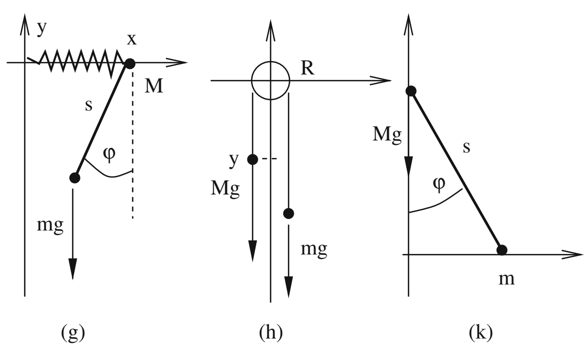

Fig. 1.20 Exercises

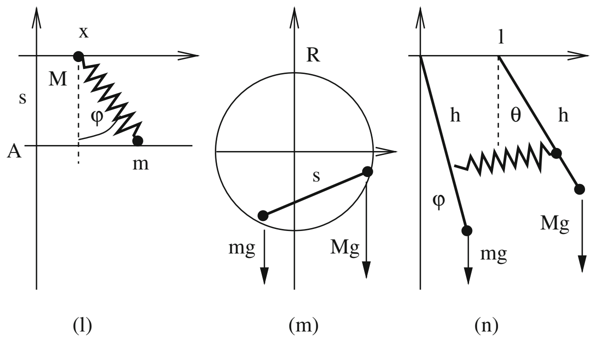

Fig. 1.21 Exercises

1.9 Action of a Constrained System: Justification of the Recipe

Holonomic constraints represent the idealization of a very strong force directed towards a surface of configuration space, and forcing a particle to move near the surface. This suggests a natural way [2] to confirm the recipe (1.184) for construction of the constrained-system Lagrangian.[9] We start from a system with a potential that produces the strong force (among others), and then take the limit of infinite force. Since the dimension of the configuration space is not essential for the discussion, we take a system with two generalized coordinates x, y and with the action being

$$S = \int d\tau \left[\frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 - U(x, y) - s(y - g(x))^2\right]. \qquad (1.196)$$

This depends on the parameter s =const. To put this in concrete terms, we suppose $U \geq 0$ in the region of interest. The second term of the potential energy grows when the particle goes away from the line $y = g(x)$. We consider a particle that starts on that line and has initial velocity tangent to the line

$$\begin{aligned} x(0) &= x_0, & \dot{x}(0) &= v_0, \\ y(0) &= g(x_0), & \dot{y}(0) &= \dot{g}(x_0). \end{aligned} \qquad (1.197)$$

As will be seen, at the limit $s \to \infty$ an action describing a particle in potential U and subject to the constraint $y = g(x)$ appears.

To confirm this, it is convenient to write the action in coordinates x, $\tilde{y} \equiv y - g(x)$

$$S = \int d\tau \left[\frac{1}{2}\dot{x}^2 + \frac{1}{2}(\dot{\tilde{y}} + \dot{g}(x))^2 - U(x, \tilde{y} + g(x)) - s\tilde{y}^2\right]. \qquad (1.198)$$

[9] We will need to use the law of conservation of energy. So the motivation works only for time-independent constraints.

Then the equations of motion are

$$\ddot{x} + (\ddot{\tilde{y}} + \ddot{g}(x))\frac{dg}{dx} + \frac{\partial}{\partial x}U(x, \tilde{y} + g(x)) = 0, \tag{1.199}$$

$$\ddot{\tilde{y}} + \ddot{g}(x) + \frac{\partial U}{\partial \tilde{y}} + 2s\tilde{y} = 0. \tag{1.200}$$

We can estimate the $\tilde{y}$ coordinate using the law of conservation of energy, $E = T + U + s\tilde{y}^2$, as follows: $\tilde{y}(\tau) = s^{-\frac{1}{2}}\sqrt{2(E - T - U)} \ll s^{-\frac{1}{2}}\sqrt{2E}$. Hence, the particle with total energy E can not move far from the curve (in $\tilde{y}$-direction) more than a distance proportional to $\frac{1}{\sqrt{s}}$. Then $\tilde{y} \stackrel{s\to\infty}{\longrightarrow} 0$, or $y(\tau) \stackrel{s\to\infty}{\longrightarrow} g(x(\tau))$, that is at the limit our particle is confined to move on the line. At this limit the Eq. (1.199) for the x-coordinate reads

$$\ddot{x} + \ddot{g}(x)\frac{dg}{dx} + \frac{\partial}{\partial x}U(x, g(x)) = 0. \tag{1.201}$$

The final observation is that this can be obtained from the action

$$S = \int d\tau \left[\frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 - U(x, y)\right], \tag{1.202}$$

where y is replaced according the constraint $y = g(x)$. Thus, we have arrived at the recipe (1.188).

1.10 Description of Constrained System by Singular Action

To describe a constrained system, our ideology was to reduce the number of variables, from q^a, $a = 1, 2, \ldots, n$ to s^α, $\alpha = 1, 2, \ldots k < n$. Unfortunately, however, this can result in the loss of some properties that were presented in the initial variables.

For example, both the unconstrained action (1.191) and the constraints (1.190) of the Thomson-Tait pendulum have a manifest rotational symmetry[10] $\vec{r} \to R\vec{r}$. This is hidden in the formulation (1.193) (in which only the rotational symmetry in the (x, y)-plane, $\varphi \to \varphi + \text{const}$ is evident).

The same happens if we use some of the initial variables to parameterize the constraint surface. For example, take a free particle on a circle of radius l. Both the unconstrained Lagrangian $\frac{1}{2}m(\dot{x}^2+\dot{y}^2)$ and the constraint $x^2+y^2 = l^2$ have the rotation symmetry $\vec{r} \to R\vec{r}$. This is not manifest in the Lagrangian $\frac{1}{2}m\left(1 - \left(\frac{x}{l}\right)^2\right)^{-1}\dot{x}^2$ which appears after using the constraint $y = \pm\sqrt{l^2 - x^2}$.

[10] Generally, a symmetry is called manifest if it is expressed by a *linear* transformation.

Since symmetries often play a fundamental role in the analysis of a theory, it would be desirable to find a way to keep them untouched. This implies that we continue to use the initial configuration space variables for the description of a constrained system. Strange as it may seem, this can be achieved following the opposite ideology: instead of reducing the configuration space we extend it, adding new variables into the formulation.

As before, we take a system with an unconstrained Lagrangian $L(q^a, \dot{q}^a)$ and the constraints $G_i(q^a) = 0$, $i = 1, 2, \ldots, n-k$. Introduce $(2n-k)$-dimensional space with the independent coordinates q^a, λ^i, and consider the action

$$S = \int d\tau \left[L(q^a, \dot{q}^a) + \lambda^i G_i(q^a) \right]. \tag{1.203}$$

Since the initial coordinates are untouched, the formulation does not spoil the symmetry properties of a theory. The price we pay is the appearance of the additional variables λ^i that have no direct physical interpretation. They do not participate in determining the configuration of the system (positions, velocities, energies of the particles and so on). So we can not (and need not!) measure them. For this reason they are called *unphysical* (or *auxiliary*) degrees of freedom.

In all other respects, we treat the auxiliary variables on equal footing with others. In particular, we write and solve equations of motion for both q^a and λ^i.

Since the action does not contain derivatives of λ, it represents an example of a singular theory. The Hessian matrix has vanishing blocks, $\frac{\partial^2 S}{\partial\dot{\lambda}^i \partial\dot{\lambda}^j} = 0$, $\frac{\partial^2 S}{\partial\dot{\lambda}^i \partial\dot{q}^a} = 0$. So its rank is less than dimension $2n-k$ of extended space.

We demonstrate now that the new formulation (1.203) implies the same evolution for q^a as the old one (1.184). Hence they are equivalent. It will be convenient to write separately equations of motion for the variables q^i and q^α (these were described in Eq. (1.187)). Applying the principle of least action, we find

$$\frac{\partial S}{\partial q^\alpha} \equiv \frac{\partial L}{\partial q^\alpha} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{q}^\alpha} + \lambda^j \frac{\partial G_j}{\partial q^\alpha} = 0, \tag{1.204}$$

$$\frac{\partial S}{\partial q^i} \equiv \frac{\partial L}{\partial q^i} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{q}^i} + \lambda^j \frac{\partial G_j}{\partial q^i} = 0, \tag{1.205}$$

$$\frac{\partial S}{\partial \lambda^i} \equiv G_i(q^i, q^\alpha) = 0. \tag{1.206}$$

From this system we get closed equations for q^α. They can be obtained as follows. We suppose that the solution $q^i = q^i(q^\alpha)$ of (1.206) has been substituted into (1.204) and (1.205). Differentiation of the identity $G_j(q^i(q^\alpha), q^\alpha) = 0$ gives $\frac{\partial G_j}{\partial q^\alpha} = -\frac{\partial G_j}{\partial q^i}\frac{\partial q^i}{\partial q^\alpha}$. Using this expression in (1.204), the latter reads

$$\left(\frac{\partial L}{\partial q^\alpha} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{q}^\alpha} \right)\Bigg|_{q^i(q^\alpha)} - \lambda^j \left. \frac{\partial G_j}{\partial q^i} \right|_{q^i(q^\alpha)} \frac{\partial q_i}{\partial q^\alpha} = 0. \tag{1.207}$$

Now Eq. (1.205) allows us to exclude the term $\lambda^j \frac{\partial G_j}{\partial q^i}$ from (1.207). The result is

$$\left(\frac{\partial L}{\partial q^\alpha} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{q}^\alpha}\right)\Bigg|_{q^i(q^\alpha)} + \left(\frac{\partial L}{\partial q^i} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{q}^i}\right)\Bigg|_{q^i(q^\alpha)} \frac{\partial q^i}{\partial q^\alpha} = 0, \quad (1.208)$$

which is precisely Eq. (1.189).

To be sure of the self-consistency of the new formulation, we discuss the structure of solutions of the entire system (1.208), (1.205) and (1.206). The last of these equations has been already solved by $q^i(q^\alpha)$. Let $q^\alpha = q^\alpha(\tau)$ be a solution to (1.208). We substitute the functions $q^\alpha(\tau)$, $q^i(q^\alpha(\tau))$ into Eq. (1.205). This gives an algebraic equation for determining λ^i. Since by construction $\det\frac{\partial G_j}{\partial q^i} \neq 0$, Eq. (1.205) can be resolved with respect to λ^i. Note a consequence: since all λ^i are determined algebraically, we need not impose initial conditions for the auxiliary variables.

1.11 Kinetic Versus Potential Energy: Forceless Mechanics of Hertz

Here we discuss one more example of using an auxiliary variable. In this case, it allows us to reformulate potential motion in n-dimensional configuration space as a free fall in (fictitious) space of $n+1$-dimensions. This explains the terminology "forceless mechanics" for the new formulation developed by H. Hertz [21].

Take a system with the generalized coordinates q^a, $a = 1, 2, \ldots, n$ and potential $U(q^a)$

$$S[q^a] = \int d\tau \left[\frac{1}{2}(\dot{q}^a)^2 - U(q)\right]. \quad (1.209)$$

This leads to the equations

$$\ddot{q}^a = -\frac{\partial U}{\partial q^a}. \quad (1.210)$$

We could also work using the generalized coordinates with the non-trivial metric $g_{ab}(q)$, as in Eq. (1.130); this would not alter the final results.

We introduce $n+1$-dimensional space with the coordinates $q^i \equiv (q^a, q^{n+1})$ and write the following potential-free action

$$S[q^i] = \int d\tau \frac{1}{2} g_{ij}(q)\dot{q}^i \dot{q}^j \equiv \quad (1.211)$$

$$\int d\tau \left[\frac{1}{2}(\dot{q}^a)^2 + \frac{1}{4U}\dot{q}^{n+1}\dot{q}^{n+1}\right]. \quad (1.212)$$

This looks like the action of a *free* particle in generalized coordinates (see (1.130)), with a metric that has only a $g_{n+1\ n+1}$ -component nontrivial

$$g_{ij} = \begin{pmatrix} 1 & 0 \dots 0 & 0 \\ 0 & 1 \dots 0 & 0 \\ \vdots & \vdots \ddots \vdots & \vdots \\ 0 & 0 \dots 1 & 0 \\ 0 & 0 \dots 0 & \frac{1}{2U} \end{pmatrix}. \tag{1.213}$$

As compared with the initial formulation, the potential is now hidden in the kinetic term. We impose initial conditions both for q^a and q^{n+1} as follows:

$$\begin{aligned} q^a(0) &= q_0^a, & \dot{q}^a(0) &= v_0^a, \\ q^{n+1}(0) &= q_0^{n+1}, & \dot{q}^{n+1}(0) &= 2U(q_0^a). \end{aligned} \tag{1.214}$$

Due to the special choice of initial condition for $\dot{q}^{n+1}$, the formulation (1.211), (1.214) leads to the same equations of motion for q^a as the initial one (1.209). To see this, write equations of motion for (1.211)

$$\ddot{q}^a = -\frac{1}{4U^2}\frac{\partial U}{\partial q^a}(\dot{q}^{n+1})^2, \tag{1.215}$$

$$\left(\frac{1}{2U}\dot{q}^{n+1}\right)^{\cdot} = 0 \quad \Rightarrow \quad \dot{q}^{n+1} = 2cU(q^a). \tag{1.216}$$

Initial condition for $\dot{q}^{n+1}$ implies $c = 1$. Substituting the result, $\dot{q}^{n+1} = 2U$, into (1.215) we obtain Eq. (1.210) of the initial formulation.

In the new formulation the potential energy acquires a geometric origin. Recall that according to Sect. 1.6, the metric that appears in Eq. (1.211) has a certain geometric interpretation: it determines distances between points of configuration space. As will be explained in Chap. 6, trajectories of a theory with such an action also have a remarkable geometric interpretation: they represent lines of minimal length with respect to the metric g_{ij}. Hence they are similar to straight lines of Euclidean space and are called *geodesic lines*. It is known that trajectories of particles in general relativity theory have the same property. So, the motion of the q^i-particle described by (1.211) is analogous to free fall in a gravitational field. An intuitive picture of motion in the new formulation is presented in Fig. 1.22 on page 56. A fictitious q^i-particle moves freely along the shortest line of $(n+1)$-dimensional space with the metric $g_{ij}(U)$. Physical trajectory is its projection on configuration space q^a and corresponds to potential motion, with the potential being $U(q^a)$.

Since q^{n+1} represents an auxiliary variable, we are not able to experimentally fix its initial condition q_0^{n+1}. Fortunately, this does not lead to inconsistency: since equations for q^a and q^{n+1} have been separated, different choices of q_0^{n+1} imply the same physical dynamics, which is given by (1.210). Geometrically, solutions

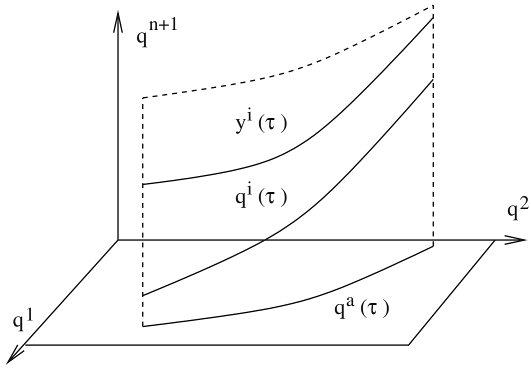

Fig. 1.22 According to Hertz, free fall in $n+1$-dimensional curved space is equivalent to potential motion in n-dimensional space

in extended space that correspond to different choices of q_0^{n+1} project on the same physical trajectory, see Fig. 1.22 on page 56.

Although we have discussed the case of a scalar potential, the construction can be adapted for the vector potential as well. An example of a force with a vector potential is electromagnetic force (with the vector potential being A_μ). Being appropriately generalized for that case, the construction leads to the *Kaluza–Klein theory*, which formally unifies four-dimensional gravity and electromagnetism into a unique five-dimensional theory [22].

1.12 Electromagnetic Field in Lagrangian Formalism

Here we apply Lagrangian formalism to the analysis of Maxwell equations describing electric and magnetic phenomena. While it is not evident in the initial formulation, Maxwell equations obey the principle of special relativity, that is they are covariant under the Poincaré transformations. We start from a description of an electromagnetic field in terms of a three-dimensional vector potential. In this case electrodynamics can be formulated on a base of nonsingular Lagrangian action. Like the Maxwell equations themselves, this three-dimensional formalism is not manifestly Poincaré invariant. We then go on to discuss the manifestly invariant formulation. This is achieved in terms of a four-dimensional vector potential and implies the use of singular Lagrangian action.

1.12.1 Maxwell Equations

Moving electric charges can be described using the charge density $\rho(t, x^a)$ and the current density vector $\vec{J}(t, x^a) = \rho(t, x^a)\vec{v}(t, x^a)$, where $\vec{v}$ is the velocity of a charge at t, x^a. They produce the electric $\vec{E}(t, x^a)$ and the magnetic $\vec{B}(t, x^a)$ fields. The fields obey the *Maxwell equations*

$$\frac{1}{c}\frac{\partial \vec{E}}{\partial t} - [\vec{\nabla}, \vec{B}] = -\frac{1}{c}\vec{J}, \tag{1.217}$$

$$\frac{1}{c}\frac{\partial \vec{B}}{\partial t} + [\vec{\nabla}, \vec{E}] = 0, \tag{1.218}$$

$$(\vec{\nabla}, \vec{E}) = \rho, \tag{1.219}$$

$$(\vec{\nabla}, \vec{B}) = 0. \tag{1.220}$$

We use the following notation:

$$\begin{aligned}
\vec{\nabla} &= \left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right) = (\partial_1, \partial_2, \partial_3),\\
\text{divergence}: \quad (\vec{\nabla}, \vec{E}) &= \partial_1 E_1 + \partial_2 E_2 + \partial_3 E_3,\\
\text{curl}: \quad [\vec{\nabla}, \vec{E}]_a &= \epsilon_{abc}\partial_b E_c,\\
\text{gradient}: \quad \vec{\nabla}\alpha &= (\partial_1\alpha, \partial_2\alpha, \partial_3\alpha),\\
\triangle\rho &= (\partial_1^2 + \partial_2^2 + \partial_3^2)\rho.
\end{aligned} \tag{1.221}$$

Examples of solutions

1. **Electric field of a point charge.** Consider the charge q at rest at the origin of a coordinate system. Then $\vec{J} = 0$, while the charge density can be described as $\rho = q\delta^3(\vec{x})$, where $\delta^3(\vec{x})$ is the Dirac δ-function. Then the total charge is $Q = \int d^3x\rho = q$. Maxwell equations with these densities are solved by

$$\vec{E} = \frac{q\vec{x}}{|\vec{x}|^3}, \qquad (\text{then } |\vec{E}| = \frac{q}{|\vec{x}|^2}), \qquad \vec{B} = 0. \tag{1.222}$$

   This is the Coulomb law: the electric field of a point charge is spherically symmetric, directed to the charge, with its strength equal to the inverse square of the distance to the charge.
   When $\vec{x} \neq 0$, the direct computation gives $\partial_a \frac{qx^a}{|\vec{x}|^3} = 0$. So the Maxwell equation $(\vec{\nabla}, \vec{E}) = q\delta^3(\vec{x})$ is satisfied for any $\vec{x} \neq 0$. For $\vec{x} = 0$ it holds as an equality among generalized functions, $\int d^3x\partial_a(\frac{qx^a}{|\vec{x}|^3})f(\vec{x}) = \int d^3xq\delta^3(x)f(\vec{x})$. This can be proved after a careful definition of the function $\frac{x^a}{|\vec{x}|^3}$ at $\vec{x} = 0$; see, for example, [23]. The field of a moving charge will be obtained in Sect. 1.12.7.
2. **Magnetic field of a straight wire.** Consider a neutral wire placed along the x^3-axis, with the current density $\vec{J} = \big(0,\ 0,\ \rho V\delta(x^1)\delta(x^2)\big)$, where ρ is the linear density of moving charges. The Maxwell equations are

solved by

$$\vec{E} = 0,$$
$$\vec{B} = \left(-\frac{1}{c}\rho V x^2[(x^1)^2 + (x^2)^2],\ \frac{1}{c}\rho V x^1[(x^1)^2 + (x^2)^2],\ 0\right)$$
$$= \frac{\rho[\vec{V}, \vec{R}]}{c|\vec{R}|^2}, \tag{1.223}$$

where $\vec{R} = (x^1, x^2, 0)$. The solution is drawn in Fig. 1.23a on page 58.

3. **Electromagnetic wave.** Maxwell equations in absence of the charge and the current densities

$$\partial_\tau \vec{E} - c[\vec{\nabla}, \vec{B}] = 0, \quad \partial_\tau \vec{B} + c[\vec{\nabla}, \vec{E}] = 0, \quad (\vec{\nabla}, \vec{E}) = 0,$$
$$(\vec{\nabla}, \vec{B}) = 0, \tag{1.224}$$

admit nontrivial solutions called waves. An example is

$$\vec{E} = (0, D\sin(\omega x^1)\sin(c\omega t), 0), \quad \vec{B} = (0, 0, D\cos(\omega x^1)\cos(c\omega t)), \tag{1.225}$$

where D, ω are constants. The solution consists of two standing waves (see Sect. 1.7), the E-wave in the (x^1, x^2)-plane, and the B-wave in the (x^1, x^3)-plane; see Fig. 1.23b on page 58. When the amplitude of $\vec{E}$ decreases, the amplitude of $\vec{B}$ increases and vice versa. In particular, at the instances $t = \frac{k\pi}{c\omega}$ we have $\vec{E} = 0$, $\vec{B} = \pm D\cos(\omega x^1)$. That is, when $\vec{E}$ vanishes, $\vec{B}$ has its maximum amplitude at each space point. Intuitively, the fields $\vec{E}$ and $\vec{B}$ turn into one another during the time evolution.

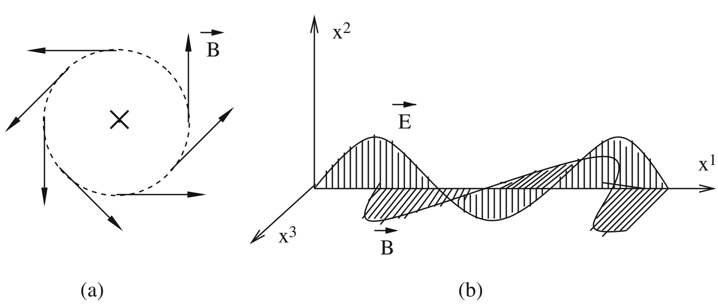

Fig. 1.23 **(a)** The magnetic field is tangent to a circle surrounding a wire. **(b)** An electromagnetic wave consists of two standing waves

Consistency condition. Conservation of charge. Maxwell equations imply an important consistency condition on ρ, $\vec{J}$ called the continuity equation. To find it, take the divergence of both sides of Eq. (1.217). Using the identity $(\vec{\nabla}, [\vec{\nabla}, \vec{B}]) = 0$ as well as Eq. (1.219), we obtain

$$\partial_t \rho + (\vec{\nabla}, \vec{J}) = 0. \tag{1.226}$$

So, the charge and current densities can not be taken as arbitrary, but must obey the continuity equation. To see the meaning of this, we write it in the integral form. Let us integrate both sides of the continuity equation over a volume V surrounded by the closed surface S

$$\partial_t \int_V d^3x\rho = -\int_V d^3x(\vec{\nabla}, \vec{J}) = -\int_S (\vec{J}, d\vec{S}). \tag{1.227}$$

The last equality is due to the Gauss theorem. So we have

$$\partial_t Q_V = -\int_S (\vec{J}, d\vec{S}), \tag{1.228}$$

where Q_V stands for the charge contained in the volume V. Hence the continuity equation states that an electric charge must be locally conserved: the rate of variation of a charge in a volume is equal to its flow (amount of charge that leaves or enters through the surface per unit of time).

If we integrate over all space, the right-hand side vanishes owing to $\vec{J} \stackrel{x^a \to \infty}{\longrightarrow} 0$ (charges can not escape to infinity). Then $\partial_t Q = 0$, that is, the total electric charge is conserved.

1.12.2 Nonsingular Lagrangian Action of Electrodynamics

The aim of this section is to show that electromagnetic forces can be described in terms of a unique vector field (called a three-dimensional vector potential) instead of $\vec{E}$, $\vec{B}$. This obeys a second-order partial differential equation, which can be obtained by applying the least action principle to a nonsingular Lagrangian action.

According to Maxwell, an electromagnetic field is described by six functions $\vec{E}$, $\vec{B}$ subject to eight equations. There are six equations of the first order with respect to time, (1.217) and (1.218). Two more Eqs. (1.219) and (1.220) do not involve the time derivative and hence represent the field analogy of kinematic constraints. We first reduce the number of equations from 8 to 6. A specific property of the Maxwell system is that the constraint equations can be replaced by properly-chosen initial conditions for the problem. Indeed, consider the following problem

$$\frac{1}{c}\frac{\partial\vec{E}}{\partial t} - [\vec{\nabla}, \vec{B}] = -\frac{1}{c}\vec{J}, \tag{1.229}$$

$$\frac{1}{c}\frac{\partial\vec{B}}{\partial t} + [\vec{\nabla}, \vec{E}] = 0, \tag{1.230}$$

with the initial conditions

$$\left[(\vec{\nabla}, \vec{E}) - \rho\right]\Big|_{t=0} = 0, \qquad (\vec{\nabla}, \vec{B})\Big|_{t=0} = 0. \tag{1.231}$$

This is equivalent to the problem (1.217), (1.218), (1.219), and (1.220). Any solution to (1.217), (1.218), (1.219), and (1.220) satisfies the Eqs. (1.229), (1.230), (1.231). Conversely, let $\vec{E}$, $\vec{B}$ be the solution to the problem (1.229), (1.230), and (1.231). Taking the divergence of Eq. (1.229), we obtain the consequence $\partial_t(\vec{\nabla}, \vec{E}) + (\vec{\nabla}, \vec{J}) = \partial_t[(\vec{\nabla}, \vec{E}) - \rho] = 0$. The initial condition (1.231)) then implies $(\vec{\nabla}, \vec{E}) - \rho = 0$, that is, Eq. (1.219). In the same way, taking the divergence of Eq. (1.230) we arrive at Eq. (1.220).

To proceed further, it is convenient to unify the vectors $\vec{E}$, $\vec{B}$ into the complex field

$$\vec{W} \equiv \vec{B} + i\vec{E}. \tag{1.232}$$

Then Eqs. (1.229), (1.230), and (1.231) can be written in a more compact form

$$\left(\frac{i}{c}\partial_t + \vec{\nabla}\times\right)\vec{W} = \frac{1}{c}\vec{J}, \qquad \left[(\vec{\nabla}, \vec{W}) - i\rho\right]\Big|_{t=0} = 0. \tag{1.233}$$

If we look for a solution to the form $\vec{W} = \left(-\frac{i}{c}\partial_t + \vec{\nabla}\times\right)\vec{A}$, the equations that appear for $\vec{A}$ turns out to be real. They read

$$\frac{1}{c^2}\partial_t^2\vec{A} - \triangle\vec{A} + \vec{\nabla}(\vec{\nabla}, \vec{A}) = \frac{1}{c}\vec{J}, \tag{1.234}$$

$$\left[\partial_t(\vec{\nabla}, \vec{A}) + c\rho\right]\Big|_{t=0} = 0, \tag{1.235}$$

where we have used the identity

$$\epsilon_{cab}\epsilon_{cmn} = \delta_{am}\delta_{bn} - \delta_{an}\delta_{bm}, \quad \text{then}$$

$$[\vec{\nabla}, [\vec{\nabla}, \vec{A}]] = -\triangle\vec{A} + \vec{\nabla}(\vec{\nabla}, \vec{A}). \tag{1.236}$$

Hence it is consistent to take $\vec{A}$ as a real function. Thus, any real solution $\vec{A}(t, x)$ of Eq. (1.234) determines a solution

$$\vec{B} = [\vec{\nabla}, \vec{A}], \qquad \vec{E} = -\frac{1}{c}\partial_t\vec{A}, \tag{1.237}$$

of the Maxwell equations. $\vec{A}$ is called a (three-dimensional) *vector potential* of the electromagnetic field[11].

Conversely, any given solution $\vec{E}$, $\vec{B}$ of the Maxwell equations can be written in the form (1.237), where $\vec{A}$ is a solution to the problem (1.234), (1.235). To see this, construct the field

$$\vec{A}(t, x^a) = -c \int_0^t d\tau \vec{E}(\tau, x^a) + \vec{K}(x^a), \tag{1.238}$$

where K is any solution to the equation

$$[\vec{\nabla}, \vec{K}] = \vec{B}(0, x^a). \tag{1.239}$$

The existence of the solution K is guaranteed by the Eq. (1.231). By direct substitution, we can verify that the field constructed obeys the Eqs. (1.237), (1.234) and (1.235).

It is not difficult to construct a Lagrangian action that implies equations of motion (1.234) for the vector potential. It is

$$\begin{aligned} S &= \int dt d^3x \left[\frac{1}{2c^2}(\partial_t \vec{A}, \partial_t \vec{A}) - \frac{1}{2}([\vec{\nabla}, \vec{A}], [\vec{\nabla}, \vec{A}]) + \frac{1}{c}(\vec{A}, \vec{J}) \right] \\ &\equiv \int dt d^3x \left[\frac{1}{2c^2} \partial_t A_a \partial_t A_a - \frac{1}{4}(\partial_a A_b - \partial_b A_a)^2 + \frac{1}{c} A_a J_a \right] \\ &\equiv \int dt d^3x \left[\frac{1}{2}(\vec{E}^2 - \vec{B}^2) + \frac{1}{c}(\vec{A}, \vec{J}) \right]. \end{aligned} \tag{1.240}$$

In the last line the substitution (1.237) is implied. To confirm that this leads to the desired equations, let us compute the variation of the action

$$\delta S = \int dt d^3x \left[\frac{\partial_t A_a}{c^2} \partial_t \delta A_a - (\partial_b A_a - \partial_a A_b) \partial_b \delta A_a + \frac{J_a}{c} \delta A_a \right]. \tag{1.241}$$

Using integration by parts, we can extract δA_a, obtaining the expression

$$\begin{aligned} \delta S = \int dt d^3x \Bigg[& \left(-\frac{1}{c^2} \partial_t^2 A_a + (\partial_b \partial_b A_a - \partial_a \partial_b A_b) + \frac{1}{c} J_a \right) \delta A_a + \\ & \frac{1}{c^2} \partial_t (\partial_t A_a \delta A_a) - \partial_b (\partial_b A_a \delta A_a - \partial_a A_b \delta A_a) \Bigg]. \end{aligned} \tag{1.242}$$

The total-derivative terms do not contribute to the variation due to the boundary conditions of the variational problem

[11] A similar procedure will be used in Sect. 2.10 where we obtain a scalar potential for the quantum mechanical wave function.

$$\int dtd^3x\partial_t(\partial_t A_a\delta A_a) = \int d^3x\ [\partial_t A_a\delta A_a]|_{t_1}^{t_2} = 0 \quad \text{since} \quad \delta A_a(t_i, x) = 0,$$
$$\int dtd^3x\partial_b(\partial_b A_a\delta A_a - \partial_a A_b\delta A_a) = \int dt(\partial_b A_a\delta A_a - \partial_a A_b\delta A_a)\, dS_b = 0 \quad \text{since} \quad A_a \xrightarrow{x^a\to\infty} 0. \quad (1.243)$$

Then the extremum condition

$$\delta S = \int dtd^3x\left[-\frac{1}{c^2}\partial_t^2 A_a + \partial_b\partial_b A_a - \partial_a\partial_b A_b + \frac{1}{c}J_a\right]\delta A_a = 0, \quad (1.244)$$

implies the Eqs. (1.234).

In short, an electromagnetic field can be described starting from the nonsingular action

$$S = \int dtd^3x\left[\frac{1}{2c^2}(\partial_t\vec{A}, \partial_t\vec{A}) - \frac{1}{2}([\vec{\nabla}, \vec{A}], [\vec{\nabla}, \vec{A}]) + \frac{1}{c}(\vec{A}, \vec{J})\right], \quad (1.245)$$

which implies the equations of motion

$$\frac{1}{c^2}\partial_t^2\vec{A} - \triangle\vec{A} + \vec{\nabla}(\vec{\nabla}, \vec{A}) = \frac{1}{c}\vec{J}. \quad (1.246)$$

They must be solved under the initial condition $[\partial_t(\vec{\nabla}, \vec{A}) + c\rho]\Big|_{t=0} = 0$. Any solution $\vec{A}(t, x)$ to the problem determines the solution

$$\vec{B} = [\vec{\nabla}, \vec{A}], \qquad \vec{E} = -\frac{1}{c}\partial_t\vec{A}, \quad (1.247)$$

to the Maxwell equations (1.217), (1.218), (1.219), and (1.220). Note that the solutions $\vec{A}(t, x^a)$ and $\vec{A}(t, x^a) + \vec{\nabla}\alpha(x^a)$, where $\alpha(x^a)$ is an arbitrary function, determine the same $\vec{E}$, $\vec{B}$. Any given solution to the Maxwell equations can be presented in the form (1.247), with the potential $\vec{A}$ given in Eq. (1.238).

Exercises

1. Show that the problem

$$G_a \equiv \frac{1}{c^2}\partial_t^2\vec{A} - \triangle\vec{A} + \vec{\nabla}(\vec{\nabla}, \vec{A}) - \frac{1}{c}\vec{J} = 0, \quad D \equiv \partial_t(\vec{\nabla}, \vec{A}) + c\rho = 0, \tag{1.248}$$

is equivalent to the problem (1.234), (1.235).
2. Verify the identity $\partial_t D - c^2(\vec{\nabla}, \vec{G}) \equiv 0$ among these equations.
3. Confirm that the action (1.245) is invariant under the transformation $\vec{A} = \vec{A}' + \vec{\nabla}\alpha$, where $\alpha(x^a)$ is an arbitrary function.

Non-relativistic particle on an electromagnetic background. The action of a charged particle on a given electromagnetic field with potential A_a is

$$S = \int dt \left[\frac{m}{2}(\dot{x}^a)^2 + \frac{e}{c}A_a(t, x^b)\dot{x}^a\right]. \tag{1.249}$$

The particle placed at the point $x^a(t)$ interacts with the potential at that point, $A_a(t, x^b(t))$.

Recall the arbitrariness $A_a \to A_a + \partial_a\alpha(x^b)$ presented in the definition of the vector potential. Replacing A_a by $A_a + \partial_a\alpha(x^b)$ in the action, we obtain an extra term that is the total derivative $\partial_a\alpha\dot{x}^a = \dot{\alpha}$. Hence it does not modify equations of motion.

The variation of the action reads

$$\delta S = \int dt \left[-m\ddot{x}^a - \frac{e}{c}\partial_t A_a + \frac{e}{c}(\partial_a A_b - \partial_b A_a)\dot{x}^b\right]\delta x^a. \tag{1.250}$$

Taking into account Eq. (1.247), the extremum condition $\delta S = 0$ implies the well-known equations of motion

$$m\ddot{\vec{r}} = e\vec{E} + \frac{e}{c}[\dot{\vec{r}}, \vec{B}]. \tag{1.251}$$

1.12.3 *Manifestly Poincaré-Invariant Formulation in Terms of a Singular Lagrangian Action*

We start by presenting the Maxwell equations (1.217), (1.218), (1.219), and (1.220) in terms of a four-component quantity $A_\mu = (A_0, A_a)$ called a (four-dimensional) *vector potential*. Consider first the homogeneous Eqs. (1.220) and (1.218). Since a magnetic field has zero divergence, $(\vec{\nabla}, \vec{B}) = 0$, it can be presented as a curl of a vector, $\vec{B} = [\vec{\nabla}, \vec{A}]$ (this is proved at the end of this section). Substituting this expression into Eq. (1.218), the latter reads $[\vec{\nabla}, \frac{1}{c}\partial_t\vec{A} + \vec{E}] = 0$. A field with zero curl can be presented as a gradient of a function, so $\frac{1}{c}\partial_t\vec{A}+\vec{E}=\vec{\nabla}A_0$. Hence any given solution of the homogeneous Maxwell equations can be presented as

$$\vec{E}=-\frac{1}{c}\frac{\partial\vec{A}}{\partial t}+\vec{\nabla}A_0, \qquad \vec{B}=[\vec{\nabla},\vec{A}], \tag{1.252}$$

through some functions A_0, $\vec{A}$. It is convenient to use four-dimensional notations introducing $x^\mu=(x^0\equiv ct,x^a)$, $A_\mu=(A_0,A_a)$. Besides, we define an anti-symmetric matrix called the field strength of the vector potential

$$F_{\mu\nu}\equiv\partial_\mu A_\nu-\partial_\nu A_\mu. \tag{1.253}$$

Then Eq. (1.252) acquires the form

$$-E_a=F_{0a}, \qquad B_a=\frac{1}{2}\epsilon_{abc}F_{bc}, \quad (\text{then } F_{ab}=\epsilon_{abc}B_c), \tag{1.254}$$

that is $\vec{E}$ and $\vec{B}$ can be identified with components of the field strength matrix

$$F_{\mu\nu}=\begin{pmatrix}0 & -E^1 & -E^2 & -E^3\\ E^1 & 0 & B^3 & -B^2\\ E^2 & -B^3 & 0 & B^1\\ E^3 & B^2 & -B^1 & 0\end{pmatrix}. \tag{1.255}$$

Exercise
Show that the homogeneous Maxwell equations in these notations reads $\partial_\mu F_{\nu\rho}+\partial_\nu F_{\rho\mu}+\partial_\rho F_{\mu\nu}=0$. As should be the case, they are satisfied by (1.253).

Substituting Eq. (1.253) into the inhomogeneous Eqs. (1.219) and (1.217), they read $\partial_\mu F^{\mu 0}=-\rho$, $\partial_\mu F^{\mu a}=-\frac{1}{c}J^a$, where $F^{\mu\nu}=\eta^{\mu\alpha}\eta^{\nu\beta}F_{\alpha\beta}$. Denoting $J^\mu=(\rho,\frac{1}{c}J^a)$, they can be written in four-dimensional form

$$S^\nu\equiv\partial_\mu F^{\mu\nu}+J^\nu=0, \quad \text{or} \quad \partial_\mu\partial^\mu A^\nu-\partial^\nu(\partial_\mu A^\mu)=-J^\nu. \tag{1.256}$$

They follow from the action

$$S=\int d^4x\left[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}+A_\mu J^\mu\right], \quad F_{\mu\nu}\equiv\partial_\mu A_\nu-\partial_\nu A_\mu. \tag{1.257}$$

Exercise
Obtain (1.256) from (1.257).

Poincaré invariance. Let us postulate that A_μ and J^μ are Minkowski - space vector fields. Under the Poincaré transformations (1.51) they transform as

$$A'_\mu(x') = A_\nu(x)\tilde{\Lambda}^\nu{}_\mu, \qquad J'^\mu(x') = \Lambda^\mu{}_\nu J^\nu(x), \tag{1.258}$$

Then $F_{\mu\nu}$ is the second-rank covariant vector and the Lagrangian $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + A_\mu J^\mu$ is a scalar function, see Sect. 1.3. Hence the action (1.257) is invariant[12] under the Poincaré transformations, while equations of motion (1.256) are manifestly Poincaré-covariant.

According to the terminology introduced in Sect. 1.6, Poincaré transformations represent a family of global symmetries (with ten parameters) of the action (1.257).

We also point out that invariance of a Lagrangian action implies covariance of the corresponding equations of motion; see Sect. 7.5.

To conclude with, we discuss two statements that were used at the beginning of this section.

Given field $\vec{B}$ with zero curl, $[\vec{\nabla}, \vec{B}] = 0$, there is a function φ such that $\vec{B} = \vec{\nabla}\varphi$. To construct it, let us fix the point $\vec{x}_0$, and let $\vec{\gamma}(l, \vec{x})$ be a curve connecting $\vec{x}_0$ with a point $\vec{x}$, that is, $\vec{\gamma}(0, \vec{x}) = \vec{x}_0$, $\vec{\gamma}(1, \vec{x}) = \vec{x}$ (take, for example, $\vec{\gamma}(l, \vec{x}) = \vec{x}_0 + l(\vec{x} - \vec{x}_0)$). Then

$$\varphi = \int_0^1 B_a(\vec{\gamma}(l, \vec{x}))\frac{\partial \gamma_a(l, \vec{x})}{\partial l}dl. \tag{1.259}$$

By direct computations, we verify that $\vec{\nabla}\varphi = \vec{B}$. Note also that the integral in Eq. (1.259) is just the line integral of the vector function, $\varphi = \int_\gamma (\vec{B}, \vec{dl})$.

Given field $\vec{B}$ with zero divergence, $(\vec{\nabla}, \vec{B}) = 0$, there is a vector $\vec{A}$ such that $\vec{B} = [\vec{\nabla}, \vec{A}]$. It is given by

$$A_a = \int_0^1 \frac{\gamma_b(l, \vec{x})}{\partial l}\frac{\gamma_c(l, \vec{x})}{\partial x^a}\epsilon_{bcd}B_d(\vec{\gamma}(l, \vec{x}))dl + \partial_a\alpha, \tag{1.260}$$

where α is an arbitrary function.

1.12.4 Notion of Local (Gauge) Symmetry

From Eqs. (1.253) and (1.254) it follows that a potential of the form $\bar{A}_\mu \equiv \partial_\mu\alpha$, where $\alpha(x^\mu)$ is an *arbitrary* space-time function, leads to zero field strength, $\bar{F}_{\mu\nu} = 0$, so it does not produce electric and magnetic fields. As a consequence, the potentials A_μ and $A'_\mu = A_\mu + \partial_\mu\alpha$, determine the same electromagnetic field.

---

[12] Remember that according to mathematical analysis, an integration measure changes as $d^4x' = |\det\Lambda| d^4x$.

The ambiguity in the representation of an electromagnetic field through a potential can be used for various reformulations of equations of motion (1.256).

For instance, for a given potential A_μ, the function α can be chosen such that $A'_0 = 0$ (it is sufficient to solve the equation $\partial_0\alpha = -A_0$). Hence there is a potential of the form $A'_\mu = (0, A'_a)$, that produces the same electromagnetic field as A_μ. Hence, if we wish, we can look for solutions of this form resolving the Eqs. (1.256). Knowledge of this solution of a special form is sufficient to reconstruct the electromagnetic field created by a given distribution of charges. When $A_0 = 0$, the Eqs. (1.256) reduce to the system (1.248) obtained in the previous section[13].

The procedure described is called *fixation of a gauge*. The condition $A_0 = 0$ is known as a *unitary gauge*.

To make the relationship between $(\vec{E}, \vec{B})$ and A_μ clearer, we define an equivalence relation on the space of functions $A_\mu(x^\nu)$. The potentials A_μ and A'_μ are equivalent, $A' \sim A$, if they differ by the divergence of some function, $A'_\mu = A_\mu + \partial_\mu\alpha$. Equivalent potentials form a set called a class of equivalent potentials. It is denoted by $\tilde{A}_\mu = \{A_\mu + \partial_\mu\alpha,\ \text{where } \alpha \text{ is an arbitrary function}\}$. According to the known theorems of algebra, given the equivalence relation, the initial space decomposes on *non-intersecting* classes of equivalent potentials, see Fig. 1.24 on page 66. Electromagnetic fields are in one-to-one correspondence with the classes, $(\vec{E}, \vec{B}) \leftrightarrow \tilde{A}_\mu$. As we have discussed, in each class there is a representative of the form $(0, A_a)$. Fixation of the unitary gauge means that solving equations of motion we look for a representative of this special form in each class.

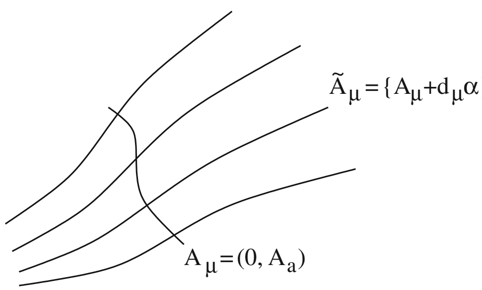

Fig. 1.24 The space of potentials decomposes into non-intersecting classes $\tilde{A}_\mu$ of equivalent potentials. Electromagnetic fields are in one-to-one correspondence with the classes, $(\vec{E}, \vec{B}) \sim \tilde{A}_\mu$. In each class there are representatives of the form $\tilde{A}_\mu = (0,\ A_a)$

[13] The equivalent potential reads $A'_\mu = (A'_0, A'_a) = (0, A_a - \partial_a \int dx^0 A_0 + \partial_a c(x^a))$, where $c(x^a)$ stands for an arbitrary function. For the free electrodynamics, the arbitrariness can be used for further specification of the potential. For example, we can find A'_a with $\partial_a A'_a = 0$. It implies the following equation for c: $\partial_a^2 c(x^b) = -\partial_a A_a + \partial_a^2 \int dx^0 A_0$. The equation is consistent, since its r.h.s. does not depend on x^0: $\partial_0(-\partial_a A_a + \partial_a^2 \int dx^0 A_0) = -\partial_\mu F^{\mu 0} = 0$, and thus can be resolved. The potential then obeys the three-dimensional wave equation. For the more detailed discussion of this point see Sect. 8.4.

Another often-used gauge is the *Lorentz gauge*, $\partial_\mu A^\mu = 0$. One of its advantages is that it does not spoil the manifest Poincaré covariance of the problem. Another advantage is that in this gauge the Eq. (1.256) acquires the form of a wave equation (see Sect. 1.7)

$$\partial_\mu \partial^\mu A^\nu = -J^\nu. \tag{1.261}$$

Exercise
Find the solution to Eq. (1.261) that produces the electromagnetic wave of Example 3 on page 58.

The action (1.257) also reflects the ambiguity in the representation of an electromagnetic field through a potential: it has the same value for all representatives of a given class. Indeed, the substitution

$$A_\mu = A'_\mu - \partial_\mu \alpha, \tag{1.262}$$

leaves the action (1.257) invariant, $S[A(A')] = S[A']$. So it represents a symmetry transformation on the space of fields. In contrast to the global Poincaré symmetry discussed above, the symmetry (1.262) involves an arbitrary function $\alpha(x^\mu)$ instead of continuous parameters. Intuitively, the transformation law varies from one point to another, transformations of a potential "here" and "there" are different. For this reason the transformation is called *local (or gauge) symmetry* of the action. In what follows we discuss some characteristic properties of locally-invariant theories.

Singular character of the locally-invariant action. Let us separate the terms with temporal derivatives in the action (1.257)

$$S = \int d^4x \left[\frac{1}{2}(\partial_0 A_b - \partial_b A_0)^2 - \frac{1}{4} F_{ab}^2 + A_\mu J^\mu \right]. \tag{1.263}$$

Note that it does not contain the square of the time derivative of A_0. Hence the Hessian matrix (see Sect. 1.6) is degenerate

$$\frac{\partial^2 L}{\partial \dot{A}_\mu \partial \dot{A}_\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.264}$$

and we deal with a singular theory.

Arbitrariness in solutions of locally-invariant theory. Remember that a nonsingular theory leads to a normal system of equations, which admits formulation of the Cauchy problem. This is impossible in a locally-invariant theory: independently of the type and the number of initial and boundary conditions imposed, a solution to the equations of locally-invariant theory is not unique, and involves arbitrary functions. Indeed, suppose that A_μ represents a solution to the Eqs. (1.256) with some initial and boundary conditions imposed. Take a function $\alpha(x^\nu)$ such that $\partial_\mu\alpha$ (together with its time derivatives, if necessary), vanishes at the initial instance and on the boundary. Then $A_\mu + \partial_\mu\alpha$ obeys the same problem as A_μ.

Structure of equations of motion. One consequence of singularity is that not all equations of motion are of the second order with respect to a temporal variable. Let us write Eqs. (1.256) in three-dimensional notations

$$S_0 \equiv \triangle A_0 - \partial_0(\partial_b A_b) = \rho, \tag{1.265}$$

$$S_a \equiv (\partial_0^2 - \triangle)A_a - \partial_a(\partial_0 A_0 - \partial_b A_b) - \frac{1}{c}J_a = 0. \tag{1.266}$$

Equation (1.265) does not contain ∂_0^2. Another consequence is the identity presented among the equations

$$\partial_\nu S^\nu \equiv 0. \tag{1.267}$$

Finally, similarly to the three-dimensional formulation, Eq. (1.265) can be replaced by an appropriate initial condition. To see this, compute the divergence of Eq. (1.266), then it reads $\partial_0[\triangle A_0 - \partial_a(\partial_b A_b) - \rho] = 0$. So, Eq. (1.265) will hold at any instance if it has been satisfied at the initial instant of time. Hence the system (1.265), (1.266) is equivalent to Eq. (1.266) supplemented by the initial condition $[\triangle A_0 - \partial_a(\partial_b A_b) - \rho]\,|_{t=0} = 0$.

There is a profound relationship between the properties enumerated above and the local invariance of a theory. This will be discussed in Chap. 8.

1.12.5 Lorentz Transformations of Three-Dimensional Potential: Role of Gauge Symmetry

While it was mentioned above that the formulation in terms of three-dimensional potential A_b obeys the special relativity principle, it was not demonstrated for the present.

Consider two observers related by the Lorentz transformation $x'^\mu = \Lambda^\mu{}_\nu x^\nu$. Suppose they study a given electromagnetic field. As we have seen, the four-dimensional potentials are related by combination of the Lorentz and the gauge transformations

$$A'_\mu = A_\nu \tilde{\Lambda}^\nu{}_\mu + \partial_\mu\alpha. \tag{1.268}$$

Suppose the observers decided to use a three-dimensional formalism to describe the given field. Then O describes it by the potential $A_\mu = (A_0 = 0,\ A_a)$, while O'

uses $A'_\mu = (A'_0 = 0,\ A'_a)$. The question is if the formalism is a Lorentz covariant, that is, whether the two descriptions are related by a Lorentz transformation. It is clear that the linear Lorentz transformation of four-dimensional formulation does not relate these A' and A, since transformation of $(0, A_a)$ leads to a potential with $A'_0 = A_b \tilde{\Lambda}^b{}_0 \neq 0$. A general gauge transformation also implies $A'_0 \neq 0$.

It can be said that neither the Lorentz nor the gauge symmetry of four-dimensional formulation survive in the gauge $A_0 = 0$. But we can look for a combination that does not spoil the gauge condition. So, given the Lorentz transformation $\tilde{\Lambda}$, we ask whether there is a $\alpha(\tilde{\Lambda})$ such that $A'_0 = 0$.

Taking $\mu = 0$-component of Eq. (1.268) and requiring $A'_0 = 0$, we find $\alpha(\tilde{\Lambda}) = -C_c \tilde{\Lambda}^c{}_0$, where C_c is a primitive function of A_c, $\partial_0 C_c = A_c$. Then the Lorentz transformation of three-dimensional potential is

$$A'_b = A_c \tilde{\Lambda}^c{}_b - \partial_b C_c \tilde{\Lambda}^c{}_0. \tag{1.269}$$

The transformation turns out to be highly non-linear (it is non local with respect to A_a!), involving the primitive function C_c of the potential.

This computation shows the *role of auxiliary variables and local symmetries associated with them. Introducing the auxiliary variables, we arrive at the formulation where the initial non-linearly realized global symmetry decomposes on a linear global symmetry plus a local symmetry.*

1.12.6 Relativistic Particle on Electromagnetic Background

The free-particle Lagrangian $\frac{m}{2}(\dot{x}^a(t))^2$ does not determine a relativistic theory. It admits solutions $x^a = x^a_0 + v^a t$ with any velocity v^a, so it does not take into account the existence of a maximum speed in nature. To resolve the problem, let us consider the action

$$S = -mc \int dt \sqrt{c^2 - (\dot{x}^a)^2}. \tag{1.270}$$

It has the following properties.

1. The action has no meaning when $\dot{\vec{x}}^{\,2} > c^2$, so we do not expect it to admit motions with $v > c$. Indeed, this implies the equations

$$\frac{d}{dt}\left(\frac{\dot{x}^a}{\sqrt{c^2 - \dot{\vec{x}}^{\,2}}}\right) = 0, \quad \text{then} \quad \frac{\dot{x}^a}{\sqrt{c^2 - \dot{\vec{x}}^{\,2}}} = b^a. \tag{1.271}$$

   This implies $\frac{\dot{\vec{x}}^{\,2}}{c^2 - \dot{\vec{x}}^{\,2}} = \vec{b}^2$, then $\sqrt{c^2 - \dot{\vec{x}}^{\,2}} = \frac{c}{\sqrt{1+\vec{b}^2}}$. Using this equality in Eq. (1.271), it reads $\dot{x}^a = \frac{cb^a}{\sqrt{1+\vec{b}^2}}$, so the general solution is

$$x^a(t) = v^a t + x_0^a, \quad \text{where} \quad v^a = \frac{cb^a}{\sqrt{1+\vec{b}^2}}, \tag{1.272}$$

and b^a, x_0^a are arbitrary constants. The square of velocity is given by $\vec{v}^2 = \frac{c^2\vec{b}^2}{1+\vec{b}^2}$. It is less then c^2 for any integration constant b^a.

2. When $\dot{x}^2 \ll c^2$, we expand the Lagrangian in a power series obtaining $-mc\sqrt{c^2 - \dot{\vec{x}}^2} = -mc^2 + \frac{1}{2}m\dot{\vec{x}}^2 + O^2\left(\frac{\dot{\vec{x}}^2}{c^2}\right)$, so in the nonrelativistic limit $c \to \infty$ it reduces to the standard Lagrangian $\frac{1}{2}m\dot{\vec{x}}^2$.
3. The action is a Poincaré invariant. To see this, we follow the procedure discussed at the end of Sect. 1.3. Let $x^0 = x^0(\tau)$, $x^a = x^a(\tau)$ be parametric equations of the curve $x^a(t)$. Then $\frac{dx^a}{dt} = \frac{dx^a}{d\tau}/\frac{dt}{d\tau}$, and the action acquires a manifestly Poincaré-invariant form

$$S = -mc\int d\tau\sqrt{\eta_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}}. \tag{1.273}$$

Remember that $f(\dot{x}_\nu\dot{x}^\nu)$ is the only Poincaré-invariant quantity without the higher derivatives, see Sect. 1.3.

Exercises

1. Confirm that this is a singular action.
2. Show that the action is invariant under the local symmetry $\tau = f(\tau')$, $x^\mu(\tau) = x'^\mu(\tau')$. It has a simple geometric interpretation as a change of a parametrization of the particle trajectory, see Chap. 6. For this reason it is called *reparametrization symmetry*.

It is instructive to solve the manifestly covariant equations that result from this action. We have

$$S^\mu \equiv \frac{d}{d\tau}\left(\frac{\dot{x}^\mu}{\sqrt{\dot{x}_\nu\dot{x}^\nu}}\right) = 0, \quad \text{then} \quad \frac{\dot{x}^\mu}{\sqrt{\dot{x}_\nu\dot{x}^\nu}} = b^\mu, \tag{1.274}$$

where $\dot{x} = \frac{dx}{d\tau}$, and b^μ are arbitrary constants subject to the restriction $b_\mu b^\mu = 1$.

Exercise
Confirm the identity $\dot{x}_\mu S^\mu \equiv 0$.

We can verify that

$$x^\mu(\tau) = b^\mu f(\tau) + d^\mu, \tag{1.275}$$

where $f(\tau)$ is an *arbitrary* function, represents a solution to Eq. (1.274). Moreover, any solution to the problem has this form; see Exercise on page 198. Hence we have found all of them.

According to the ideology discussed at the end of Sect. 1.3, the physical trajectory $x^a(t)$ is obtained excluding τ from the parametric equations $x^0 = b^0 f(\tau)+d^0$, $x^a = b^a f(\tau) + d^a$. This leads to the expression

$$x^a(t) = v^a t + x_0^a, \quad \text{where } v^a = \frac{cb^a}{\sqrt{1+\vec{b}^2}}, \; x_0^a = d^a - \frac{d^0 b^a}{\sqrt{1+\vec{b}^2}}, \tag{1.276}$$

which reproduces our previous result (1.272).

The motion of the relativistic particle on a given electromagnetic background can be described starting from the action

$$S = \int d\tau \left[-mc\sqrt{\eta_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}} + \frac{e}{c}A_\mu \dot{x}^\mu \right]. \tag{1.277}$$

Exercise
Verify that the action is invariant both under the gauge and the reparametrization transformations.

This leads to the equations

$$mc\partial_\tau\left(\frac{\eta_{\mu\nu}\partial_\tau x^\nu}{\sqrt{\partial_\tau x_\nu \partial_\tau x^\nu}}\right) + \frac{e}{c}F_{\mu\nu}\partial_\tau x^\nu = 0, \tag{1.278}$$

or, in three-dimensional notations

$$\begin{aligned} mc\partial_\tau\left(\frac{\partial_\tau x^0}{\sqrt{\partial_\tau x_\nu \partial_\tau x^\nu}}\right) - \frac{e}{c}E_a \partial_\tau x^a = 0, \\ -mc\partial_\tau\left(\frac{\partial_\tau x^a}{\sqrt{\partial_\tau x_\nu \partial_\tau x^\nu}}\right) + \frac{e}{c}E_a\partial_\tau x^0 + \frac{e}{c}\epsilon_{abc}\partial_\tau x^b B^c = 0. \end{aligned} \tag{1.279}$$

Equations for the physical trajectory $x^a(t)$ can be obtained if we take $\frac{x^0}{c}$ as the parameter along the trajectory, $\tau = \frac{x^0}{c} = t$. Then $\partial_\tau x^\mu = (c, \; \dot{x}^a(t))$, and the equations acquire the form

$$\begin{aligned} G \equiv mc^3\frac{d}{dt}\left(\frac{1}{\sqrt{c^2-\dot{\vec{x}}^2}}\right) - eE_a\dot{x}^a = 0, \\ S^a \equiv mc\frac{d}{dt}\left(\frac{\dot{x}^a}{\sqrt{c^2-\dot{\vec{x}}^2}}\right) - eE_a - \frac{e}{c}\epsilon_{abc}\dot{x}^b B^c = 0. \end{aligned} \tag{1.280}$$

Note that $\dot{x}^a S^a \equiv G$, so the first equation is a consequence of others and can be omitted from consideration. At the non-relativistic limit, $\dot{\vec{x}}^2 \ll c^2$, the term $\dot{\vec{x}}^2$ inside the square root can be omitted, and the remaining three equations coincide with Eqs. (1.251) obtained above.

1.12.7 Poincaré Transformations of Electric and Magnetic Fields

Remember that E_a and B_a have been identified with matrix elements of the field strength $F_{\mu\nu}$, which transforms as

$$F'_{\mu\nu}(x') = F_{\alpha\beta}(x)\tilde{\Lambda}^{\alpha}{}_{\mu}\tilde{\Lambda}^{\beta}{}_{\nu}. \tag{1.281}$$

Using this equation as well as (1.254), we obtain the following expression for the transformation of an electromagnetic field

$$\begin{aligned} E'_a(x') &= E_d(x)(\tilde{\Lambda}^d{}_a\tilde{\Lambda}^0{}_0 - \tilde{\Lambda}^d{}_0\tilde{\Lambda}^0{}_a) + B_d(x)\epsilon_{dbc}\tilde{\Lambda}^b{}_a\tilde{\Lambda}^c{}_0, \\ B'_a(x') &= E_d(x)\tilde{\Lambda}^d{}_b\epsilon_{bca}\tilde{\Lambda}^0{}_c + \frac{1}{2}B_d(x)\epsilon_{dmn}\tilde{\Lambda}^m{}_b\tilde{\Lambda}^n{}_c\epsilon_{bca}. \end{aligned} \tag{1.282}$$

Here x' and x are related by the Poincaré transformation (1.51).

For a spacial rotation, the Lorentz matrix is $\tilde{\Lambda}^0{}_0 = 1$, $\tilde{\Lambda}^a{}_0 = \tilde{\Lambda}^0{}_a = 0$, $\tilde{\Lambda}^a{}_b = R_{ab}$, $R^T R = 1$. Substituting this into Eq. (1.282) we immediately obtain (remember that $\epsilon_{abc}R_{a\alpha}R_{b\beta}R_{c\gamma} = \epsilon_{\alpha\beta\gamma}$, therefore $\epsilon_{abc}R_{a\alpha}R_{b\beta} = \epsilon_{\alpha\beta\gamma}R_{c\gamma}$)

$$E'_a = E_b R_{ba}, \qquad B'_a = B_b R_{ba}. \tag{1.283}$$

As was expected, $\vec{E}$ and $\vec{B}$ behave like three-dimensional vectors under the spacial rotations.

To find matrix elements of $\tilde{\Lambda}$ for the case of Lorentz boost, we lower the indexes in Eq. (1.81) obtaining

$$x'_0 = \gamma\left(x_0 + \frac{V}{c}x_1\right), \; x'_1 = \gamma\left(\frac{V}{c}x_0 + x_1\right), \; x'_{2,3} = x_{2,3}, \tag{1.284}$$

where we have denoted $\gamma = \left(1 - \frac{V^2}{c^2}\right)^{-\frac{1}{2}}$. Hence the $\tilde{\Lambda}$-matrix has the following non-zero components

$$\tilde{\Lambda}^0{}_0 = \tilde{\Lambda}^1{}_1 = \gamma, \; \tilde{\Lambda}^0{}_1 = \tilde{\Lambda}^1{}_0 = \gamma\frac{V}{c}, \; \tilde{\Lambda}^a{}_b = \delta^a{}_b, \; a, b = 2, 3. \tag{1.285}$$

Substituting them into Eq. (1.282) we obtain

$$E_1' = E_1, \; E_2' = \gamma\left(E_2 - \frac{V}{c}B_3\right), \; E_3' = \gamma\left(E_3 + \frac{V}{c}B_2\right); \tag{1.286}$$

$$B_1' = B_1, \; B_2' = \gamma\left(\frac{V}{c}E_3 + B_2\right), \; B_3' = \gamma\left(-\frac{V}{c}E_2 + B_3\right). \tag{1.287}$$

Making the changes $E' \leftrightarrow E$, $B' \leftrightarrow B$, $V \to -V$, we obtain the inverse transformation

$$E_1 = E_1', \quad E_2 = \gamma\left(E_2' + \frac{V}{c}B_3'\right), \quad E_3 = \gamma\left(E_3' - \frac{V}{c}B_2'\right);$$
$$B_1 = B_1', \; B_2 = \gamma\left(-\frac{V}{c}E_3' + B_2'\right), \; B_3 = \gamma\left(\frac{V}{c}E_2' + B_3'\right). \tag{1.288}$$

According to Eqs. (1.286) and (1.287), when O registers only an electric field $\vec{E}$, the observer O' will register both electric and magnetic fields. When $\vec{B} = 0$, Eq. (1.287) can be written in the vector form

$$\vec{B}' = \frac{1}{c}[\vec{V}, \vec{E}']. \tag{1.289}$$

Hence the vectors $\vec{E}'$ and $\vec{B}'$ are mutually orthogonal in the O'-frame.

For a boost with $V \ll c$, we can disregard $\frac{V^2}{c^2}$-terms in Eqs. (1.286) and (1.287). The approximate expressions can be written in the vector form

$$\vec{E}' \approx \vec{E} - \frac{1}{c}[\vec{V}, \vec{B}], \quad \vec{B}' \approx \vec{B} + \frac{1}{c}[\vec{V}, \vec{E}]. \tag{1.290}$$

The Lorentz boost transformation can be used to find new solutions to the Maxwell equations from a known solution. We present an example of this kind below.

Example Electromagnetic field of a moving charge. Consider the charge q that moves with constant velocity V along the x^1-axis of the frame $O(x^\mu)$, passing through its origin at $t = 0$. Introduce also the frame $O(x'^\mu)$ which moves with the charge. Since the charge is at rest in O', its electromagnetic field is $\vec{E}' = \frac{q\vec{x}'}{|\vec{x}|^3}$, $\vec{H}' = 0$. Then the field in O can be obtained from Eq. (1.288)

$$E_1(x) = \frac{qx'^1}{|\vec{x}'|^3}, \quad E_2(x) = \gamma\frac{qx'^2}{|\vec{x}'|^3}, \quad E_3(x) = \gamma\frac{qx'^3}{|\vec{x}'|^3},$$
$$B_1(x) = 0, \quad B_2(x) = -\gamma\frac{V}{c}\frac{qx'^3}{|\vec{x}'|^3}, \quad B_3(x) = \gamma\frac{V}{c}\frac{qx'^2}{|\vec{x}'|^3}, \tag{1.291}$$

where we need to substitute x'^{μ} through x^{μ} according to Eq. (1.81)

$$x'^0 = \gamma(x^0 - \frac{V}{c}x^1),\ x'^1 = \gamma(x^1 - \frac{V}{c}x^0),\ x'^2 = x^2,\ x'^3 = x^3. \tag{1.292}$$

We first compute

$$\begin{aligned}|\vec{x}'|^3 &= \gamma^3 \left[(x^1 - Vt)^2 + \frac{1}{\gamma^2}[(x^2)^2 + (x^3)^2]\right]^{\frac{3}{2}} \\ &= \gamma^3 \left[(x^1 - Vt)^2 + (x^2)^2 + (x^3)^2 + \left(\frac{1}{\gamma^2} - 1\right)[(x^2)^2 + (x^3)^2]\right]^{\frac{3}{2}}.\end{aligned} \tag{1.293}$$

Introduce the vector from the position of charge q to the observation point, $\vec{R} = (x^1 - Vt,\ x^2,\ x^3)$; see Fig. 1.25 on page 75. Then $(x^2)^2 + (x^3)^2 = |\vec{R}|^2 \sin^2\theta$, where θ is the angle between $\vec{V}$ and $\vec{R}$. So $|\vec{x}'|^3 = \gamma^3 |\vec{R}|^3 \left(1 - \frac{V^2}{c^2}\sin\theta\right)^{\frac{3}{2}}$. Now the Eqs. (1.291) read

$$\vec{E} = \frac{q\vec{R}}{|\vec{R}|^3} \frac{1 - \frac{V^2}{c^2}}{\left(1 - \frac{V^2}{c^2}\sin^2\theta\right)^{\frac{3}{2}}}, \qquad \vec{B} = \frac{q}{c}[\vec{V}, \vec{E}]. \tag{1.294}$$

These expressions give the electromagnetic field of a moving charge. The electric field is directed to the charge, with a magnitude that increases with θ. The electric field in the direction of motion, ($\theta = 0$), has the minimal magnitude $|\vec{E}| = \frac{q}{|\vec{R}|^2}\left(1 - \frac{V^2}{c^2}\right)$, and in the orthogonal direction, $\theta = \frac{\pi}{2}$, it acquires the maximum value $|\vec{E}| = \frac{q}{|\vec{R}|^2}\frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}$. The electric and the magnetic field fields of the moving charge are shown in Figs. 1.26b on page 75 and 1.27 on page 75.

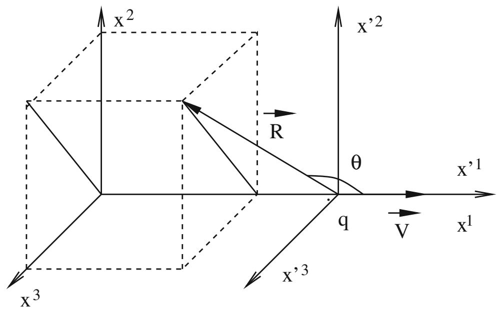

Fig. 1.25 The electric field of a moving charge depends on the angle θ

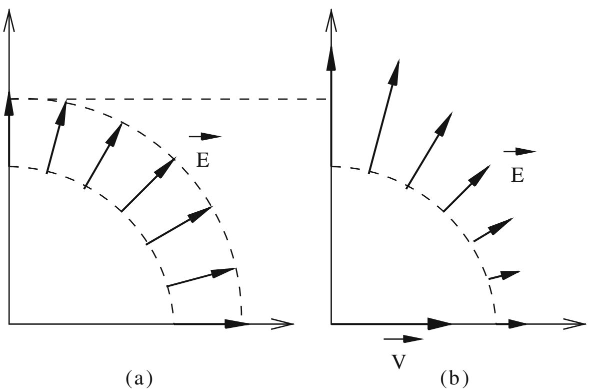

Fig. 1.26 **(a)** Electric field of a charge at rest. **(b)** Field of a moving charge

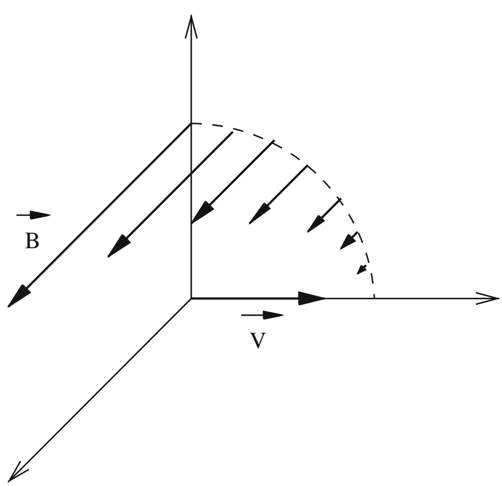

Fig. 1.27 Magnetic field of a moving charge

Chapter 2
Hamiltonian Formalism

2.1 Derivation of Hamiltonian Equations

As we have discussed, Lagrangian formulation of classical mechanics is based on Euler–Lagrange (Newton) equations of motion, which represent a system of second-order differential equations, written for a set of variables that describe the *position* of a physical system of interest. Hamiltonian formulation suggests an equivalent description in terms of first-order equations written for independent variables describing the *position and velocity* of the system. The aim of this section is to establish an equivalence of the two descriptions.

2.1.1 Preliminaries

Hamiltonian equations can be obtained from Lagrangian ones by successive application of two well-known procedures in a theory of differential equations: reduction of order and change of variables. Both procedures are intended to obtain an equivalent system of equations from a given system. So, we recall here some elementary facts from the theory of ordinary differential equations, which will be used below.

Reduction of the order of a system. A second-order system of n equations for n independent variables $q^a(\tau)$,

$$F^a(q^a, \dot{q}^b, \ddot{q}^c) = 0, \tag{2.1}$$

is equivalent to the first-order system of $2n$ equations for $2n$ independent variables $q^a(\tau), v^b(\tau)$

$$\dot{q}^a = v^a, \qquad F^a(q^a, v^b, \dot{v}^c) = 0, \tag{2.2}$$

in the following sense:

(a) If $q^a(\tau)$ obeys Eq. (2.1), then the functions $q^a(\tau), v^a(\tau) \equiv \dot{q}^a(\tau)$ obey Eq. (2.2);
(b) If the functions $q^a(\tau), v^a(\tau)$ obey Eq. (2.2), then $q^a(\tau)$ obeys Eq. (2.1).

A. Deriglazov, *Classical Mechanics*, DOI 10.1007/978-3-642-14037-2_2,

In other words, there is a one-to-one correspondence among solutions to the systems. The system (2.2) is referred to as the *first-order form* of the system (2.1).

Normal form of a system. We restrict ourselves to the first-order system

$$G^i(z^j, \dot{z}^k) = 0. \tag{2.3}$$

It is said to be presented in the *normal form* if all the equations are solved algebraically with respect to higher derivatives

$$\dot{z}^i = g^i(z^j). \tag{2.4}$$

Any system with $\det \frac{\partial G^i}{\partial \dot{z}^j} \neq 0$ can (locally) be rewritten in the normal form. According to the theory of differential equations, a normal system has well established properties. In particular, under known restrictions to functions g^i, the theorem for the existence and uniqueness of a solution holds: let z_0^i be given numbers, then locally there exists a unique solution $z^i(\tau)$ of the system (2.4) that obeys the initial conditions $z^i(0) = z_0^i$. Physically it means the causal dynamics and, in turn, a possibility of interpretation of the system (2.3) as the equations of motion for some physical system of classical mechanics.

Change of variables. Let $z'^i(z^j)$ be given functions, with the property

$$\det \frac{\partial z'^i}{\partial z^j} \neq 0. \tag{2.5}$$

Starting from original parametrization z^i of the configuration space for the system (2.3), functions $z'^i(z^j)$ can be used to define another parametrization z'^i, namely

$$z'^i = z'^i(z^j). \tag{2.6}$$

We use the same letter z'^i to denote the function and the new coordinate, as long as this does not lead to any misunderstanding. According to the condition (2.5), *change of variables* $z^i \longrightarrow z'^i$ is invertible: the expressions (2.6) can be resolved with relation to z^i, with the result being

$$z^i = z^i(z'^j). \tag{2.7}$$

Once the functions $z'^i(z^j)$ have been chosen, we can use the new coordinates to analyze the system (2.3). Namely, the system

$$G^i(z^j(z'^k), \dot{z}^j(z'^k)) = 0, \tag{2.8}$$

where $\dot{z}^j(z'^k) = \frac{\partial z^i}{\partial z'^k}\dot{z}'^k$, is equivalent to the initial system (2.3): if $z^i(\tau)$ obeys the system (2.3), then $z'^i(\tau) \equiv z'^i(z^j(\tau))$ obeys (2.8), and vice versa.

Below we prefer to use the notation

$$G^i(z^j, \dot{z}^j)\Big|_{z=z(z')} = 0, \tag{2.9}$$

instead of (2.8), since sometimes caution is needed in making use of the substitution, see, for example, Eqs. (2.27) and (2.28) below.

Hamiltonian system. Let q^a, p_a, $a = 1, 2, \ldots, n$ be independent variables. The normal system

$$\dot{q}^a = Q^a(q, p, \tau), \qquad \dot{p}_a = P_a(q, p, \tau), \tag{2.10}$$

with the given functions Q, P is called the *Hamiltonian system*, if there is a function $H(q, p, \tau)$, such that

$$Q^a = \frac{\partial H}{\partial p_a}, \qquad P_a = -\frac{\partial H}{\partial q^a}. \tag{2.11}$$

In accordance with this, the Hamiltonian system can be written in the form

$$\dot{q}^a = \frac{\partial H}{\partial p_a}, \qquad \dot{p}_a = -\frac{\partial H}{\partial q^a}. \tag{2.12}$$

Equation (2.11) implies the necessary conditions for the system to be a Hamiltonian one

$$\frac{\partial Q^a}{\partial q^b} = -\frac{\partial P_b}{\partial p_a}, \qquad \frac{\partial Q^a}{\partial p_b} = \frac{\partial Q^b}{\partial p_a}, \qquad \frac{\partial P_a}{\partial q^b} = \frac{\partial P_b}{\partial q^a}. \tag{2.13}$$

2.1.2 *From Lagrangian to Hamiltonian Equations*

Let q^a, $a = 1, 2, \ldots, n$ represent generalized coordinates of the configuration space for a mechanical system with the Lagrangian being $L(q^a, \dot{q}^a)$. Then the dynamics is governed by the second-order Euler-Lagrange equation

$$\frac{d}{d\tau}\left(\frac{\partial L(q, \dot{q})}{\partial \dot{q}^a}\right) - \frac{\partial L(q, \dot{q})}{\partial q^a} = 0. \tag{2.14}$$

For any Lagrangian system there is an equivalent Hamiltonian system. We demonstrate this mathematically notable fact for the particular case of a *nonsingular Lagrangian*

$$\det \frac{\partial^2 L(q, \dot{q})}{\partial \dot{q}^a \partial \dot{q}^b} \neq 0. \tag{2.15}$$

In this case, the system (2.14) can be rewritten in the first-order normal form. Then in specially chosen coordinates it acquires the Hamiltonian form. It basically gives the Hamiltonian formulation of mechanics.

Our presentation below is somewhat more detailed as compared with other textbooks. This has been done for two reasons. First, equivalence, which represents one of the basic facts of classical mechanics, will be manifest in our discussion. Second, our treatment of the subject turns out to be useful for singular Lagrangians, revealing an algebraic structure of the Hamiltonian formulation for these systems [24].

Computing the derivative with respect to τ in Eq. (2.14), the latter can be written as

$$M_{ab}\ddot{q}^b = K_a, \tag{2.16}$$

where it was denoted

$$M_{ab}(q,\dot{q}) \equiv \frac{\partial^2 L(q,\dot{q})}{\partial \dot{q}^a \partial \dot{q}^b}, \quad K_a(q,\dot{q}) \equiv \frac{\partial L}{\partial q^a} - \frac{\partial^2 L}{\partial \dot{q}^a \partial q^b}\dot{q}^b. \tag{2.17}$$

Let us start with construction of the first-order form for the system (2.16). We find it instructive to present here a less formal reasoning, as compared to that of Sect. 2.1.1. We introduce $2n$-dimensional *configuration-velocity space* parameterized by *independent* coordinates q^a, v^b (sometimes the coordinates v^b are called *generalized velocities*). Let us define evolution in this space according to the equations

$$M_{ab}\ddot{q}^b = K_a, \qquad v^a = \dot{q}^a, \tag{2.18}$$

with $M(q,\dot{q})$, $K(q,\dot{q})$ given by Eq. (2.17). As before, time dependence of the coordinates $q^a(\tau)$ is determined by Lagrangian equations (2.16), while $v^a(\tau)$ accompanies $\dot{q}^a(\tau)$: $v^a(\tau)$ is determined from the known $q^a(\tau)$, taking its derivative. Evidently, systems (2.16) and (2.18) are equivalent. Further, we can use one of the equations of the system in other equations, obtaining an equivalent system. Substitution of the second equation from (2.18) into the first one gives the desired first order system

$$\dot{q}^a = v^a, \qquad \bar{M}_{ab}\dot{v}^b = \bar{K}_a, \tag{2.19}$$

where $\bar{M}$, $\bar{K}$ are obtained from (2.17) by the replacement $\dot{q} \to v$, for example

$$\bar{M}_{ab} \equiv M_{ab}(q,\dot{q})|_{\dot{q}^a \to v^a} = \frac{\partial L(q,v)}{\partial v^a \partial v^b}. \tag{2.20}$$

According to Eq. (2.15), the matrix $\bar{M}$ is invertible. Applying the inverse matrix $\tilde{\bar{M}}$, the Eqs. (2.19) can be presented in the normal form $\dot{q} = v$, $\dot{v} = \tilde{\bar{M}}\bar{K}$. The right-hand sides of these equations do not obey Eq. (2.13). So in terms of the variables q, v the system is not a Hamiltonian one.

Making the variable change $q \to q(q', v')$, $v \to v(q', v')$ in Eq. (2.19), we could look for the new variables that imply the Hamiltonian form of the system. The point here is that there is a wide class of so-called *canonical transformations* that preserve the Hamiltonian form of an arbitrary Hamiltonian system (see Sect. 2.7 below). Hence the variables under discussion are not unique.[1] The remarkable observation made by W. R. Hamilton was that it is sufficient to make the change of variables of the form (with v' conventionally denoted as p)

$$\begin{pmatrix} q^a \\ v^b \end{pmatrix} \leftrightarrow \begin{pmatrix} q'^a \\ p_b \end{pmatrix}, \quad \text{where} \quad q'^a \equiv q^a, \quad p_b = \frac{\partial L(q, v)}{\partial v^b}, \tag{2.21}$$

to transform the system (2.19) into the Hamiltonian one. Due to Eq. (2.15) we have $\det \frac{\partial p_b(q,v)}{\partial v^c} \neq 0$. The latter condition guarantees invertibility of the transformation (2.21). Let us denote the inverse transformation as

$$v^a = v^a(q, p). \tag{2.22}$$

This implies the identities

$$\left.\frac{\partial L}{\partial v^a}\right|_{v(q,p)} \equiv p_a, \qquad \frac{\partial p_a}{\partial v^b} = \frac{\partial^2 L}{\partial v^a \partial v^b} = M_{ab}(q, v). \tag{2.23}$$

Let us confirm that in terms of the variables q, p the system (2.19) acquires the Hamiltonian form.

According to Sect. 2.1.1, the dynamics for the new variables is obtained from (2.19) by substitution $v \to v(q, p)$. We have

$$\dot{q}^a = v^a(q, p), \tag{2.24}$$

$$\begin{aligned} \bar{M}_{ab}\big|_{v(q,p)} \frac{\partial v^b}{\partial p_c} \dot{p}_c &= \bar{K}_a\big|_{v(q,p)} - \bar{M}_{ab}\big|_{v(q,p)} \frac{\partial v^b}{\partial q^c} v^c(q, p) \\ = \left.\frac{\partial L(q, v)}{\partial q^a}\right|_{v(q,p)} &- \left(\left.\frac{\partial^2 L(q, v)}{\partial q^c \partial v^a}\right| - \left.\frac{\partial^2 L(q, v)}{\partial v^a \partial v^b}\right| \frac{\partial v^b}{\partial q^c} \right) v^c(q, p). \end{aligned} \tag{2.25}$$

In the last equation, the left hand side is just p_a, as is implied[2] by Eq. (2.23), while the expression inside the brackets vanishes since it is $\frac{\partial}{\partial q^c} \left(\left.\frac{\partial L}{\partial v^a}\right|_{v(q,p)} \right) = \frac{\partial p_a}{\partial q^c} = 0$. Then the Eqs. (2.24) and (2.25) acquire the form

---

[1] If the change $q(q', v')$, $v(q', v')$ transforms the system (2.19) into the Hamiltonian one, and $q'(q'', v'')$, $v'(q'', v'')$ is the canonical transformation, then the change $q(q'(q'', v''), v'(q'', v''))$, $v(q'(q'', v''), v'(q'', v''))$ transforms (2.19) into the Hamiltonian system as well.

[2] Recall that the Jacobi matrices of direct and inverse transformations are opposites: from the identity $z^i(z'^j(z^k)) = z^i$ we have $\left.\frac{\partial z^i}{\partial z'^j}\right|_{z'(z)} \frac{\partial z'^j}{\partial z^k} = \delta^i{}_k$. See also Exercise 2.1.2 on page 83.

$$\dot{q}^a = v^a(q,p), \qquad \dot{p}_a = \left.\frac{\partial \bar{L}(q,v)}{\partial q^a}\right|_{v(q,p)}. \tag{2.26}$$

To substitute $v(q,p)$ in the last equation, let us compute

$$\begin{aligned}\frac{\partial}{\partial q^a}L(q,v(q,p)) &= \left.\frac{\partial L(q,v)}{\partial q^a}\right|_{v(q,p)} + \left.\frac{\partial L(q,v)}{\partial v^b}\right|_{v(q,p)}\frac{\partial v^b}{\partial q^a} \\ &= \left.\frac{\partial L(q,v)}{\partial q^a}\right|_{v(q,p)} + p_b\frac{\partial v^b}{\partial q^a},\end{aligned} \tag{2.27}$$

which implies

$$\left.\frac{\partial L(q,v)}{\partial q^a}\right|_{v(q,p)} = -\frac{\partial}{\partial q^a}\Big(p_b v^b(q,p) - L(q,v(q,p))\Big). \tag{2.28}$$

Let us denote

$$H(q,p) = p_b v^b(q,p) - L(q,v(q,p)), \tag{2.29}$$

where $v(q,p)$ is given in implicit form by Eq. (2.21). Then the expression (2.28) reads

$$\left.\frac{\partial \bar{L}(q,v)}{\partial q^a}\right|_{v(q,p)} = -\frac{\partial H(q,p)}{\partial q^a}. \tag{2.30}$$

The function $H(q,p)$ is called the *Hamiltonian* of the physical system. To complete the derivation of the Hamiltonian equations, note the following property of the Hamiltonian:

$$\frac{\partial H}{\partial p_a} = v^a(q,p) + p_b\frac{\partial v^b}{\partial p_a} - \left.\frac{\partial L(q,v)}{\partial v^b}\right|_{v(q,p)}\frac{\partial v^b}{\partial p_a} = v^a(q,p). \tag{2.31}$$

Using these results, the equations of motion (2.26) acquire the Hamiltonian form

$$\dot{q}^a = \frac{\partial H}{\partial p_a}, \qquad \dot{p}_a = -\frac{\partial H}{\partial q^a}, \tag{2.32}$$

and are known as *Hamiltonian equations of motion*. Note that the first equation is the Eq. (2.24) written in another notation.

The coordinates p_a defined by Eq. (2.21) are called *conjugated momenta* for q^a. The configuration-velocity space parameterized by the coordinates q^a, p_b is referred to as the *phase space* of the system.

The passage (2.29) from $L(q,v)$ to $H(q,p)$ is known as *Legendre transformation*. Its basic properties are presented by Eqs. (2.31), (2.30). Note its meaning: if the variable change $v^a \to p_b$ (the variables q^a are considered as parameters) is "generated" by the function $L(v)$ according to Eq. (2.21), $p_a = \frac{\partial L}{\partial v^a}$, then the Legendre transformation gives the generating function H of the inverse transformation (2.22), $v^a = \frac{\partial H}{\partial p_a}$. See also Exercise 5 below.

To sum up, in this section we have demonstrated that for the case of a nonsingular system, the Lagrangian equations of motion (2.14) for the configuration space variables q^a are equivalent to the Hamiltonian equations (2.32) for independent phase-space variables q^a, p_b. According to our procedure, the Hamiltonian formulation of mechanics is the first order form of the Lagrangian formulation, further rewritten using the special coordinates q^a, p_b of the configuration-velocity space. Schematically we write

$$q^a \to (q^a, v^b) \leftrightarrow (q^a, p_b). \tag{2.33}$$

Exercises

1. Check that the function $v^a(q, p)$ defined by (2.21) obeys the equation $\frac{\partial v^b(q,p)}{\partial p_c} = \tilde{\bar{M}}^{bc}(q, v)\Big|_{v(q,p)}$, where $\tilde{\bar{M}}$ is the inverse matrix for $\bar{M}$.
2. Derive the identity $\frac{\partial v^a}{\partial q^c} = -\tilde{M}^{ab} \frac{\partial^2 \bar{L}}{\partial v^b \partial q^c}\Big|_{v(q,p)}$.
3. Work out the Lagrangian equations (2.18) from the Hamiltonian ones (2.32) and (2.29).
4. Confirm that all the results of this section remain true for the time-dependent Lagrangian $L(q, \dot{q}, \tau)$.
5. Legendre transformation.

   (a) Let the functions $f_i(x^j)$ be generated by the function $F(x^j)$, that is $f_i = \frac{\partial F}{\partial x^i}$, and let g_i be the inverse function of f_i. Verify that $g_i(x^j)$ is generated by $x^i g_i - F(g)$, $g_i = \frac{\partial}{\partial x^i}(x^i g_i - F(g))$.
   (b) Observe that for a one-dimensional case the Legendre transformation gives a simple formula for the indefinite integral of the inverse function.
   (c) If F depends on the parameters y^i, and $f_i(x, y) = \frac{\partial F(x,y)}{\partial x^i}$, then derivatives of the generating functions with respect to y are the same, $\frac{\partial F}{\partial y^i}\Big|_{x \to g(x,y)} = \frac{\partial}{\partial y^i}(x^j g_j - F(g, y))$.

2.1.3 Short Prescription for Hamiltonization Procedure, Physical Interpretation of Hamiltonian

The passage from a Lagrangian to a Hamiltonian description of a system is referred to as the *Hamiltonization procedure*. Note that the resulting Hamiltonian equation

(2.32) do not contain the velocities v^a. Then we expect the existence of a formal recipe for the Hamiltonization procedure that, in particular, does not mention the velocities. Let

$$S = \int d\tau L(q^a, \dot{q}^a), \tag{2.34}$$

be the Lagrangian action of some nonsingular system. Inspection of the previous section allows us to formulate the recipe as follows.

(1) Write the conjugated momenta for the variables q^a according to the equations (see Eq. (2.21))

$$p_a = \frac{\partial L(q, \dot{q})}{\partial \dot{q}^a}. \tag{2.35}$$

(2) Resolve the equations algebraically in relation to $\dot{q}^a$: $\dot{q}^a = v^a(q, p)$, and find the Hamiltonian (see Eq. (2.29))

$$H(q, p) = \left(p_b \dot{q}^b - L(q, \dot{q})\right)\Big|_{\dot{q}=v(q,p)}. \tag{2.36}$$

(3) Write the Hamiltonian equations (2.32).

According to the previous section, the resulting equations are *equivalent* to the Lagrangian equations of motion for the action (2.34).

Note that the function $H(q, p)$ turns out to be a basic object of Hamiltonian formalism. To reveal the physical interpretation of the Hamiltonian, let us consider a particle in the presence of a potential $U(x)$. The corresponding action is

$$S = \int d\tau \left[\frac{1}{2} m (\dot{x}^a)^2 - U(x^a)\right]. \tag{2.37}$$

To construct the Hamiltonian formulation, we have the momenta $p_a = m\dot{x}^a$. This implies $\dot{x}^a = \frac{1}{m} p_a$, and leads to the Hamiltonian $H(x, p) = \frac{1}{2m}(p^a)^2 + U(x)$. Making the inverse change, we obtain the position-velocity function: $E(x, \dot{x}) \equiv H(x, p)|_{p=m\dot{x}} = \frac{1}{2} m(\dot{x}^a)^2 + U(x^a)$ which represents the total energy of the particle. The reasoning works equally for a system of particles. Thus the Hamiltonian of nonsingular Lagrangian theory in Cartesian coordinates represents the total energy of a system written in terms of the phase space variables.[3]

Exercise
Bearing in mind the ambiguity presented in the Hamiltonization procedure (see the discussion just before Eq. (2.21)), let us define momenta for the model

[3] The case of generalized coordinates will be discussed below; see Exercise 3 on page 140.

(2.37) according to the rule $\dot{x}^a = \frac{1}{m}p_a + A_a(x)$, where $A_a(x)$ is a given function. Write the Hamiltonian equations and work out conditions for A_a which imply their canonical form (that is the form (2.32) with a function $\tilde{H}$). Write the corresponding Hamiltonian $\tilde{H}$. Does it have an interpretation as the energy of the particle? Derive the Lagrangian equations from the Hamiltonian ones.

2.1.4 Inverse Problem: From Hamiltonian to Lagrangian Formulation

Let $H(q,p)$ be the Hamiltonian of some non-singular Lagrangian system. The problem is to restore the corresponding Lagrangian, that is, to construct a function $L(q,\dot{q})$ which would lead to the given $H(q,p)$ after the Hamiltonization procedure. For this purpose we have the phase-space expression (2.29), which determines the desired L as a function of q, p: $L(q, v(q,p)) = p_a v^a - H(q,p)$. According to Sect. 2.1.2, phase space and configuration-velocity space quantities are related by the change of variables (2.21) and (2.22). Then L, as a function of q, v, is obtained by making this change in the previous expression

$$L(q,v) = \left(p_a v^a - H(q,p)\right)\Big|_{p(q,v)}. \tag{2.38}$$

To find the transition functions $p(q,v)$, it is sufficient to recall Eq. (2.31), which determines the inverse functions: $v^a(q,p) = \frac{\partial H(q,p)}{\partial p_a}$. Thus we resolve the equations $v^a = \frac{\partial H(q,p)}{\partial p_a}$ in relation of p: $p_a = p_a(q,v)$, which gives the desired transition functions.

The resulting formal prescription can be formulated without mentioning the velocities: starting from a given $H(q,p)$, solve the part of Hamiltonian equations $\dot{q}^a - \frac{\partial H(q,p)}{\partial p_a} = 0$ with respect to p: $p = p(q,\dot{q})$. Then $L(q,\dot{q}) = [p_a\dot{q}^a - H(q,p)]|_{p(q,\dot{q})}$.

2.2 Poisson Bracket and Symplectic Matrix

Here we introduce standard notation and conventions used to deal with Hamiltonian equations. Let $\{A(q,p), B(q,p), \ldots\}$ be a set of phase-space functions.

Definition 1 *The Poisson bracket* is an application that with any two phase-space functions A, B associates a third function denoted $\{A, B\}$, according to the rule

$$\{A, B\} = \frac{\partial A}{\partial q^a}\frac{\partial B}{\partial p_a} - \frac{\partial B}{\partial q^a}\frac{\partial A}{\partial p_a}. \tag{2.39}$$

The definition implies the following properties of the Poisson bracket:

(a) antisymmetry

$$\{A, B\} = -\{B, A\}; \tag{2.40}$$

(b) linearity with respect to both arguments, as a consequence of (2.40). linearity with respect to second argument is

$$\{A, \lambda B + \eta C\} = \lambda\{A, B\} + \eta\{A, C\}, \qquad \lambda, \eta = \text{const}; \tag{2.41}$$

(c) Leibnitz rule

$$\{A, BC\} = \{A, B\}C + B\{A, C\}; \tag{2.42}$$

(d) Jacobi identity

$$\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0. \tag{2.43}$$

Exercise
Verify (2.43) by direct computations. Hint: consider separately all the terms involving, for example, two derivatives of B.

Poisson brackets among phase-space variables are called *fundamental brackets*. They are:

$$\{q^a, p_b\} = \delta^a{}_b, \qquad \{q^a, q^b\} = 0, \qquad \{p_a, p_b\} = 0. \tag{2.44}$$

Poisson brackets can be used to rewrite Hamiltonian equations in the form:

$$\dot{q}^a = \{q^a, H\}, \qquad \dot{p}_a = \{p_a, H\}. \tag{2.45}$$

Hence the Poisson bracket of $q,\ p$ with the Hamiltonian determines their rate of variation with time. Moreover, the same is true for any phase-space function: if $q^a(\tau), p_b(\tau)$ is a solution to the Hamiltonian equations, the rate of variation of the function $A(q(\tau), p(\tau))$ can be computed as:

$$\begin{aligned}\dot{A}(q, p) &= \frac{\partial A}{\partial q^a}\dot{q}^a + \frac{\partial A}{\partial p_a}\dot{p}_a = \frac{\partial A}{\partial q^a}\{q^a, H\} + \frac{\partial A}{\partial p_a}\{p_a, H\} \\ &= \{A, H\}.\end{aligned} \tag{2.46}$$

Thus $\{A(q, p), H\} = 0$ implies that A is a *conserved quantity*, that is, it has a fixed value throughout any given solution. As an example, let us apply this result to compute the rate of variation of a Hamiltonian. We have $\dot{H} = \{H, H\} = 0$, due to the antisymmetry of the Poisson bracket. Hence the Hamiltonian is the conserved quantity, which gives a further argument in support of its interpretation as the total energy.

Below it will be convenient to work with phase-space quantities by using the following notation. For the phase-space coordinates we use the unique symbol: $(q^a, p_b) \equiv z^i$, $i = 1, 2, \ldots, 2n$, or, equivalently, for $a, b = 1, 2, \ldots, n$ we have $z^a = q^a$ and $z^{n+b} = p_b$. Thus Latin indices from the middle of the alphabet run from 1 to $2n$. Let us also introduce the $2n \times 2n$-dimensional *symplectic matrix* composed of four $n \times n$ blocks

$$\omega^{ij} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix}. \tag{2.47}$$

In more detail, for $a, b = 1, 2, \ldots, n$ one writes $\omega^{ab} = 0$, $\omega^{a,n+b} = \delta^{ab}$, $\omega^{n+a,b} = -\delta^{ab}$, $\omega^{n+a,n+b} = 0$. The symplectic matrix is antisymmetric: $\omega^{ij} = -\omega^{ji}$ and invertible, with the inverse matrix being

$$\omega_{ij} = \begin{pmatrix} \mathbf{0} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}. \tag{2.48}$$

In this notation the Poisson brackets (2.39) and (2.44) acquire a more compact form

$$\{A, B\} = \frac{\partial A}{\partial z^i}\omega^{ij}\frac{\partial B}{\partial z^j}, \qquad \{z^i, z^j\} = \omega^{ij}, \tag{2.49}$$

while the Hamiltonian equations can be written as

$$\dot{z}^i = \omega^{ij}\frac{\partial H}{\partial z^j}, \quad or \quad \dot{z}^i = \{z^i, H\}. \tag{2.50}$$

Exercise
Verify the Jacobi identity with use of the representation (2.49).

2.3 General Solution to Hamiltonian Equations

As a first application of Hamiltonian formalism, we find here a general solution to Hamiltonian equations with an arbitrary time-independent Hamiltonian in terms of power series with respect to τ.

As a preliminary step, consider the differential operator defined by the formal series

$$e^{a\partial_x} = 1 + a\partial_x + \frac{1}{2}a\partial_x(a\partial_x) + \ldots = \sum_{n=0}^{\infty}\frac{1}{n!}(a\partial_x)^n, \quad a = \text{const.} \tag{2.51}$$

This obeys the properties $e^{a\partial_x}x = x+a$, $e^{a\partial_x}G(x) = G(e^{a\partial_x}x)$, as can be verified by expansion in power series of both sides of these equalities. There is a generalization of the last equality for the case of a function $a(x)$

$$e^{a(x)\partial_x}G(x) = G\left(e^{a(x)\partial_x}x\right). \tag{2.52}$$

Exercise
Verify the validity of Eq. (2.52) up to the third order of power expansion.

Due to the identity (2.52), the function $f(\tau, x) = e^{\tau a(x)\partial_x}x$ turns out to be a formal solution to the equation

$$\frac{\partial f}{\partial \tau} = a(f). \tag{2.53}$$

Besides, this obeys the initial condition $f(0, x) = x$. This observation can be further generalized for the case of several variables, the functions $f^i(\tau, x^j) = e^{\tau a^k(x^j)\partial_k}x^i$ obey the system

$$\frac{\partial f^i}{\partial \tau} = a^i(f^j). \tag{2.54}$$

Note that the Hamilton equations $\dot{z}^i = \{z^i, H\}$ represent a system of this type. So its solution is

$$z^i(\tau) = e^{\tau\{z_0^k, H(z_0)\}\frac{\partial}{\partial z_0^k}}z_0^i. \tag{2.55}$$

In particular, the position of a system as a function of $2n$ constants is given by

$$q^a(\tau, q_0, p_0) = e^{\tau\{z_0^k, H(z_0)\}\frac{\partial}{\partial z_0^k}}q_0^a. \tag{2.56}$$

For the Hamiltonian $H = \frac{\vec{p}^2}{2m} + U(q)$ it implies

$$q^a(\tau, q_0, p_0) = e^{\tau(\frac{1}{m}\vec{p}_0\cdot\vec{\nabla} - \vec{\nabla}U\cdot\vec{\nabla}_{p0})}q_0^a. \tag{2.57}$$

We illustrate the formula (2.55) with several examples.

1. For a *free particle* with the Hamiltonian $H = \frac{\vec{p}^2}{2m}$ we have

$$\begin{aligned}
x^a(t) &= e^{t\frac{1}{m}\vec{p}_0\cdot\vec{\nabla}} x_0^a \\
&= \left(1 + t\frac{1}{m}\vec{p}_0\cdot\vec{\nabla} + \frac{t^2}{2!m^2}(\vec{p}_0\cdot\vec{\nabla})(\vec{p}_0\cdot\vec{\nabla}) + \ldots\right) x_0^a \\
&= x_0^a + \frac{1}{m}p_{0a}t, \\
p_a(t) &= e^{t\frac{1}{m}\vec{p}_0\cdot\vec{\nabla}} p_{0a} \\
&= \left(1 + t\frac{1}{m}\vec{p}_0\cdot\vec{\nabla} + \frac{t^2}{2!m^2}(\vec{p}_0\cdot\vec{\nabla})(\vec{p}_0\cdot\vec{\nabla}) + \ldots\right) p_{0a} = p_{0a}.
\end{aligned} \tag{2.58}$$

2. For a *one-dimensional harmonic oscillator* with the Hamiltonian $H = \frac{p^2}{2m} + \frac{1}{2}kx^2$ we obtain

$$\begin{aligned}
x(t) &= e^{t(\frac{1}{m}p_0\partial_x - kx\partial_p)} x_0 \\
&= x_0 + p_0\frac{1}{m}t - x_0\frac{k}{m}\frac{t^2}{2!} - p_0\frac{k}{m^2}\frac{t^3}{3!} + x_0\frac{k^2}{m^2}\frac{t^4}{4!} + p_0\frac{k^2}{m^3}\frac{t^5}{5!} - \ldots .
\end{aligned} \tag{2.59}$$

Bringing together even and odd degrees of t, it gives the expected result

$$\begin{aligned}
x(t) &= x_0\left(1 - \frac{k}{m}\frac{t^2}{2!} + \frac{k^2}{m^2}\frac{t^4}{4!} - \ldots\right) \\
&\quad + p_0\left(\frac{1}{m}t - \frac{k}{m^2}\frac{t^3}{3!} + \frac{k^2}{m^3}\frac{t^5}{5!} - \ldots\right) \\
&= x_0\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}\left(\sqrt{\frac{k}{m}}t\right)^{2n} + \frac{p_0}{\sqrt{km}}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}\left(\sqrt{\frac{k}{m}}t\right)^{2n+1} \\
&= x_0\cos\sqrt{\frac{k}{m}}t + \frac{p_0}{\sqrt{km}}\sin\sqrt{\frac{k}{m}}t.
\end{aligned} \tag{2.60}$$

3. *Kepler's problem*. Consider a particle with the initial position and velocity being $x^a(0) = x_0^a, \dot{x}^a(0) = v_0^a$, subject to a central field potential with origin at the center of a coordinate system. As we have seen in Sect. 1.6, trajectory of motion is a second-order plane curve (ellipse, hyperbola or parabola) with the polar equation being (see Eq. (1.146))

$$r = \frac{p}{1 + e\cos(\theta + \theta_0)}. \tag{2.61}$$

Here the constat p, the eccentricity e, and the inclination angle θ_0 of the semi-axis are determined through initial values of the problem. We reproduce this result using the formula (2.55).

The system is described by the action

$$S = \int dt \left(\frac{m}{2} \dot{\vec{x}}^2 + \frac{\alpha}{|\vec{x}|} \right). \tag{2.62}$$

Remember that due to the conservation of angular momentum, the particle trajectory lies on the plane of vectors $\vec{x}_0$, $\vec{v}_0$. Choosing the cartesian axis x^1, x^2 on the plane, we discard the third variable in the action, $x^3(t) = 0$. To proceed further, let us rewrite the action in polar coordinates, $x^1 = r\cos\theta$, $x^2 = r\sin\theta$

$$S = \int dt \left(\frac{m}{2} \dot{r}^2 + \frac{m}{2} r^2 \dot{\theta}^2 + \frac{\alpha}{r} \right). \tag{2.63}$$

Denoting conjugated momenta for r, θ as p, p_θ, the Hamiltonian reads

$$H = \frac{1}{2m} p^2 + \frac{1}{2mr^2} p_\theta^2 - \frac{\alpha}{r}. \tag{2.64}$$

It leads to the equations

$$\begin{aligned} \dot{r} &= \frac{1}{m} p, \qquad & \dot{p} &= \frac{1}{mr^3} p_\theta^2 - \frac{\alpha}{r^2}, \\ \dot{\theta} &= \frac{1}{mr^2} p_\theta, \qquad & \dot{p}_\theta &= 0. \end{aligned} \tag{2.65}$$

The last equation implies $p_\theta = l = \text{const}$. We are interested in finding a form of trajectory, $r(\theta)$, $p(\theta)$. The relevant equations can be obtained from (2.65). Considering $r = r(t)$, $p = p(t)$, $\theta = \theta(t)$ as parametric equations of the trajectory, we write $r' \equiv \frac{dr}{d\theta} = \frac{\dot{r}}{\dot{\theta}}$, $p' \equiv \frac{dp}{d\theta} = \frac{\dot{p}}{\dot{\theta}}$. Using (2.65) in these expressions, we obtain the equations

$$r' = \frac{r^2}{l} p, \qquad p' = \frac{l}{r} - \frac{\alpha m}{l}. \tag{2.66}$$

They acquire a more simple form if we introduce the variable $q = \frac{1}{r}$. Then

$$q' = -\frac{1}{l} p, \qquad p' = lq - \frac{\alpha m}{l}. \tag{2.67}$$

They form a Hamiltonian system,[4] with the Hamiltonian being

[4] Notice that (2.66) is not a Hamiltonian system. The regular way to construct Hamiltonian equations for a trajectory will be discussed in Sect. 6.1.2 below.

$$H(r,p) = -\frac{1}{2l}p^2 - \frac{l}{2}q^2 + \frac{\alpha m}{l}q. \tag{2.68}$$

Using Eq. (2.55) with this Hamiltonian, we obtain the solution

$$\begin{aligned}
q(\theta) &= e^{\theta(-\frac{p_0}{l}\partial_{q0}+(lq_0-\frac{\alpha m}{l})\partial_{p0}))}q_0 \\
&= q_0 - \frac{1}{l}p_0\theta - \left(q_0 - \frac{\alpha m}{l^2}\right)\frac{\theta^2}{2!} + \frac{1}{l}p_0\frac{\theta^3}{3!} - \left(q_0 - \frac{\alpha m}{l^2}\right)\frac{\theta^4}{4!} \\
&= \frac{\alpha m}{l^2} + \left(q_0 - \frac{\alpha m}{l^2}\right)\sum_{n=0}^{\infty}(-1)^n\frac{\theta^{2n}}{(2n)!} - \frac{p_0}{l}\sum_{n=0}^{\infty}(-1)^n\frac{\theta^{2n+1}}{(2n+1)!} \\
&= \frac{\alpha m}{l^2} + \left(q_0 - \frac{\alpha m}{l^2}\right)\cos\theta - \frac{p_0}{l}\sin\theta.
\end{aligned} \tag{2.69}$$

Returning to the variable $r = \frac{1}{q}$, we have

$$\begin{aligned}
\frac{l^2}{\alpha m r(\theta)} &= 1 + \left(\frac{l^2}{\alpha m r_0} - 1\right)\cos\theta - \frac{l p_0}{\alpha m}\sin\theta \\
&\equiv 1 + A\cos\theta - B\sin\theta \\
&= 1 + \sqrt{A^2+B^2}\left(\frac{A}{\sqrt{A^2+B^2}}\cos\theta - \frac{B}{\sqrt{A^2+B^2}}\sin\theta\right).
\end{aligned} \tag{2.70}$$

Comparing $A^2 + B^2$ with the Hamiltonian (2.64), we obtain $A^2 + B^2 = 1 + \frac{2l^2E}{\alpha^2 m}$, where $E = \frac{p_0^2}{2m} + \frac{l^2}{2mr_0^2} - \frac{\alpha}{r_0}$ represents the total energy. Besides, since $\left(\frac{A}{\sqrt{A^2+B^2}}\right)^2 + \left(\frac{B}{\sqrt{A^2+B^2}}\right)^2 = 1$, there is an angle θ_0 such that $\frac{A}{\sqrt{A^2+B^2}} = \cos\theta_0$, $\frac{B}{\sqrt{A^2+B^2}} = \sin\theta_0$. Taking this into account, the equation of the trajectory acquires the form (2.61)

$$r(\theta) = \frac{l^2(\alpha m)^{-1}}{1 + \sqrt{1 + \frac{2l^2E}{\alpha^2 m}}\cos(\theta + \theta_0)}. \tag{2.71}$$

2.4 Picture of Motion in Phase Space

Here we illustrate some advantages of Hamiltonian formalism as compared with the Lagrangian one. In particular it will be seen that a general solution to Hamiltonian equations has useful interpretations in the framework of hydrodynamics and differential geometry.

General solution as the phase-space flux. Hamiltonian equations (2.50) represent a normal system of $2n$ first-order differential equations for $2n$ variables $z^i(\tau)$.

According to the general theory of differential equations, the theorem of existence and uniqueness of a solution holds for the case: for given numbers z_0^i, locally there is a unique solution $z^i(\tau)$ of the system, which obeys the *initial conditions*: $z^i(0) = z_0^i$. Let us recall also the definition of a general solution: $2n$ functions of $2n+1$ variables $z^i(\tau, c^j)$ are called a *general solution* of the system (2.50), if: (a) they obey the system for all c^i; (b) for given initial conditions z_0^i, there are numbers $\tilde{c}^j$ such that $z^i(0, \tilde{c}^j) = z_0^i$.

Owing to the above-mentioned theorem, a general solution to the normal system contains all particular solutions (trajectories) of the system, any one of them appearing after the appropriate fixation of the constants c^i.

These results imply a remarkable picture of motion in phase space: trajectories of the Hamiltonian system (2.50) do not intercept each other. To confirm this, let us suppose that two trajectories have interception at some point z_0^i. These numbers can be taken as initial conditions of the problem (2.50), and, according to the theorem, there is only one trajectory which passes through z_0^i, contrary to the initial supposition. Thus, trajectories of a Hamiltonian system in phase space form a flow, similarly to the picture of the motion of a fluid. Moreover, the "fluid" turns out to be incompressible, see Sect. 4.4.1. Note that it is very different from the corresponding picture of motion in the configuration space; see Fig. 2.1 on page 92.

Geometric interpretation of the symplectic matrix. In contrast to Lagrangian equations, Hamiltonian ones have a simple interpretation in the framework of differential geometry. Let us consider the right-hand sides of Hamiltonian equations as components H^i of a vector field in the phase space: $H^i(z^k) \equiv \omega^{ij} \frac{\partial H}{\partial z^j}$. Then the Hamiltonian equations $\dot{z}^i = H^i(z)$ state that any solution to equations of motion is a trajectory of this vector field (according to differential geometry, a line is the trajectory of a given vector field, if vectors of the field are tangent vectors to the line at each point). *Hamiltonian vector field* H^i also has certain interpretation. Let $H(z) = const$ represent a surface of constant energy. Then the vector field $H_i = \frac{\partial H}{\partial z^i} \equiv (grad\ H)_i|_{H=const}$ is normal to the surface at each point. The scalar product of H^i with the vector $grad\ H$ vanishes: $H^i(grad\ H)_i = \partial_j H \omega^{ji} \partial_i H = 0$,

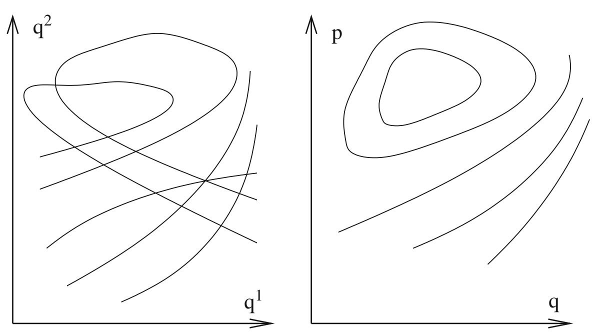

Fig. 2.1 Trajectory flows on configuration and phase spaces

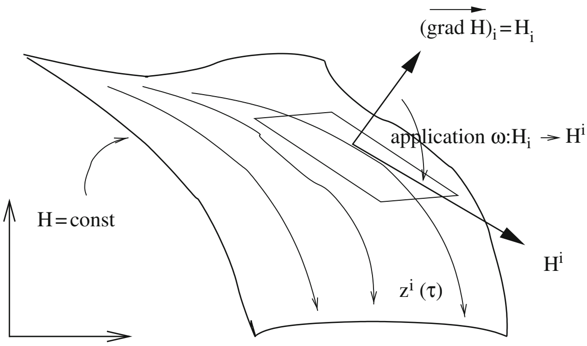

Fig. 2.2 Solutions lie on the surfaces of constant energy of the phase space. They are trajectories of the Hamiltonian vector field, constructed from H: $H^i(z^k) = \omega^{ij}(grad\ H)_j$

that is, the Hamiltonian vector field H^i is tangent to the surface. Hence each trajectory $z^i(\tau)$ lies on one of the surfaces of constant energy, as should be the case, see Fig. 2.2 on page 93. Now, observe the remarkable role played by the symplectic matrix ω^{ij}. There is a whole hyperplane of the vectors, which are normal to $grad\ H$ at a given point. It is the matrix ω that transforms the normal vector $grad\ H$ into the tangent vector to a trajectory!

Note that in terms of the coordinates z^i the vector field H^i is *divergenceless*: $\partial_i H^i = 0$. Now, let us consider the field $H^i(z^k)$ in the coordinates: $z_j \equiv z^l \omega_{lj}$. First, the Hamiltonian, as a function of z_i, is $H(z_i) \equiv H(z^i(z_j))$.

Exercise
Write $H(z^i) = q - p^2$ in terms of the variables z_i.

Further, since $\omega^{ik}\frac{\partial}{\partial z^k} z_j = \delta^i{}_j$, the derivative associated with z_j is $\partial^i = \frac{\partial}{\partial z_i} \equiv \omega^{ik}\frac{\partial}{\partial z^k}$. The Hamiltonian vector field $H^i(z_k)$ in these coordinates is $H^i(z_k) = \partial^i H(z_k)$, and turns out to be *curl-free (conservative)*: $\partial^j H^i - \partial^i H^j = 0$. This fact will be explored in Chap. 4.

2.5 Conserved Quantities and the Poisson Bracket

Definition 2 A function $Q(z^i, \tau)$ is called an *integral of motion*, if for any solution $z^i(\tau)$ of the Hamiltonian equations, Q retains a constant value:

$$Q(z(\tau), \tau) = c, \quad \text{or} \quad \frac{d}{d\tau} Q = 0 \quad \text{on-shell}. \tag{2.72}$$

Here "on-shell" stands for "for an arbitrary solution to equations of motion", while "off-shell" means "for an arbitrary function $z^i(\tau)$". Of course c may change when we pass from one trajectory to another. In the current literature, integral of motion is also referred to as *first integral, constant of motion, conserved quantity, conservation law, charge or dynamical invariant* – according to taste. Hereafter we use the term (conserved) charge, as the shortest among these expressions.

There is an important necessary and sufficient condition for a quantity Q to be a conserved charge.

Assertion $Q(z^i, \tau)$ represents a conserved charge if and only if

$$\frac{\partial Q}{d\tau} + \{Q, H\} = 0 \qquad \text{off-shell.} \tag{2.73}$$

In particular, the quantity $Q(z^i)$ (without manifest dependence on τ) is conserved if and only if its bracket with the Hamiltonian vanishes

$$\{Q, H\} = 0. \tag{2.74}$$

Proof We write identically

$$\frac{dQ}{d\tau} = \frac{\partial Q}{\partial \tau_0} + \{Q, H\} + \frac{\partial Q}{\partial z^i}(\dot{z}^i - \{Q, H\}), \tag{2.75}$$

so the condition (2.73) implies (2.72). Conversely, supposing that (2.72) is true, we have $\frac{\partial Q}{d\tau} + \{Q, H\} = 0$, on-shell. Given a phase-space point z_0^i and a value τ_0, let $z^i(\tau)$ represent the trajectory that passes through z_0^i at the instant τ_0. Inserting the solution into the equation and taking $\tau = \tau_0$ we obtain $\frac{\partial Q(\tau_0, z_0)}{\partial \tau} + \{Q(\tau_0, z_0), H(z_0, \tau_0)\} = 0$ for any τ_0, z_0, as has been stated.

An example of a charge is the Hamiltonian of a conservative system (see page 87). The search for the charges turns out to be an important task. From a pragmatic point of view, knowledge of them allows us to simplify (sometimes to solve) equations of motion of a system (it is sufficient to recall that conservation of angular momentum allows us to reduce the three-dimensional Kepler problem to a two-dimensional one). Let us point out also that in quantum theory the concept of a trajectory does not survive and is replaced by an abstract state space associated with the system. But the notion of conserved charges survives, and they play a crucial role in the interpretation of the state space, establishing a correspondence between the states and physical particles.

A powerful method for obtaining charges for a system which exhibits certain symmetries is provided by the Noether theorem, which is discussed in Chap. 7. Here we describe some general properties of a set of charges.

If Q is a charge, an arbitrary function $f(Q)$ will also be a charge. If Q_1, Q_2 are charges, their product and linear combinations with numerical coefficients also represent charges. It is convenient to introduce the notion of independent charges as follows: the charges $Q_\alpha(z^i, \tau)$, $\alpha = 1, 2, \dots, k \le 2n$ are called *functionally independent*, if

$$rank \frac{\partial Q_\alpha}{\partial z^i} = k. \tag{2.76}$$

This implies that the expressions $Q_\alpha(z^i, \tau) = c_\alpha$ can be resolved with respect to k variables z^α among z^i:

$$z^\alpha = G^\alpha(z^a, c_\alpha, \tau), \tag{2.77}$$

where z^a are the remaining variables of the set z^i. As will be discussed in Sect. 7.9, knowledge of k functionally independent charges immediately reduces the order of equations of motion by k units (that is, there is an equivalent system with the total number of derivatives being $2n - k$).

It is a simple matter to confirm the existence of $2n$ independent charges for a given dynamical system. Let the functions $f^i(\tau, c_j)$ represent a general solution to the Hamiltonian equations. This implies, in particular, that det $\frac{\partial f^i}{\partial c_j} \neq 0$. If we write the equations $z^i = f^i(\tau, c_j)$, they can be resolved with respect to c: $Q_j(z^i, \tau) = c_j$, giving the functions $Q_j(z^i, \tau)$. By construction, substitution of any solution $z^i(\tau)$ into Q turns them into constants. Hence we have obtained $2n$ independent combinations of z^i, τ, namely $Q_j(z^i, \tau)$, which do not depend on time for their solutions, and thus represent the conserved charges.

Of course, in practice the problem is the opposite: it is interesting to reveal as many charges as possible by independent methods, and use them to search for a general solution to equations of motion. In particular, inverting the previous discussion, we conclude that the knowledge of $2n$ independent charges is equivalent to knowledge of the general solution.

The set of charges is endowed with a remarkable algebraic structure in relation to the Poisson bracket: the bracket of two charges is also a charge. This is proved by direct computation

$$\begin{aligned}
&\frac{d}{d\tau}\{Q_1, Q_2\} = \frac{\partial}{\partial\tau}\{Q_1, Q_2\} + \{\{Q_1, Q_2\}, H\} = \\
&\left\{\frac{\partial Q_1}{\partial\tau}, Q_2\right\} + \left\{Q_1, \frac{\partial Q_2}{\partial\tau}\right\} - \{\{Q_2, H\}, Q_1\} - \{\{H, Q_1\}, Q_2\} = \\
&\left\{\frac{\partial Q_1}{\partial\tau} + \{Q_1, H\}, Q_2\right\} + \left\{Q_1, \frac{\partial Q_2}{\partial\tau} + \{Q_2, H\}\right\} = 0.
\end{aligned} \tag{2.78}$$

Here the Jacobi identity was used for the transition from the first to the second line. The last line is equal to zero since Q_1, Q_2 obey Eq. (2.73). Thus $Q_3 \equiv \{Q_1, Q_2\}$ is conserved. Of course, it can be identically null or can be functionally dependent on Q_1, Q_2. If not, the Poisson bracket can be used to generate new charges from the known ones.

As an illustration, consider a free-moving particle, with the Hamiltonian $H = \frac{1}{2m}(p^i)^2$, $i = 1, 2, 3$, and the corresponding Hamiltonian equations $\dot{x}^i = \frac{1}{m}p^i$, $\dot{p}^i = 0$. Besides the Hamiltonian, the conserved charges are the momenta $p^i = c^i = \text{const}$ (as follows from their equations), and angular momentum $L^i = \epsilon^{ijk}x^j p^k = d^i = \text{const}$ (since on-shell $\dot{L}^i = \frac{1}{2m}\epsilon^{ijk}p^j p^k \equiv 0$). H can be omitted, since it forms a functionally dependent set with p^i. As to the remaining six charges, only five of them are functionally independent (imagine that all they are independent. Then it would be possible to solve the equations $Q_i(z^i) = c_i$, obtaining a general solution to the equations of motion in the form $z^i = f^i(c_i)$, and arriving at the rather strange result that the particle cannot move! Of course, their dependence can be verified by direct computation of the corresponding Jacobian). By choosing p^i and L^2, L^3 as independent quantities, we find the dynamics of p^i, x^2, x^3 in terms of x^1: $p^i = c^i$, $x^2 = \frac{c^2}{c^1}x^1 - \frac{d^3}{c^1}$, $x^3 = \frac{c^3}{c^1}x^1 + \frac{d^2}{c^1}$. Thus, to find a general solution to the equations of motion, we need to solve only one of them, namely $\dot{x}^1 = \frac{c^1}{2m}$, which gives the time-dependent charge $x^1 = \frac{c^1}{2m}\tau + b$.

Exercises

1. Compute the number of functionally independent charges for the case of a free particle in n-dimensional space, $n > 3$.
2. Confirm the Poisson bracket algebra of the charges:

$$\{L^i, L^j\} = \epsilon^{ijk}L^k, \qquad \{L^i, p^j\} = \epsilon^{ijk}p^k. \tag{2.79}$$

2.6 Phase Space Transformations and Hamiltonian Equations

In many interesting cases, the Lagrangian equations can be solved with use of the coordinate transformations $q \to q'(q)$ in the configuration space. In particular, if the system in question exhibits certain symmetries, they can be taken into account to search for adapted coordinates. This often leads to separation of variables in Lagrangian equations, much simplifying the problem. Well-known examples are the use of polar coordinates in the Kepler problem and the use of center-of-mass variables in the two-body problem. The Hamiltonian formulation gives supplementary possibilities due to the fact that a set of transformations in the phase space is much larger, allowing us to mix position and velocity variables: $q \to q'(q, p)$, $p \to p'(q, p)$. In this section we find out how Hamiltonian equations transform under phase-space transformations. It will be seen that an arbitrary transformation spoils the canonical form of the Hamiltonian equations. So it is reasonable to choose the subset which preserves their form. Transformations of this subset are called canonical transformations, and are discussed in the next section.

Let $\varphi^i(z^j, \tau)$ represent $2n$ given functions of $2n+1$ variables, with the property

$$\det \frac{\partial \varphi^i}{\partial z^j} \neq 0, \quad \text{for any} \quad \tau. \tag{2.80}$$

Starting from the original parametrization z^i of phase space, functions φ^i with fixed τ_0 can be used to define another parametrization z'^i, namely

$$z'^i = \varphi^i(z^j, \tau_0). \tag{2.81}$$

According to the condition (2.80), *transformation of the coordinates*[5] $z^i \longrightarrow z'^i$ is invertible: the expressions (2.81) can be resolved with relation to z^i, with the result being

$$z^i = \psi^i(z'^j, \tau_0). \tag{2.82}$$

By construction, there are identities

$$\varphi^i(\psi(z', \tau_0), \tau_0) \equiv z'^i, \qquad \psi^i(\varphi(z, \tau_0), \tau_0) \equiv z^i. \tag{2.83}$$

From this we obtain more identities

$$\begin{aligned}
\left.\frac{\partial \varphi^k(z, \tau)}{\partial z^i}\right|_{z=\psi(z',\tau)} \frac{\partial \psi^i(z', \tau)}{\partial z'^j} &= \delta^k{}_j, \\
\left.\frac{\partial \varphi^i(z, \tau)}{\partial \tau}\right|_{z=\psi(z',\tau)} &= -\left.\frac{\partial \varphi^i(z, \tau)}{\partial z^j}\right|_{z=\psi(z',\tau)} \frac{\partial \psi^j(z', \tau)}{\partial \tau}, \\
\left.\frac{\partial \psi^i(z', \tau)}{\partial \tau}\right|_{z'=\varphi(z,\tau)} &= -\left.\frac{\partial \psi^i(z', \tau)}{\partial z'^j}\right|_{z'=\varphi(z,\tau)} \frac{\partial \varphi^j(z, \tau)}{\partial \tau}.
\end{aligned} \tag{2.84}$$

The first identity relates *Jacobi matrices* of inverse and direct transformations: the matrices turn out to be opposites. The second identity relates derivatives with respect to τ of the direct (φ) and the inverse (ψ) transformations. The third identity differs from the second one by changing $\varphi \leftrightarrow \psi$, as should be the case (it is a matter of convenience which transformation is called the "direct" and the "inverse" one).

---

[5] In configuration space, transformations of the type (2.81) with manifest dependence on τ are well-known and have a clear physical interpretation. For example, a Galilean transformation $x^i \to x'^i = x^i + v^i\tau + a^i$ gives the relationship between the coordinates of inertial frames, with relative velocity being v^i.

Coordinate transformation (2.81) implies an *induced map* in the space of functions (curves) $z^i(\tau)$

$$\varphi^i : z^i(\tau) \longrightarrow z'^i(\tau) \equiv \varphi^i(z^j(\tau), \tau). \tag{2.85}$$

So, changing the description in terms of $z^i(\tau)$ to a description in terms $z'^i(\tau)$, it is said that a *phase-space transformation* has been carried out: $z^i \to z'^i$. The function ψ gives an inverse transformation

$$\psi^i : z'^i(\tau) \longrightarrow z^i(\tau) \equiv \psi^i(z'^j(\tau), \tau), \tag{2.86}$$

and we have $\varphi^i(\psi(z'(\tau), \tau), \tau) \equiv z'^i(\tau),\ \psi^i(\varphi(z(\tau), \tau), \tau) \equiv z^i(\tau)$.

Comment As has already been mentioned, for a fixed τ the expression (2.81) has a clear geometric interpretation as a phase-space transformation $z^i \to z'^i$. Although it is not strictly necessary, it is convenient to discuss two different interpretations of (2.81), (2.85) with a varying τ.

(A) A geometric interpretation of (2.81) can be obtained in the *extended phase space* $\mathbb{R}^{2n+1} = \mathbb{R}(\tau) \bigotimes \mathbb{R}^{2n}(z^i)$ with coordinates $(\tau,\ z^i)$. In this space we can change the parametrization as follows

$$\begin{pmatrix} \tau \\ z^i \end{pmatrix} \leftrightarrow \begin{pmatrix} \tau' \\ z'^i \end{pmatrix}, \quad \begin{cases} \tau' = f(z, \tau) \\ z'^i = \varphi^i(z, \tau). \end{cases} \tag{2.87}$$

Now, Eq. (2.81) is a particular case of this coordinate transformation, with $\tau' \equiv \tau$. Let $\tau = \tau(s),\ z^i = z^i(s)$ represent parametric equations of a curve in $\mathbb{R}^{2n+1}$. The particular case is $\tau = s, z^i = z^i(s)$, where one of the coordinates, namely τ, was chosen as the parameter. In this sense the functions $z^i(\tau)$ can be considered as parametric equations of a curve in $\mathbb{R}^{2n+1}$. Then $z'^i = \varphi^i(z(\tau), \tau)$ are equations of this curve in the coordinate system (τ, z'^i).

(B) Sometimes it will be useful to treat certain combinations of the functions $\varphi^i(z^j, \tau)$ (2.81) as components of a (time dependent) vector field on the phase space z^i. In particular, we discuss below curl-free vector fields, see Eq. (3.14).

Starting from Hamilton equations

$$\dot{z}^i = \omega^{ij} \frac{\partial H(z^k)}{\partial z^j}, \qquad i, j = 1, 2, \ldots, 2n, \tag{2.88}$$

let us ask, what is the form they acquire in parametrization (2.81). In other words, if $z(\tau)$ obeys (2.88), what equations hold for $z'^i(\tau)$ defined by Eq. (2.85)? The answer is given by the following:

Affirmation Let $z'^i = \varphi^i(z^j, \tau)$ be a phase-space transformation. The system (2.88) and the following one

$$\dot{z}'^k = \left(\frac{\partial\varphi^k}{\partial z^i}\omega^{ij}\frac{\partial\varphi^l}{\partial z^j}\frac{\partial H(\psi(z',\tau))}{\partial z'^l} + \frac{\partial\varphi^k}{\partial\tau} \right)\Bigg|_{z=\psi(z',\tau)}, \tag{2.89}$$

are equivalent in the following sense:

(a) if $z^i(\tau)$ obeys (2.88) then $z'^i(\tau) \equiv \varphi^i(z(\tau),\tau)$ obeys (2.89). (b) if $z'^i(\tau)$ obeys (2.89) then $z^i(\tau) \equiv \psi^i(z'(\tau),\tau)$ obeys (2.88).

Let us prove item (a). According to (2.8), equations for z' arise after substitution of z in the form (2.86) into (2.88). This leads immediately to Eq. (2.89). The same result can be obtained by direct computation of the derivative (bearing in mind the identity $\dot{z}^i(\tau) \equiv \omega^{ij}\frac{\partial H(z)}{\partial z^j}\Big|_{z(\tau)}$)

$$\begin{aligned}
\dot{z}'^k(\tau) = (\varphi^k(z(\tau),\tau))^{\cdot} = \frac{\partial\varphi^k(z,\tau)}{\partial z^i}\Bigg|_{z(\tau)}\dot{z}^i(\tau) + \frac{\partial\varphi^k(z,\tau)}{\partial\tau}\Bigg|_{z(\tau)} = \\
\left(\frac{\partial\varphi^k(z,\tau)}{\partial z^i}\omega^{ij}\frac{\partial H(\psi(\varphi(z,\tau),\tau))}{\partial z^j} + \frac{\partial\varphi^k(z,\tau)}{\partial\tau}\right)\Bigg|_{z(\tau)} = \\
\left(\frac{\partial\varphi^k(z,\tau)}{\partial z^i}\omega^{ij}\ \frac{\partial H(\psi(z',\tau))}{\partial z'^l}\Bigg|_{z'=\varphi(z,\tau)}\frac{\partial\varphi^l(z,\tau)}{\partial z^j} + \frac{\partial\varphi^k(z,\tau)}{\partial\tau}\right)\Bigg|_{z(\tau)} = \\
\left(\left(\frac{\partial\varphi^k}{\partial z^i}\omega^{ij}\frac{\partial\varphi^l}{\partial z^j}\frac{\partial H(\psi(z',\tau))}{\partial z'^l} + \frac{\partial\varphi^k}{\partial\tau}\right)\Bigg|_{z=\psi(z',\tau)}\right)\Bigg|_{z'(\tau)}
\end{aligned} \tag{2.90}$$

For the transition to the last line we have used the equality: $z(\tau) = \psi(z'(\tau),\tau) = \psi(z',\tau)\big|_{z'(\tau)}$. Item (b) can be proved in a similar fashion.

Hereafter we use simplified notation, similar to that used in differential geometry. Instead of $z'^i = \varphi^i(z^j,\tau)$ and $z^i = \psi^i(z'^j,\tau)$ we write

$$z'^i = z'^i(z^j,\tau), \qquad z^i = z^i(z'^j,\tau), \tag{2.91}$$

Thus the new coordinate (value of function) and the transition function itself are denoted by the same symbol. The notation for partial derivatives[6] is

$$\begin{aligned}
\frac{\partial}{\partial z^i} \equiv \partial_i, \qquad \omega^{ij}\frac{\partial}{\partial z^j} \equiv \partial^i, \\
\frac{\partial}{\partial z'^i} \equiv \partial'_i, \qquad \omega^{ij}\frac{\partial}{\partial z'^j} \equiv \partial'^i, \\
\frac{\partial}{\partial\tau} \equiv \partial_\tau.
\end{aligned} \tag{2.92}$$

Also, we sometimes omit the operation of substitution:

---

6 ∂^i represents the usual partial derivative with respect to variable $z_l \equiv z^k\omega_{kl}$, since $\partial^i(z^k\omega_{kl}) = \delta^i_l$.

$$A(z)|_{z(z')} \longrightarrow A(z) \; or \; A(z)|. \tag{2.93}$$

If the left and right hand sides of an expression have wrong "balance of variables", we need to substitute $z(z')$ on the left or on the right hand side. In this notation we can write, for example

$$z'^i(z(z', \tau_0), \tau_0) \equiv z'^i \quad \text{instead of} \quad (2.83). \tag{2.94}$$

The identities (2.84) can now be written as follows

$$\begin{aligned} \frac{\partial z'^k}{\partial z^i}\frac{\partial z^i}{\partial z'^j} &= \delta^k{}_j \\ \frac{\partial z'^i(z,\tau)}{\partial \tau} &= -\frac{\partial z'^i}{\partial z^j}\frac{\partial z^j(z',\tau)}{\partial \tau}, \\ \frac{\partial z^i(z',\tau)}{\partial \tau} &= -\frac{\partial z^i}{\partial z'^j}\frac{\partial z'^j(z,\tau)}{\partial \tau}, \end{aligned} \tag{2.95}$$

where, for example, the last equation implies substitution of $z'(z,\tau)$ on l.h.s. and in the first term on r.h.s. Equivalently, we can substitute $z(z',\tau)$ in the last term on r.h.s. Hamiltonian equations for z'^i (2.89) acquire the form

$$\dot{z}'^k = \{z'^k, \; z'^l\}_z\Big|_{z(z',\tau)} \frac{\partial H(z(z',\tau))}{\partial z'^l} + \frac{\partial z'^k(z,\tau)}{\partial \tau}\Big|_{z(z',\tau)}, \tag{2.96}$$

where $\{z'^k, \; z'^l\}_z$ is the Poisson bracket computed with respect to z.

2.7 Definition of Canonical Transformation

From comparison of Eqs. (2.88) and (2.89) we conclude that phase-space transformation generally does not preserve the initial form of Hamiltonian equations. It justifies the following

Definition 3 The transformation $z'^i = \varphi^i(z^j, \tau)$ is called *canonical* if for any Hamiltonian system the corresponding induced transformation (2.85) preserves the canonical form of the Hamiltonian equations:

$$\dot{z}^i = \omega^{ij}\frac{\partial H}{\partial z^j} \xrightarrow{z \to z'} \dot{z}'^i = \omega^{ij}\frac{\partial \tilde{H}(z',\tau)}{\partial z'^j}, \; \text{any } H, \; \text{some } \tilde{H}. \tag{2.97}$$

It will be seen below that $\tilde{H}$ is related to H according to a simple rule (in particular, for the case of time-independent canonical transformation, we have $\tilde{H}(z') = cH(z(z')), \; c = \text{const}$).

Transformations that do not alter a given Hamiltonian system are called *canonoid transformations*.

By construction, the composition of canonical transformations is also a canonical transformation : if $z \to z'=z'(z,\tau)$, and $z' \to z''=z''(z',\tau)$ are canonical, then $z \to z'' = z''(z'(z,\tau),\tau)$ is a canonical transformation. The set of canonical transformations form a group, with a product defined by this law of composition. This allows us to describe the ambiguity present in the Hamiltonization procedure: besides Eq. (2.21), any change of the form $(q, v) \to (q'(q, p(q, v), \tau), p'(q, p(q, v), \tau)$, where $p(q, v) = \frac{\partial L}{\partial v}$ and $q'(q, p, \tau)$, $p'(q, p, \tau)$ is a canonical transformation, transforms Eq. (2.19) into the Hamiltonian system.

From Eqs. (2.96) and (2.97) it follows that the canonical transformation $z'(z, \tau)$ obeys

$$\begin{aligned} &\{z'^k,\ z'^l\}_z\Big|_{z(z',\tau)} \partial'_l H(z(z',\tau)) + \partial_\tau z'^k(z,\tau)\Big|_{z(z',\tau)} \\ &= \omega^{kl}\partial'_l \tilde{H}(z',\tau), \quad \forall\, H \quad and \quad some\ \tilde{H}. \end{aligned} \tag{2.98}$$

From this expression we immediately obtain two useful consequences. First, taking derivative ∂'_k of Eq. (2.98) we have

$$\partial'_k\left(\{z'^k, z'^l\}_z\Big|_{z(z',\tau)}\right)\partial'_l H(z(z',\tau)) + \partial'_k(\partial_\tau z'^k(z,\tau)|_{z(z',\tau)}) = 0. \tag{2.99}$$

Since this is true for any H, the first and second terms vanish separately. In particular, the derivative of the Poisson bracket must be zero, hence $\{z'^k, z'^l\}_z\Big|_{z(z',\tau)} = c^{kl}(\tau)$, where c^{kl} does not depend on z^i. So, the substitution of $z(z',\tau)$ can be omitted, and we have

$$\{z'^k, z'^l\}_z = c^{kl}(\tau). \tag{2.100}$$

Second, denoting the left hand side of Eq. (2.98) by $J^{k'}$, it can be written as $J'^k = \partial'^k \tilde{H}$. From this it follows:

$$\partial'^i J'^j = \partial'^j J'^i. \tag{2.101}$$

In greater detail, the identity is

$$\begin{gathered} \partial'^i\left(\partial_\tau z'^j(z,\tau)\Big|_{z(z',\tau)}\right) - (i \leftrightarrow j) + \\ \left(\partial'^i W^{jl} - (i \leftrightarrow j)\right)\partial'_l H - \\ \left(W^{ik}\omega^{jl} - (i \leftrightarrow j)\right)\partial'^2_{kl} H = 0, \end{gathered} \tag{2.102}$$

where

$$W^{ij} \equiv \{z'^i, z'^j\}_z \Big|_{z(z',\tau)}. \tag{2.103}$$

Since this is true for any H, we write separately

$$\partial'^i \left(\partial_\tau z'^j(z,\tau)\Big|_{z(z',\tau)} \right) - (i \leftrightarrow j) = 0,$$
$$\partial'^a W^{bd} - \partial'^b W^{ad} = 0,$$
$$W^{ik}\omega^{jl} - W^{jk}\omega^{il} + W^{il}\omega^{jk} - W^{jl}\omega^{ik} = 0. \tag{2.104}$$

The Eqs. (2.100) and (2.104) hold for an arbitrary canonical transformation and will be the starting point for our analysis below. In particular, it will be shown in Chap. 4, that the system (2.104) is equivalent to a simple statement that the symplectic matrix is invariant under the canonical transformation (disregarding the constant c):

$$\frac{\partial z'^k}{\partial z^i}\omega^{ij}\frac{\partial z'^l}{\partial z^j} = c\omega^{kl}, \quad c = \text{const}, \tag{2.105}$$

Transformations with $c = 1$ are called *univalent canonical transformations*.

2.8 Generalized Hamiltonian Equations: Example of Non-canonical Poisson Bracket

Here we discuss the form that the Hamiltonian equations acquire in an arbitrary parametrization of the configuration-velocity space.

In Sect. 2.2 the Hamiltonian equations were written in terms of the Poisson bracket

$$\dot{z}^i = \{z^i, H\}, \qquad \{z^i, z^j\} = \omega^{ij}, \tag{2.106}$$

with the numeric matrix ω^{ij}, see Eq. (2.47). According to Eq. (2.89), after the time-independent transformation $z^i \to z'^i = z'^i(z^j)$, the equations read

$$\dot{z}'^i = W^{ij}\frac{\partial \tilde{H}(z')}{\partial z'^j} \equiv \{z'^i, \tilde{H}(z')\}^{(W)}, \tag{2.107}$$

where $\tilde{H}(z') \equiv H(z(z'))$, and W is now a z'-dependent matrix

$$W^{ij} = \frac{\partial z'^i}{\partial z^k}\omega^{kl}\frac{\partial z'^j}{\partial z^l}\Big|_{z(z')}. \tag{2.108}$$

This was used in Eq. (2.107) to define a *non-canonical Poisson bracket*

$$\{A(z'), B(z')\}^{(W)} \equiv \frac{\partial A}{\partial z'^i} W^{ij} \frac{\partial B}{\partial z'^j}. \tag{2.109}$$

It can be shown that this obeys all the properties (2.40), (2.41), (2.42), and (2.43) of the Poisson bracket. The Eqs. (2.107) with the non-canonical Poisson bracket are known as *generalized Hamiltonian equations*.

As an example, let us discuss the transformation

$$(q^a, p_b) \to (q'^a = q^a, p'_b = p_b + b_b(q)), \tag{2.110}$$

where $b_b(q)$ is a given function. The Eqs. (2.106) acquire the form

$$\begin{aligned} \dot{q}^a &= \frac{\partial \tilde{H}(q, p')}{\partial p'_a}, \\ \dot{p}'_a &= -\frac{\partial \tilde{H}(q, p')}{\partial q^a} + \left(\frac{\partial b_a}{\partial q^b} - \frac{\partial b_b}{\partial q^a}\right) v^b(q^a, p'_b - b_b), \end{aligned} \tag{2.111}$$

where

$$\begin{aligned} \tilde{H}(q, p') &\equiv H(q^a, p'_b - b_b(q)) \\ &= \left((p'_a - b_a)v^a - L(q, v)\right)\big|_{v(q, p'-b)}. \end{aligned} \tag{2.112}$$

Let us point out that the same result appears if we carry out the Hamiltonization procedure of the Eqs. (2.19) using the variable change

$$p'_a = \frac{\partial L(q, v)}{\partial v^a} + b_a(q), \tag{2.113}$$

instead of the standard one (2.21).

The normal system (2.111) is equivalent to the Euler-Lagrange equations for any given function $b_a(q)$.

In this case, the symplectic form of the non-canonical bracket (2.109) is given by the expression

$$W^{ij} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & W_{ab} \end{pmatrix}, \qquad W_{ab}(q) \equiv \frac{\partial b_a}{\partial q^b} - \frac{\partial b_b}{\partial q^a}. \tag{2.114}$$

This implies the fundamental brackets

$$\{q^a, q^b\}^{(W)} = 0, \ \{q^a, p'_b\}^{(W)} = \delta_{ab}, \ \{p'_a, p'_b\}^{(W)} = W_{ab}(q), \tag{2.115}$$

which obey the properties (2.40), (2.41), (2.42), and (2.43). The Eqs. (2.111) acquire the form

$$\dot{z}^i = \{z^i, \tilde{H}(z)\}^{(W)}, \quad \text{where} \quad z^i = (q^a, p'_b), \tag{2.116}$$

with the Hamiltonian (2.112).

Non-canonical brackets (2.109), (2.114), and (2.115) naturally appear in the description of a system with velocity-dependent interactions. As an example, consider the Lagrangian action of non-relativistic particle on an external electromagnetic background (see Sect. 1.12.2))

$$S = \int d\tau \left(\frac{1}{2}(\dot{q}^a)^2 + \dot{q}^a A_a(q) \right). \tag{2.117}$$

The standard definition of momentum $p_a = \frac{\partial L}{\partial \dot{q}^a} = \dot{q}_a + A_a(q)$ leads to the Hamiltonian

$$H(q, p) = \frac{1}{2}(p_a - A_a)^2, \tag{2.118}$$

which implies the Hamiltonian equations

$$\dot{q}^a = p_a - A_a \equiv \{q^a, H\}, \quad \dot{p}_a = (p_b - A_b)\frac{\partial A_b}{\partial q^a} \equiv \{p_a, H\}, \tag{2.119}$$

with the canonical Poisson bracket.

Now, using Eq. (2.113) as a definition of momentum: $p'_a = \dot{q}_a + A_a(q) + b_a(q)$, it is natural to take $b_a = -A_a$, which leads to the expression $p'_a = \dot{q}_a$. Then Eq. (2.112) gives the Hamiltonian

$$H(q, p) = \frac{1}{2}(p'_a)^2. \tag{2.120}$$

According to Eq. (2.111) the Hamiltonian equations are

$$\dot{q}^a = p_a \equiv \{q^a, H\}', \quad \dot{p}'_a = -F_{ab}p'_b \equiv \{p'_a, H\}', \tag{2.121}$$

with the non-canonical Poisson bracket

$$\{q^a, q^b\}' = 0, \quad \{q^a, p'_b\}' = \delta_{ab}, \quad \{p'_a, p'_b\}' = F_{ab}(q). \tag{2.122}$$

Here F is a field strength of the vector potential: $F_{ab} = \frac{\partial A_a}{\partial q^b} - \frac{\partial A_b}{\partial q^a}$. It is easy to see that both (2.119) and (2.121) imply the same Lagrangian equations $\ddot{q}^a = -F_{ab}\dot{q}^b$.

Note that the Hamiltonian (2.120) formally coincides with the free-particle one. In this sense, in the second formulation the interaction is encoded in the non-canonical Poisson bracket.[7] Inclusion of the velocity-dependent interactions into the bracket was suggested in [6].

Let us return to the Eqs. (2.110) and ask about the existence of a function $b_b(q)$ that preserves the canonical form of Hamiltonian equations. The Eqs. (2.111) will be in the canonical form if the last term vanishes.

Exercise
Show that $\left(\frac{\partial b_a}{\partial q^b} - \frac{\partial b_b}{\partial q^a}\right) v^b = 0$ implies $\frac{\partial b_a}{\partial q^b} - \frac{\partial b_b}{\partial q^a} = 0$.

In turn, the latter equation implies that $b_a = \frac{\partial g}{\partial q^a}$ for a function g. So, for a change of the form

$$q'^a = q^a, \qquad p'_a = p_a + \frac{\partial g(q)}{\partial q^a}, \tag{2.123}$$

we obtain the canonical equations with the Hamiltonian

$$H' = H\left(q^a, p'_b - \frac{\partial g}{\partial q^b}\right). \tag{2.124}$$

According to the terminology of Sect. 2.7, Eq. (2.123) represents an example of a canonical transformation.

Exercises

1. Show that the symplectic form (2.114) is invertible; find the inverse matrix W_{ij}. Show that the latter obeys the equation

$$\partial_{[k} W_{ij]} \equiv \partial_k W_{ij} + \partial_i W_{jk} + \partial_j W_{ki} = 0. \tag{2.125}$$

A two-form with this property is called a *closed form*.

---

[7] While in the second formulation the vector potential A_a is not explicitly presented (Eqs. (2.120), (2.121), and (2.122) depend on F, that is an electric $\vec{E}$ and magnetic $\vec{B}$ fields), it nevertheless reappears in the course of quantization. In fact, the operators, which reproduce the brackets (2.122), should contain A: $q^a \to \hat{q}^a = q^a$, $p'_a \to \hat{p}_a = \frac{\partial}{\partial q^a} + A_a$. This leads to the same Schrodinger equation as in the initial formulation, with explicit dependence on A. In turn, dependence of the Schrodinger equation on A has interesting consequences. Contrary to the conclusions of classical mechanics, four-vector electromagnetic potential can affect the motion of charged particles, even in the region where the electric and the magnetic fields vanish. This effect [25, 26], known as the Aharonov-Bohm effect, has been confirmed by experiment.

2. Let $W^{ij}(z)$ be an antisymmetric invertible matrix, with the inverse matrix obeying the Eq. (2.125). Show that the bracket (2.109), constructed from this W, obeys properties (2.40), (2.41), (2.42), and (2.43).
3. Show that the bracket (2.109) with W given by Eq. (2.108) obeys properties (2.40), (2.41), (2.42), and (2.43).
4. Find the non-canonical bracket and the generalized Hamiltonian equations in the initial parametrization (q^a, v^b) of the configuration-velocity space (see Sect. 2.1.2).

2.9 Hamiltonian Action Functional

Similarly to Lagrangian equations, Hamiltonian ones can be obtained from the principle of least action applied to an appropriately chosen action functional. The *Hamiltonian action functional* is given by

$$S_H = \int d\tau \left(p_a \dot{q}^a - H(q^a, p_b, \tau)\right). \tag{2.126}$$

A variational problem is formulated as follows. We look for a curve $z^i(\tau)$ with fixed initial and final positions $q^a(\tau_1) = q_1^a,\ q^a(\tau_2) = q_2^a$ and arbitrary momenta, that would give a minimum for the functional (see Fig. 2.3 on page 106). The variation of the functional is

$$\delta S_H = \int d\tau \left(\left(\dot{q}^a - \frac{\partial H}{\partial p_a}\right)\delta p_a - \left(\dot{p}_a + \frac{\partial H}{\partial q^a}\right)\delta q^a + \left.(p_a \delta q^a)\right|_{t_1}^{t_2}\right). \tag{2.127}$$

Owing to the boundary conditions we have: $\delta q^a(t_1) = \delta q^a(t_2) = 0$, so the last term vanishes. Therefore $\delta S_H = 0$ implies the Hamiltonian Eqs. (2.32).

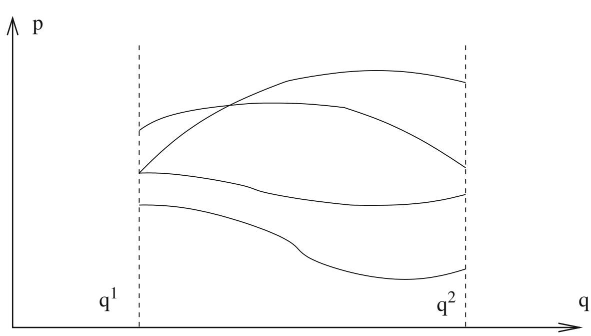

Fig. 2.3 Variational problem for the Hamiltonian action functional

Exercises

1. The addition of a total derivative term to the Lagrangian does not alter the Lagrangian equations of motion. Is the same true for the Hamiltonian action? See also the exercise on page 140.
2. Disregarding the boundary term, the Hamiltonian action can be written in the form

$$S_H = \int d\tau \left(\frac{1}{2} z^i \omega_{ij} \dot{z}^j - H(z^i) \right). \tag{2.128}$$

Is it possible to formulate a consistent variational problem for this functional, which should lead to the Hamiltonian equations?

2.10 Schrödinger Equation as the Hamiltonian System

Hamiltonian action appears in applications more often than one might expect. As an example, consider the quantum mechanics of a particle subject to the potential $V(t, x^i)$. The *Schrödinger equation* for the complex *wave function* $\Psi(t, x^i)$

$$i\hbar\dot{\Psi} = -\frac{\hbar^2}{2m}\triangle\Psi + V\Psi, \tag{2.129}$$

is equivalent to the system of two equations for two real functions (the real and imaginary parts of Ψ, $\Psi = \varphi + ip$). We have

$$\hbar\dot{\varphi} = -\left(\frac{\hbar^2}{2m}\triangle - V \right) p, \tag{2.130}$$

$$\hbar\dot{p} = \left(\frac{\hbar^2}{2m}\triangle - V \right) \varphi. \tag{2.131}$$

Remind the notation $\triangle = \frac{\partial^2}{\partial x^{i2}}$, $\vec{\nabla} = \frac{\partial}{\partial x^i}$, $\dot{\varphi} \equiv \partial_t \varphi = \frac{\partial \varphi(t, x^i)}{\partial t}$. We can treat $\varphi(t, x^i)$ and $p(t, x^i)$ as coordinate and conjugate momenta of the field φ at the space point x^i. Then the system has the Hamiltonian form $\dot{\varphi} = \{\varphi, H\}$, $\dot{p} = \{p, H\}$, with the Hamiltonian being

$$H = \frac{1}{2\hbar} \int d^3x \left(\frac{\hbar^2}{2m} (\vec{\nabla}\varphi\vec{\nabla}\varphi + \vec{\nabla}p\vec{\nabla}p) + V(\varphi^2 + p^2) \right). \tag{2.132}$$

Hence the Eqs. (2.130) and (2.131) arise from the variation problem with the Hamiltonian action obtained according to Eq. (2.126)

$$S_H = \int dt d^3x \, [\pi\dot{\varphi} - \frac{1}{2\hbar}\left(\frac{\hbar^2}{2m}(\vec{\nabla}\varphi\vec{\nabla}\varphi + \vec{\nabla}p\vec{\nabla}p) + V(\varphi^2 + p^2)\right)]. \quad (2.133)$$

Disregarding the boundary term (see Exercise 2 in this relation), this can also be rewritten in terms of the wave function Ψ and its complex conjugate Ψ^*

$$S_H = \int dt d^3x \left[\frac{i\hbar}{2}(\Psi^*\dot{\Psi} - \dot{\Psi}^*\Psi) - \frac{\hbar^2}{2m}\vec{\nabla}\Psi^*\vec{\nabla}\Psi - V\Psi^*\Psi\right]. \quad (2.134)$$

2.10.1 Lagrangian Action Associated with the Schrödinger Equation

Due to the Hamiltonian nature of the Schrodinger equation, it is natural to search for a Lagrangian formulation of the system (2.130) and (2.131), that is a second-order equation with respect to the time derivative[8] for the real function $\varphi(t, x^i)$. According to Sect. 2.1.4, we need to solve (2.130) with respect to p and then to substitute the result either in Eq. (2.131) or into the Hamiltonian action (2.133). This leads immediately to the rather formal non-local expression $p = \hbar\left(-\frac{\hbar^2}{2m}\triangle - V\right)^{-1}\partial_t\varphi$. So, the Schrödinger system cannot be obtained starting from a (nonsingular) Lagrangian. Nevertheless, for the case of time-independent potential $V(x^i)$, there is a Lagrangian field theory with the property that any solution to the Schrödinger equation can be constructed from a solution to this theory. To find it let us look for solutions of the form

$$\Psi = -\left(\frac{\hbar^2}{2m}\triangle - V\right)\phi + i\hbar\dot{\phi}, \quad (2.135)$$

where $\phi(t, x^i)$ is a *real* function. Inserting (2.135) into (2.129) we conclude that Ψ will be a solution to the Schrödinger equation if ϕ obeys the equation

$$\hbar^2\ddot{\phi} + \left(\frac{\hbar^2}{2m}\triangle - V\right)^2\phi = 0, \quad (2.136)$$

which follows from the Lagrangian action

[8] In fact, the problem has already been raised by Schrödinger [27]. Equation (2.136) below was tested by Schrödinger as a candidate for the wave function equation and then abandoned.

$$S = \int dt d^3x \left[\frac{\hbar}{2}\dot{\phi}\dot{\phi} - \frac{1}{2\hbar}\left(\frac{\hbar^2}{2m}\triangle - V\right)\phi\left(\frac{\hbar^2}{2m}\triangle - V\right)\phi \right]. \quad (2.137)$$

This can be treated as the classical theory of field ϕ on the given external background $V(x^i)$. The action contains Planck's constant as a parameter. After the rescaling $(t, x^i, \phi) \to (\hbar t, \hbar x^i, \sqrt{\hbar}\phi)$ it appears in the potential only, $V(\hbar x^i)$, and thus plays the role of a coupling constant of the field ϕ with the background.

The formula (2.135) implies that after introduction of the field ϕ into the formalism, its mathematical structure becomes analogous to that of electrodynamics. The dynamics of the magnetic $\vec{B}$ and electric $\vec{E}$ fields is governed by first-order Maxwell equations with respect to the time variable. Equivalently, we can use the vector potential A_a, which obeys the second-order equations following from the Lagrangian action discussed in Sect. 1.12.2. A_a represents the potential for magnetic and electric fields, generating them according to $\vec{B} = [\vec{\nabla}, \vec{A}]$, $\vec{E} = -\frac{1}{c}\partial_t\vec{A}$. Similarly to this, the field ϕ turns out to be a potential for the wave function, generating its real and imaginary parts according to Eq. (2.135), see also Fig. 2.4 on the page 109.

In quantum mechanics the quantity $\Psi^*\Psi$ has an interpretation as a probability density, that is the expression $\Psi^*(t, x^i)\Psi(t, x^i)d^3x$ represents the probability of finding a particle in the volume d^3x around the point x^i at the instant t. According to the formula (2.135), we write

$$\Psi^*\Psi = \hbar^2(\dot{\phi})^2 + \left[\left(-\frac{\hbar^2}{2m}\triangle + V\right)\phi\right]^2 = 2\hbar E, \quad (2.138)$$

Electrodynamics	Quantum mechanics
There is the Lagrangian formulation in terms of A_a	The same in terms of ϕ
A_a represents the potential for magnetic and electric fields, one has $\vec{B} = \nabla \times \vec{A}$, $\vec{E} = -\frac{1}{c}\partial_t\vec{A}$	$\Psi = \varphi + ip = -(\triangle - V)\phi + i\hbar\partial_t\phi$
While the Maxwell equations are written in terms of $\vec{B}$, $\vec{E}$, the field $\vec{E}$ is the conjugate momenta for $\vec{A}$ but not for $\vec{B}$	While the Schrödinger equation is written in terms of φ, p, the field p is the conjugate momenta for ϕ but not for φ
Maxwell equations form the generalized Hamiltonian system with the Hamiltonian $\sim \vec{E}^2 + \vec{B}^2$	Schrödinger equation forms the generalized Hamiltonian system with the Hamiltonian $\sim p^2 + \varphi^2$

Fig. 2.4 Real field ϕ as the wave function potential

where $E = T + U$ is the energy density of the field φ. Eq. (2.138) states that the probability density is the energy density of the wave potential ϕ. So the preservation of probability is just an energy conservation law of the theory (2.137).

It is instructive to compare also the Hamiltonian equations of the theory (2.137)

$$\hbar\dot{\phi} = p, \qquad \hbar\dot{p} = -\left(\frac{\hbar^2}{2m}\triangle - V\right)^2 \phi, \tag{2.139}$$

with the Schrödinger system. Note the following correspondence among solutions to these systems: (a) If the functions φ, p obey Eq. (2.130), (2.131), then the functions $\phi \equiv \varphi$, $-\left(\frac{\hbar^2}{2m}\triangle - V\right) p$ obey Eq. (2.139). (b) If the functions ϕ, p obey Eqs. (2.139), then $\varphi \equiv -\left(\frac{\hbar^2}{2m}\triangle - V\right)\phi$, p obey Eqs. (2.130) and (2.131). The kernel of the map $(\varphi, p) \to (\phi, p)$ is composed of pure imaginary time-independent wave functions $\Psi = i\Pi(x^i)$, where Π is any solution to the stationary Schrödinger equation $\left(\frac{\hbar^2}{2m}\triangle - V\right)\Pi = 0$.

Any solution to the field theory (2.137) determines a solution to the Schrödinger equation according to Eq. (2.135). We should ask whether an arbitrary solution to the Schrödinger equation can be presented in the form (2.135). An affirmative answer can be obtained as follows.

Let $\Psi = \varphi + ip$ be a solution to the Schrödinger equation. Consider the expression (2.135) as an equation for determining ϕ

$$\dot{\phi} = \frac{1}{\hbar}p, \tag{2.140}$$

$$\left(\frac{\hbar^2}{2m}\triangle - V\right)\phi = -\varphi, \tag{2.141}$$

Here the right-hand sides are known functions. Take Eq. (2.141) at $t = 0$, $\left(\frac{\hbar^2}{2m}\triangle - V\right)\phi = -\varphi(0, x^i)$. The elliptic equation can be solved (at least for the analytic function $\varphi(x^i)$); let us denote the solution as $C(x^i)$. Then the function

$$\phi(t, x^i) = \frac{1}{\hbar}\int_0^t d\tau p(\tau, x^i) + C(x^i), \tag{2.142}$$

obeys the Eqs. (2.140) and (2.141). They imply the desired result: any solution to the Shrödinger equation can be presented through the field ϕ and its momenta according to (2.135). Finally, note that Eqs. (2.140) and (2.141) together with Eqs. (2.130) and (2.131) imply that ϕ obeys Eq. (2.136).

Let us finish this section with one more comment. As we have seen, treating a Schrödinger system as a Hamiltonian one, it is impossible to construct the corresponding Lagrangian formulation owing to the presence of the spatial derivatives of momentum in the Hamiltonian. To avoid this problem, we can try to treat the

Schrödinger system as a generalized Hamiltonian system. We rewrite (2.130) in the form

$$\dot{\varphi} = \{\varphi, H'\}', \qquad \dot{p} = \{p, H'\}', \tag{2.143}$$

where H' is the "free field" generalized Hamiltonian

$$H' = \int d^3x \frac{1}{2\hbar}(p^2 + \varphi^2) = \int d^3x \frac{1}{2\hbar}\Psi^*\Psi, \tag{2.144}$$

and the non-canonical Poisson bracket is specified by

$$\begin{gathered} \{\varphi, \varphi\}' = \{p, p\}' = 0, \\ \{\varphi(t, x), p(t, y)\}' = -\left(\frac{\hbar^2}{2m}\triangle - V\right)\delta^3(x - y). \end{gathered} \tag{2.145}$$

In contrast to H, the Hamiltonian H' does not contain the spatial derivatives of momentum.

A non-canonical bracket turns out to be a characteristic property of singular Lagrangian theories discussed in Chap. 8. There we obtain a more systematic treatment of the observations made above: there is a singular Lagrangian theory subject to second class constraints underlying both the Schrödinger equation and the classical field theory (2.137).

2.10.2 Probability as a Conserved Charge Via the Noether Theorem

In quantum mechanics the quantity $\Psi^*\Psi$ has an interpretation of a probability density, that is the expression $\Psi^*(t, x^i)\Psi(t, x^i)d^3x$ represents the probability of finding a particle in the volume d^3x around the point x^i at the instant t. Consistency of the interpretation implies that, for the given solution $\Psi(t, x^i)$, the probability of finding the particle anywhere in space, $P(t) = \int_{\mathbb{R}^3} d^3x \Psi^*\Psi$ must be the same number at any instant (the number can be further normalized to be 1), or $\frac{dP}{dt} = 0$ for any solution. That is, P must be the conserved charge of the theory. In Chap. 7 we will discuss the Noether theorem that gives a deep relationship among the symmetry properties of an action and the existence of conserved charges for the corresponding equations of motion. Here we obtain this relationship for a particular example of the Schrödinger equation, showing that the preservation of probability can be considered as a consequence of a symmetry presented in the Hamiltonian action (2.134).

Given the number θ, let us make the following substitution

$$\Psi \to e^{i\theta}\Psi, \tag{2.146}$$

in the Hamiltonian action (2.134). Since this involves only products of a wave function with its complex conjugate, $\Psi^*\Psi$, the action does not change

$$S_H[e^{i\theta}\Psi] - S_H[\Psi] = 0. \tag{2.147}$$

According to Sect. 1.6, the action is invariant under (2.146). The symmetry transformation has a simple geometric interpretation as a rotation through the angle θ of a two-dimensional vector space spanned by (φ, p).

What are the consequences of the invariance? Take an expansion of $e^{i\theta}\Psi$ in the power series at $\theta = 0$, keeping only a linear term, $e^{i\theta}\Psi = \Psi + \delta\Psi$, where $\delta\Psi = i\Psi\theta$. Then Eq. (2.147) implies (confirm that!)

$$S_H[\Psi + \delta\Psi]|_{O(\theta)} - S_H[\Psi] = 0, \tag{2.148}$$

That is variation of the action vanishes as well. On other hand the variation can be computed according to the known formula (1.105), so we have

$$\delta S_H = \int dtd^3x \left[\left(i\hbar\dot{\Psi} + \frac{\hbar^2}{2m}\triangle\Psi - V\Psi\right)\delta\Psi^* + (c.c)\delta\Psi + \right.$$
$$\left.\frac{i\hbar}{2}\partial_t(\Psi^*\delta\Psi - \delta\Psi^*\Psi) + \frac{-\hbar^2}{2m}\partial_i(\delta\Psi^*\vec{\nabla}\Psi + \vec{\nabla}\Psi^*\delta\Psi)\right] = 0, \tag{2.149}$$

where $(c.c.)$ stands for a complex conjugation of the previous bracket. Supposing Ψ obeys the Schrödinger equation, the first and the second terms vanish, and (omitting the factor $-\hbar$) we conclude that

$$\partial_t J + \partial_i J^i = 0, \tag{2.150}$$

where

$$J = \Psi^*\Psi,$$
$$J^i = \frac{-i\hbar}{2m}(\Psi^*\vec{\nabla}\Psi - \vec{\nabla}\Psi^*\Psi). \tag{2.151}$$

Hence invariance of the action implies the *continuity equation* (2.150) that holds on solutions to the Schrödinger equation. It is further used to construct the conserved charge P integrating the quantity J

$$P = \int d^3x J = \int d^3x \Psi^*\Psi. \tag{2.152}$$

The total probability is preserved as a consequence of the continuity equation

$$\frac{dP}{dt} = \int_{\mathbb{R}^3} d^3x \partial_t J = -\int_{\mathbb{R}^3} d^3x \partial_i J^i = \int_{\partial\mathbb{R}^3} \vec{J} d\vec{S} = 0. \tag{2.153}$$

The third equality is due to Gauss's theorem while the last one follows from the standard supposition that Ψ vanishes in spatial infinity (a particle cannot escape to infinity during a finite time interval).

Exercises

1. Confirm the preservation of probability, $\frac{dP}{dt} = 0$, by direct computation with use of the Schrödinger equation.
2. Obtain the charge P using the Hamiltonian action functional (2.133).

2.11 Hamiltonization Procedure in Terms of First-Order Action Functional

Here we describe a very elegant Hamiltonization recipe [10, 28] based on manipulations with the Lagrangian action. Let

$$S = \int d\tau L(q^a, \dot{q}^a), \tag{2.154}$$

be a Lagrangian action of a non-singular system. Let us introduce an *extended phase space* parameterized by independent coordinates q^a, p_a, v^a. Starting from the Lagrangian given in Eq. (2.154), we can construct the following *first-order action* on the extended space

$$S_1 = \int d\tau \left[L(q^a, v^a) + p_a(\dot{q}^a - v^a)\right]. \tag{2.155}$$

This implies the equations of motion

$$\dot{q}^a = v^a, \qquad \dot{p}_a = \frac{\partial L(q, v)}{\partial q^a}, \qquad p_a = \frac{\partial L(q, v)}{\partial v^a}. \tag{2.156}$$

The last equation determines the conjugate momenta (see (2.21)), while the first two equations coincide with the first-order equations of motion for the initial action (2.154), see Eq. (2.26). So the action (2.155) represents an equivalent formulation for the theory (2.154). In this formulation, equations for canonical momenta (2.21) appear as part of the equations of motion. The remainder of the Hamiltonization recipe consists of using the third equation to expel v^a from the first two equations. The corresponding computations coincide with those made in Sect. 2.1.2, starting from Eq. (2.26), and give the Hamiltonian equations (2.32).

We finish this section with a comment on the formal relationship between the different actions. Let us take the first order action as a basic object. The Lagrangian action can be obtained from the first order one by using the first equation from (2.156).

Solving the last equation from (2.156), $v = v(q, p)$, and substituting the result into S_1, we obtain the Hamiltonian action.

We can also substitute p_a of the last equation from (2.156) into S_1 obtaining the following action in v, p space

$$S_v = \int d\tau \left[L(q^a, v^a) + \frac{\partial L(q, v)}{\partial v^a}(\dot{q}^a - v^a) \right]. \tag{2.157}$$

The corresponding equations of motion are

$$\begin{aligned} \frac{\delta S_v}{\delta q^a} &= \frac{\partial L}{\partial q^a} - \frac{d}{d\tau}\frac{\partial L}{\partial v^a} + \frac{\partial^2 L}{\partial q^a \partial v^b}(\dot{q}^b - v^b) = 0, \\ \frac{\delta S_v}{\delta v^a} &= \frac{\partial^2 L}{\partial v^a \partial v^b}(\dot{q}^b - v^b) = 0. \end{aligned} \tag{2.158}$$

For non-degenerate theory, they imply the Lagrangian equations for q^a. Hence the action S_v can also be used to analyze the system.

2.12 Hamiltonization of a Theory with Higher-Order Derivatives

Here we discuss a theory which involves the higher-order derivatives. Inclusion of the higher derivatives into equations of motion is one of the ways to treat with the problem of divergences in perturbative quantum gravity theory. If such terms are added to the Einstein gravity, then the resulting quantum theory is renormalizable [29]. Detailed discussion of the subject can be found in [30].

2.12.1 First-Order Trick

We start from a particular example of the action

$$S = \int d\tau L(q_1, \dot{q}_1, \ddot{q}_1), \tag{2.159}$$

where we use the notation $q_1 \equiv (q^1, q^2, \ldots, q^n)$ for the configuration-space vector. This leads to the equations of motion of the fourth order

$$\frac{\partial L}{\partial q_1} - \frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{q}_1} - \frac{d}{d\tau}\frac{\partial L}{\partial \ddot{q}_1}\right) = 0. \tag{2.160}$$

We suppose that the theory is nondegenerate, that is $\det \frac{\partial^2 L}{\partial \ddot{q}^a \partial \ddot{q}^b} \neq 0$.

The simplest way to obtain a Hamiltonian formulation is to apply the first-order trick of previous section to the $\ddot{q}_1$. We introduce the extended configuration space with the coordinates q_1, s, q_2, and write the action

$$S_1 = \int d\tau \left[L(q_1, \dot{q}_1, s) + q_2(\ddot{q}_1 - s) \right] \tag{2.161}$$

$$= \int d\tau \left[L(q_1, \dot{q}_1, s) - \dot{q}_2 \dot{q}_1 - q_2 s \right]. \tag{2.162}$$

In contrast to (2.159), this leads to the second-order equations of motion

$$\ddot{q}_1 = s, \qquad q_2 = \frac{\partial L}{\partial s}, \tag{2.163}$$

$$\frac{\partial L}{\partial q_1} - \frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{q}_1}\right) + \ddot{q}_2 = 0. \tag{2.164}$$

Using Eqs. (2.163) in (2.164) we reproduce the initial higher-order Eqs. (2.159), hence the two actions are equivalent.

Equation (2.162) represents one more example of a singular action. So, its Hamiltonian formulation is obtained according to the formalism that will be discussed in Chap. 8. For the later use, we present the final result for Hamiltonian equations of the variables q_1, q_2, p_1, p_2

$$\dot{q}_1 = -p_2, \qquad \dot{p}_1 = \frac{\partial L}{\partial q_1},$$
$$\dot{q}_2 = -p_1 - \frac{\partial L}{\partial p_2}, \qquad \dot{p}_2 = -s, \tag{2.165}$$

where $L = L(q_1, -p_2, s)$. The reader can verify that they imply (2.163) and (2.164).

We exclude the variable s from these equations. According to the rank condition $\det \frac{\partial^2 L}{\partial \ddot{q}^a \partial \ddot{q}^b} \neq 0$, the second equation from (2.163), $q_2 = \frac{\partial L(q_1, -p_2, s)}{\partial s}$, can be resolved with respect to s, $s = s(q_1, q_2, p_2)$. Substituting the function $s(q_1, q_2, p_2)$ into Eqs. (2.165), they read

$$\dot{q}_1 = -p_2, \qquad \dot{p}_1 = \left.\frac{\partial L}{\partial q_1}\right|_s = \frac{\partial L}{\partial q_1} - q_2 \frac{\partial s}{\partial q_1},$$
$$\dot{q}_2 = -p_1 - \left.\frac{\partial L}{\partial p_2}\right|_s = -p_1 - \frac{\partial L}{\partial p_2} + q_2 \frac{\partial s}{\partial p_2}, \qquad \dot{p}_2 = -s, \tag{2.166}$$

where on the r.h.s. we have $L \equiv L(q_1, -p_2, s(q_1, q_2, p_2))$. They follow from the Hamiltonian:

$$H = -p_1 p_2 - L(q_1, -p_2, s(q_1, q_2, p_2)) + q_2 s(q_1, q_2, p_2). \tag{2.167}$$

2.12.2 Ostrogradsky Method

Now consider the Hamiltonization of a theory with an action that depends on time derivatives up to N-th order. The procedure has been developed by Ostrogradsky [31]. Consider the action

$$S = \int d\tau L(q_1, \dot{q}_1, \ddot{q}_1, \dots, \overset{(N)}{q_1}), \qquad q_1 \equiv (q^1, q^2, \dots, q^n). \tag{2.168}$$

Disregarding a total derivative, variation of the action reads

$$\delta S = \int d\tau \sum_{i=0}^{N} \frac{\partial L}{\partial \overset{(i)}{q_1}} \delta \overset{(i)}{q_1} = \int d\tau \left(\sum_{i=0}^{N} (-1)^i \frac{d^i}{d\tau^i} \frac{\partial L}{\partial \overset{(i)}{q_1}} \right) \delta q_1, \tag{2.169}$$

so the Lagrangian equations are

$$\frac{\partial L}{\partial q_1} - \frac{d}{d\tau} \left(\frac{\partial L}{\partial \dot{q}_1} - \frac{d}{d\tau} \frac{\partial L}{\partial \ddot{q}_1} + \dots + (-1)^{N-1} \frac{d^{N-1}}{d\tau^{N-1}} \frac{\partial L}{\partial \overset{(N)}{q_1}} \right) = 0. \tag{2.170}$$

Computing derivatives with respect to τ we conclude that the equations have the following structure

$$L_{ab} \overset{(2N)}{q_1^b} = K_a(q_1, \dot{q}_1, \dots, \overset{(2N-1)}{q_1}), \qquad L_{ab} \equiv \frac{\partial^2 L}{\partial \overset{(N)}{q_1^a} \partial \overset{(N)}{q_1^b}}. \tag{2.171}$$

Hence the Lagrangian with N-th order derivatives implies $2N$-th order equations. The theory is called non degenerate if

$$\det L_{ab} \neq 0. \tag{2.172}$$

In this case the equations can be written in the normal form, with higher derivatives separated on l.h.s. of the equations. Specification of the position q_1 as well as its $2N - 1$ derivatives at some instant implies unique solution to the Cauchy problem.

To present the system in the Hamiltonian form, we introduce $2N \times n$ dimensional phase space spanned by the coordinates $q_i, p_i, i = 1, 2, \dots, N$. Let us specify their dynamics as follows. The variable q_1 obeys the Eq. (2.170), while other variables accompany its evolution according the equations

$$q_i = \overset{(i-1)}{q_1}, \qquad\qquad i = 2, 3, \dots, N,$$
$$p_i = \sum_{j=i}^{N} (-1)^{j-i} \frac{d^{j-i}}{d\tau^{j-i}} \frac{\partial L}{\partial \overset{(j)}{q_1}}, \qquad i = 1, 2, \dots, N. \tag{2.173}$$

In particular, expression for the momenta p_N reads

$$p_N = \frac{\partial L(q_1, \overset{(i)}{q_1})}{\partial \overset{(N)}{q_1}} = \frac{\partial L(q_1, q_2, \dots, q_N, \dot{q}_N)}{\partial \dot{q}_N}. \tag{2.174}$$

According to the condition (2.172), it can be resolved algebraically with respect to $\dot{q}_N$. Let us denote the solution by s_N

$$\dot{q}_N = s_N(q_1, q_2, \dots, q_N, p_N). \tag{2.175}$$

Using the solution, we rewrite the system (2.170) and (2.173) in the first order normal form

$$\dot{q}_{i-1} = q_i, \tag{2.176}$$
$$\dot{q}_N = s_N(q_1, q_2, \dots, q_N, p_N), \tag{2.177}$$
$$\dot{p}_1 = \left.\frac{\partial L}{\partial q_1}\right|_{s_N} = -\frac{\partial}{\partial q_i}\left(p_N s_N - L(q_i, s_N)\right), \tag{2.178}$$
$$\dot{p}_i = -p_{i-1} + \left.\frac{\partial L}{\partial q_i}\right|_{s_N}$$
$$\equiv -\frac{\partial}{\partial q_i}\left(p_{i-1} q_i + p_N s_N - L(q_i, s_N)\right). \tag{2.179}$$

Here $i = 2, 3, \dots, N$. At last, introducing the Hamiltonian

$$H(q_i, p_j) = p_1 q_2 + p_2 q_3 + \dots + p_{N-1} q_N + p_N s_N - L(q_i, s_N), \tag{2.180}$$

the system (2.176), (2.177), (2.178), and (2.179) acquires the Hamiltonian form

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \equiv \{q_i, H\}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} \equiv \{p_i, H\}. \tag{2.181}$$

The Poisson brackets are defined by

$$\{q_i^b, p_{jb}\} = \delta^i_j \delta_{ab}. \tag{2.182}$$

Equations (2.181) follow from the Hamiltonian action functional

$$S_H = \int d\tau (p_i \dot{q}^i - H). \tag{2.183}$$

In resume, for an N-th order Lagrangian, the Hamiltonian formulation implies introducing $2N \times n$ dimensional phase space with the Poisson brackets (2.182). The working recipe for construction the corresponding Hamiltonian can be formulated as follows. Define the momenta p_N according to the Eq. (2.174) and resolve it with respect to $\overset{(N)}{q_1}$. Then the Hamiltonian is

$$H(q_i, p_j) = \sum_{i=1}^{N} p_i \overset{(i)}{q_1} - L(q_1, \overset{(i)}{q_1}), \qquad i = 1, 2, \ldots, N, \tag{2.184}$$

where one substitutes q_{i+1} instead of $\overset{(i)}{q_1}$, $i = 1, 2, \ldots, N-1$, and s_N of Eq. (2.175) instead of $\overset{N}{q_1}$.

In conclusion, we point out that Eqs. (2.176), (2.177) can not be resolved with respect to the momenta, that is the Ostrogradsky equations (2.181) can not be obtained from a Lagrangian (without higher derivatives). To avoid the difficulty, one needs to make an appropriate canonical transformation. For instance, for the case of the Lagrangian $L(q_1, \dot{q}_1, \ddot{q}_1)$ it is sufficient to make the transformation $q_2 \to -p_2$, $p_2 \to q_2$. After that, the Hamiltonian (2.180) and the Ostrogradsky equations (2.181) turn out into Eqs. (2.167) and (2.166).

Chapter 3
Canonical Transformations of Two-Dimensional Phase Space

It is common in textbooks on classical mechanics to discuss canonical transformations on the basis of the integral form of the canonicity conditions and a theory of integral invariants [1, 12, 14]. We prefer to deduce all the properties of canonical transformations by direct analysis of the canonicity conditions given by Eqs. (2.100) and (2.104). We start the discussion from the case of two-dimensional phase space $z^i = (q, p)$, where all the basic properties of canonical transformations can be obtained by elementary calculations. For convenience, we have made the subject matter of the next chapter independent from this one, so the reader can o mit this and continue from the next chapter.

3.1 Time-Independent Canonical Transformations

3.1.1 Time-Independent Canonical Transformations and Symplectic Matrix

It is worth noting that time-independent canonical transformations are an important tool to analyze the structure of a general singular theory.

Discarding the dependence on τ in Eq. (2.91) we arrive at the time-independent[1] coordinate transformation $z'^i = z'^i(z^j)$ or, $q' = q'(q, p)$, $p' = p'(q, p)$. In terms of the new coordinates, the Hamiltonian equations acquire the form (see (2.96))

$$\dot{z}'^k = \{z'^k, z'^l\}_z\Big|_{z(z')} \frac{\partial H(z(z'))}{\partial z'^l}, \tag{3.1}$$

while the definition of canonical transformation (2.97) implies (see (2.98))

$$\{z'^k, z'^l\}_z\Big|_{z(z')} \frac{\partial H(z(z'))}{\partial z'^l} = \omega^{kl} \frac{\partial \tilde{H}(z')}{\partial z'^l}, \quad \text{any} \quad H, \quad \text{some} \quad \tilde{H}. \tag{3.2}$$

---

[1] Sometimes these are called *contact transformations*.

A. Deriglazov, *Classical Mechanics*, DOI 10.1007/978-3-642-14037-2_3,

As the first basic result, we show that the group of canonical transformations can be identified with a group of coordinate transformations, leaving invariant (disregarding the multiplicative constant) the symplectic matrix ω^{ij}. More exactly, one has:

Assertion Transformation $z'^i = z'^i(z^j)$ is canonical if and only if

$$\frac{\partial z'^k}{\partial z^i}\omega^{ij}\frac{\partial z'^l}{\partial z^j} = c\omega^{kl}, \quad or \quad \{z'^k, z'^l\}_z = c\omega^{kl}, \quad c = const. \tag{3.3}$$

Proof Let the transformation be canonical, hence it obeys the system (3.2). In more detail, one has two equations

$$\{q', p'\}\big| \frac{\partial H}{\partial p'} = \frac{\partial \tilde{H}}{\partial p'}, \quad -\{q', p'\}\big| \frac{\partial H}{\partial q'} = -\frac{\partial \tilde{H}}{\partial q'}. \tag{3.4}$$

Computing the derivative of the first (second) equation with respect to q' (p') correspondingly, and adding the resulting expressions, one obtains

$$\frac{\partial}{\partial q'}\left(\{q', p'\}|\right)\frac{\partial H}{\partial p'} - \frac{\partial}{\partial q'}\left(\{q', p'\}|\right)\frac{\partial H}{\partial q'} = 0. \tag{3.5}$$

Since this is true for any H, one concludes $\frac{\partial}{\partial q'}\{q', p'\}| = 0$, $\frac{\partial}{\partial q'}\{q', p'\}| = 0$, which in turn implies $\{q', p'\} = c =$const. The remaining Poisson brackets are $\{q', q'\} = 0$, $\{p', p'\} = 0$. Combining the brackets, one has the desired result: $\{z'^k, z'^l\}_z = c\omega^{kl}$. Besides, substitution of Eq. (3.3) into Eq. (3.2) gives a relationship between the original and the transformed Hamiltonians

$$\tilde{H}(z') = cH(z(z')). \tag{3.6}$$

The inverse affirmation is evident: Eq. (3.3) implies (3.2) with $\tilde{H}$ given by Eq. (3.6).
Comments 1. Equation (3.3) can be rewritten in an equivalent form

$$\left.\frac{\partial z^i}{\partial z'^j}\right|_{z(z')} = c^{-1}\omega^{ik}\frac{\partial z'^l}{\partial z^k}\omega_{lj}, \tag{3.7}$$

and shows how an inverse of the matrix $\partial_k z'^l$ can be computed.

2. Let us define a Poisson bracket in relation to z' variables as follows: $\{z'^i, z'^j\}_{z'} = \omega^{ij}$. For the case of univalent canonical transformation (c=1), Eq. (3.3) can be written as

$$\{z'^i(z), z'^j(z)\}_z = \{z'^i, z'^j\}_{z'}. \tag{3.8}$$

In accordance with this, for any two phase-space functions one obtains

$$\{A(z), B(z)\}_z|_{z(z')} = \{A(z(z')), B(z(z'))\}_{z'}. \tag{3.9}$$

These expressions mean that univalent canonical transformation and computation of the Poisson bracket are commuting operations. For this reason, Eqs. (3.8) and (3.9) are sometimes referred to as a property of *invariance of the Poisson bracket* under univalent canonical transformation.

3.1.2 Generating Function

Let $q \to q' = q'(q, p),\ p \to p' = p'(q, p)$ be canonical transformation. Suppose that the second equation can be resolved with respect to p: $p' = p'(q, p) \Rightarrow p = p(q, p')$. Transformations with this property are called *free canonical transformations*. Using the latter equation, one can represent the variables q', p in terms of q, p':

$$q' = q'(q, p(q, p')) \equiv q'(q, p'), \qquad p = p(q, p'). \tag{3.10}$$

By construction, these expressions can in turn be solved with respect to q', p'. So, one can deal with a canonical transformation in the form (3.10), where q, p' are considered as independent variables, instead of its original form, with q, p being independent. The identities (2.95) acquire the form

$$\begin{gathered}
\left.\frac{\partial q(q', p')}{\partial q'}\right|_{q'(q,p')} \frac{\partial q'(q, p')}{\partial q} = 1, \\
\left.\frac{\partial q(q', p')}{\partial q'}\right|_{q'(q,p')} \frac{\partial q'(q, p')}{\partial p'} = -\frac{\partial q(q, p')}{\partial p'}.
\end{gathered} \tag{3.11}$$

In this section we demonstrate that there is a simple way to construct a free canonical transformation starting from any given function $S(q, p')$, see Eq. (3.15) below.

Assertion For a given transformation $z^i \to z'^i(z)$, the following conditions are equivalent:

(a) the transformation is canonical:

$$\{z^i, z^j\}_{z'} = c^{-1}\omega^{ij}, \quad c = \text{const}, \tag{3.12}$$

(b) there is a function $F(q', p')$ such that

$$cp\frac{\partial q}{\partial q'} - p' = \frac{\partial F}{\partial q'}, \qquad cp\frac{\partial q}{\partial p'} = \frac{\partial F}{\partial p'}, \tag{3.13}$$

where $q = q(q', p')$, $p = p(q', p')$.

Proof Let the transformation be canonical. The system (3.12) contains only one nontrivial equation: $\{q, p\}_{z'} = c^{-1}$, or $\frac{\partial q}{\partial q'}\frac{\partial p}{\partial p'} - \frac{\partial q}{\partial p'}\frac{\partial p}{\partial q'} = c^{-1}$, which can otherwise be rewritten as

$$\frac{\partial}{\partial p'}\left(cp\frac{\partial q}{\partial q'} - p'\right) - \frac{\partial}{\partial q'}\left(cp\frac{\partial q}{\partial p'}\right) = 0. \tag{3.14}$$

This means that a vector field with the components $F_1 = cp\frac{\partial q}{\partial q'} - p'$, $F_2 = cp\frac{\partial q}{\partial p'}$ is curl-free, $\partial_1 F_2 - \partial_2 F_1 = 0$. Then there is the potential $F(q', p')$ which obeys Eq. (3.13). The inverse affirmation is also true: differentiating Eq.(3.13) with respect to q' and p' and adding the resulting expressions, one obtains $\{q, p\}_{z'} = c^{-1}$.

Assertion Let $z^i \to z'^i(z)$ be a free canonical transformation, hence it can be presented in the form (3.10). There is a function $S(q, p')$ such that $\frac{\partial^2 S}{\partial q \partial p'} \neq 0$, and

$$q'(q, p') = \frac{\partial S}{\partial p'}, \qquad cp(q, p') = \frac{\partial S}{\partial q}. \tag{3.15}$$

The function S is called the *generating function of the canonical transformation*.

Proof The following function

$$S(q, p') = F(q'(q, p'), p') + p'q'(q, p') \tag{3.16}$$

obeys the desired conditions, as can be demonstrated by direct computations with use of Eqs. (3.13) and (3.11). Notice that S is defined on (q, p') space.

Thus we have seen that with a given canonical transformation one can associate the corresponding generating function. It is natural to ask whether a given function $S(q, p')$ defines a canonical transformation. This seems to be true. In particular, the Assertion above can be inverted in the following sense:

Assertion Let $S(q, p')$ be a function with $\frac{\partial^2 S}{\partial q \partial p'} \neq 0$. Let us solve the algebraic equations $q' = \frac{\partial S(q,p')}{\partial p'}$, $cp = \frac{\partial S(q,p')}{\partial q}$ with respect to q, p (one solves the first equation for q and substitutes the result into the second one). Then the solution

$$q = q(q', p'), \qquad p = c^{-1}\left.\frac{\partial S}{\partial q}\right|_{q(q',p')} \equiv p(q', p'), \tag{3.17}$$

is the free canonical transformation.

Proof It is sufficient to demonstrate that $\{q, p\}_{z'} = c^{-1}$; see Eq. (3.12). Let us denote $q' - \frac{\partial S(q,p')}{\partial p'} \equiv G(q', q, p')$. From the identity $G(q', q(q', p'), p') \equiv 0$, one finds the consequences

$$\frac{\partial q}{\partial q'} = \frac{1}{S_{qp'}}, \qquad \frac{\partial q}{\partial p'} = -\frac{S_{p'p'}}{S_{qp'}}, \tag{3.18}$$

where it was denoted $S_{qq} = \frac{\partial^2 S}{\partial^2 q}\Big|_{q(q',p')}$, and so on. Further, the last equation from (3.17) implies

$$\frac{\partial p}{\partial q'} = c^{-1} S_{qq} \frac{\partial q}{\partial q'} = c^{-1} \frac{S_{qq}}{S_{qp'}},$$
$$\frac{\partial p}{\partial p'} = c^{-1} \left(S_{qp'} + S_{qq} \frac{\partial q}{\partial p'} \right) = c^{-1} \left(S_{qp'} - \frac{S_{qq} S_{p'p'}}{S_{qp'}} \right). \tag{3.19}$$

These expressions allows one to compute the desired Poisson bracket, with the result being $\{q, p\}_{z'} = c^{-1}$.

Exercise
Do this calculation.

3.2 Time-Dependent Canonical Transformations

Here we repeat the analysis of Sect. 3.1 for the case of time-dependent transformations in two-dimensional phase space. As compared with the previous case, the only difference in the final results is, in fact, a non-trivial form of a transformed Hamiltonian, see Eq. (3.29) below. Owing to this property, the time-dependent canonical transformations can be used for the simplification of Hamiltonian equations, see below.

3.2.1 Canonical Transformations and Symplectic Matrix

For the case of time-dependent transformation $q' = q'(q, p, \tau)$, $p' = p'(q, p, \tau)$, the Hamiltonian equations in terms of q', p' acquire the form (2.96)), while the definition of canonical transformation (2.97) implies Eqs. (2.98), (2.99), (2.100), (2.101), (2.102), (2.103), and (2.104). As before, the set of canonical transformations can be identified with the set of coordinate transformations leaving invariant (disregarding a constant) the symplectic matrix ω^{ij}:

Assertion The transformation $z'^i = z'^i(z^b, \tau)$ is canonical if and only if:

$$\frac{\partial z'^k}{\partial z^i} \omega^{ij} \frac{\partial z'^l}{\partial z^j} = c\omega^{kl}, \quad \text{or} \quad \{z'^k, z'^l\}_z = c\omega^{kl}, \quad c = \text{const}. \tag{3.20}$$

Proof (A) Let the transformation be canonical, hence it obeys the system (2.98). Repeating analysis of section (3.1.1) one arrives at the system

$$\frac{\partial}{\partial q'}(\{q', p'\}|) = 0, \qquad \frac{\partial}{\partial p'}(\{q', p'\}|) = 0, \tag{3.21}$$

$$\frac{\partial}{\partial q'}\left(\left.\frac{\partial q'(z,\tau)}{\partial\tau}\right|_{z(z',\tau)}\right) + \frac{\partial}{\partial p'}\left(\left.\frac{\partial p'(z,\tau)}{\partial\tau}\right|_{z(z',\tau)}\right) = 0. \tag{3.22}$$

Equation (3.21) implies $\{q', p'\} = c(\tau)$, or

$$c(\tau) = \frac{\partial q'}{\partial q}\frac{\partial p'}{\partial p} - \frac{\partial p'}{\partial q}\frac{\partial q'}{\partial p}. \tag{3.23}$$

Equation (3.22) states that a vector field with the components $N_1 = \frac{\partial p'}{\partial\tau}|$, $N_2 = -\frac{\partial q'}{\partial\tau}|$ is curl-free, so there is the potential $N(q', p', \tau)$

$$\frac{\partial p'}{\partial\tau} = \left.\frac{\partial N}{\partial q'}\right|, \qquad -\frac{\partial q'}{\partial\tau} = \left.\frac{\partial N}{\partial p'}\right|. \tag{3.24}$$

Let us demonstrate that this implies $\frac{dc}{d\tau} = 0$, that is, $c =$const. Differentiating Eq. (3.24) one obtains

$$-\frac{\partial^2 q'}{\partial z^j\partial\tau} = \left.\frac{\partial^2 N}{\partial z'^i\partial p'}\right|\frac{\partial z'^i}{\partial z^j}, \qquad \frac{\partial^2 p'}{\partial z^j\partial\tau} = \left.\frac{\partial^2 N}{\partial z'^i\partial q'}\right|\frac{\partial z'^i}{\partial z^j}. \tag{3.25}$$

Therefore the derivative of Eq. (3.23) with respect to τ turns out to be zero, as a consequence of Eq. (3.25).

B) Suppose that the transformation $z'^i = z'^i(z, \tau)$ obeys Eq. (3.20). First, note that the Assertion on page 121 is true for the present case of time-dependent transformations as well (since in the corresponding proof only partial derivatives with respect to z'^a were used). Equation (3.20) thus implies Eqs. (3.13), and differentiating the latter with respect to τ we obtain:

$$c\frac{\partial p}{\partial\tau}\frac{\partial q}{\partial z'^a} + cp\frac{\partial^2 q}{\partial\tau\partial z'^a} = \frac{\partial^2 F(z', \tau)}{\partial z'^a\partial\tau}. \tag{3.26}$$

Second, under condition (3.20), Hamiltonian equations for z' (2.96) acquire the form

$$\begin{aligned} \dot{z}'^i &= c\omega^{ij}\frac{\partial H(z(z',\tau))}{\partial z'^j} - c\omega^{ij}\frac{\partial z^l}{\partial z'^j}\omega_{lk}\frac{\partial z^k}{\partial\tau} = \\ & c\omega^{ij}\frac{\partial H(z(z',\tau))}{\partial z'^j} + \omega_{ij}\frac{\partial}{\partial z'^j}\left(\frac{\partial F}{\partial\tau} - cp\frac{\partial q}{\partial\tau}\right), \end{aligned} \tag{3.27}$$

where Eqs. (2.95) and (3.7) were used in the first line, and Eq. (3.26) was used in the transition from the first to the second line. Thus condition (3.20) implies the canonical form of the Hamilton equations

$$\dot{z}'^i = \omega^{ij}\frac{\partial}{\partial z'^j}\left(cH(z(z',\tau) - cp\frac{\partial q}{\partial \tau} + \frac{\partial F}{\partial \tau}\right), \tag{3.28}$$

which completes the proof.

Besides, comparing this result with Eq. (2.97), one obtains a relationship between the original and the transformed Hamiltonians:

Consequence Let $z^a \to z'^i = z'^i(z,\tau)$ be a canonical transformation. Then there is a function F such that

$$\tilde{H}(z',\tau) = cH(z(z',\tau)) - cp(z',\tau)\frac{\partial q(z',\tau)}{\partial \tau} + \frac{\partial F(z',\tau)}{\partial \tau}. \tag{3.29}$$

Transformation properties of the Hamiltonian action under canonical transformation will be discussed in Sect. 4.5.

Comment As compared to the time-independent canonical transformations, the transformed Hamiltonian now acquires some extra terms. It allows one to formulate the following problem: find the canonical transformation $z'^i = z'^i(z,\tau)$ that simplifies the Hamiltonian as much as possible, for example[2] $\tilde{H} = 0$. The desired canonical transformation can be found in some interesting cases by using the Hamilton–Jacobi method; see Sect. 4.7 below. In the new coordinates, the Hamiltonian equations would be trivial: $\dot{z}'^i = 0$, and can immediately be solved: $z'^i = C^i$. Further, solving the algebraic equations $z'^i(z,\tau) = C^i$ (where the functions $z'^i(z^j,\tau)$ are known from the canonical transformation), one obtains a general solution to the equations of motion in the initial parametrization: $z^i = z^i(\tau, C^j)$.

3.2.2 Generating Function

Assertion Let $q \to q' = q'(q,p,\tau),\ p \to p' = p'(q,p,\tau)$ be free canonical transformation, hence from these expressions one writes

$$q' = q'(q, p(q,p',\tau),\tau) \equiv q'(q,p',\tau), \qquad p = p(q,p',\tau). \tag{3.30}$$

Then

(a) there is a function $S(q,p',\tau)$, with $\frac{\partial S}{\partial q \partial p'} \neq 0$, such that

$$q'(q,p',\tau) = \frac{\partial S}{\partial p'}, \qquad cp(q,p',\tau) = \frac{\partial S}{\partial q}; \tag{3.31}$$

---

[2] note that this is not possible in the time-independent case: if $H(z)$ depends essentially on all the variables, then the same is true for $\tilde{H} = H(z(z'))$, see Eq. (3.6)).

(b) the transformed Hamiltonian (3.29) in terms of the variables q, p' acquires the form

$$\tilde{H}(z', \tau)\Big|_{q'(q,p',\tau)} = cH(q, p(q, p', \tau)) + \frac{\partial S(q, p', \tau)}{\partial \tau}. \tag{3.32}$$

Proof (a) The proof is similar to that given for Eq. (3.15), since only partial derivatives with respect to q, p were used there. (b) To substitute $q'(q, p', \tau)$ into Eq. (3.29) one needs two identities. First, from $q(q'(q, p', \tau)p', \tau) \equiv q$ it follows:

$$\frac{\partial q(q', p', \tau)}{\partial \tau}\Bigg|_{q'(q,p',\tau)} = -\frac{\partial q}{\partial q'}\Bigg|_{q'} \frac{\partial q'(q, p', \tau)}{\partial \tau}. \tag{3.33}$$

Second, from the expression:

$$\frac{\partial}{\partial \tau} F(q'(q, p', \tau), p', \tau) = \frac{\partial F(z', \tau)}{\partial q'}\Bigg|_{q'(q,p',\tau)} \frac{\partial q'}{\partial \tau} + \frac{\partial F(z', \tau)}{\partial \tau}\Bigg|_{q'(q,p',\tau)} \tag{3.34}$$

one finds:

$$\begin{aligned} &\frac{\partial F(z', \tau)}{\partial \tau}\Bigg|_{q'(q,p',\tau)} = \\ &\left(-cp\, \frac{\partial q(z', \tau)}{\partial q'}\Bigg|_{q'(q,p',\tau)} + p'\right) \frac{\partial q'(q, p', \tau)}{\partial \tau} \\ &\qquad + \frac{\partial}{\partial \tau} F(q'(q, p', \tau), p', \tau), \end{aligned} \tag{3.35}$$

where Eq. (3.13) was used. Equation (3.32) follows from Eq. (3.29) by using these equalities as well as the manifest form of S, see Eq. (3.16).

As before, this result can be inverted in the following sense:

Assertion Let $S(q, p', \tau)$ be a function with $\frac{\partial^2 S}{\partial q \partial p'} \neq 0$, for any τ. Let us solve the algebraic equations $q' = \frac{\partial S(q,p',\tau)}{\partial p'}$, $cp = \frac{\partial S(q,p',\tau)}{\partial q}$ in relation to q, p. Then the solution

$$q = q(q', p', \tau), \qquad p = c^{-1} \frac{\partial S}{\partial q}\Bigg|_{q(q',p',\tau)} \equiv p(q', p', \tau), \tag{3.36}$$

is the free canonical transformation.

The proof is the same as before (see page 122), since only partial derivatives with respect to q', p' were used there.

Chapter 4
Properties of Canonical Transformations

As we have seen in Sect. 2.7, the canonical form of Hamiltonian equations is not preserved by general phase-space transformations. Those that leave the form of the equations unaltered were called canonical transformations. In this chapter, we discuss their properties for the case of phase space of an arbitrary dimension.

We start from the demonstration that the equation $\{z'^i, z'^j\}_z = \{z^i, z^j\}_z$, which represents the invariance of the Poisson bracket under a transformation $z \to z'(z, \tau)$, can be rewritten in the following *equivalent* form: $\partial'^i E^j(z') - \partial'^j E^i(z') = 0$. This means that E^i are components of a conservative vector field, and therefore there is a potential E, such that $E^i = \partial'^i E$. Thus, the invariance of the Poisson bracket is equivalent to the statement, that the transition functions $z'^i(z, \tau)$ can be used to construct a conservative field. In turn, this allows us to prove the following two facts. First, canonical transformations are the only ones that leave the Poisson bracket invariant (up to a constant). This gives a simple rule for checking whether a given transformation is a canonical one. Second, with any canonical transformation[1] can be associated a generating function. Its partial derivatives give the transition functions of the transformation. The generating function can be obtained from the above-mentioned potential according to a simple rule. Among other things, it gives a simple way to construct examples of canonical transformations.

Further, it will be seen that the Hamiltonian has a rather non-trivial transformation law under time-dependent canonical transformation (it does not transform as a scalar function). This implies the possibility of looking for the transformation which trivializes the Hamiltonian function (and thus trivializes the equations of motion) in the new coordinate system. By this means, the problem of finding a general solution to Hamiltonian equations can be replaced by the problem of finding a generating function of the transformation. The generating function obeys the so-called Hamilton–Jacobi equation, which can be solved in many interesting cases.

[1] Below, we discuss only a free canonical transformation. For an arbitrary canonical transformation, the situation is similar, see [14].

A. Deriglazov, *Classical Mechanics*, DOI 10.1007/978-3-642-14037-2_4,

4.1 Invariance of the Poisson Bracket (Symplectic Matrix)

General phase-space transformation alters a form of the Hamiltonian equations according to (2.89)

$$\dot{z}'^k = \left(\frac{\partial z'^k}{\partial z^i} \omega^{ij} \frac{\partial z'^l}{\partial z^j} \right) \bigg|_{z(z',\tau)} \frac{\partial H(z(z',\tau))}{\partial z'^l} + \frac{\partial z'^k}{\partial \tau} \bigg|_{z(z',\tau)} . \tag{4.1}$$

From this equation we expect that form-invariance may be closely related with the symmetry properties of the symplectic matrix. In fact, at least for time-independent transformations, invariance of ω: $\partial_i z'^k \omega^{ij} \partial_j z'^l = \omega^{kl}$ implies form-invariance of Hamiltonian equations. We can also speak of an invariance of the Poisson bracket, since the above equation can also be written as $\{z'^k, z'^l\}_z = \{z^k, z^l\}_z$; see (2.96). In this section we establish an exact relationship: the set of transformations, which preserves the canonical form of Hamiltonian equations, coincides with a set which leaves the Poisson bracket invariant (up to a constant).

Assertion Transformation $z^i \to z'^i = z'^i(z^j, \tau)$ is canonical if and only if it obeys:

$$\frac{\partial z'^k}{\partial z^i} \omega^{ij} \frac{\partial z'^l}{\partial z^j} = c\omega^{kl}, \quad \text{or} \quad \{z'^k, z'^l\}_z = c\omega^{kl}, \quad c = \text{const}. \tag{4.2}$$

Comment Denoting the Jacobi matrix $\frac{\partial z'^k}{\partial z^i}$ of the transformation as $J^k{}_j$, Eq. (4.2) is of the form $J\omega J^T = c\omega$. Taking the determinant of both sides for the case of univalent canonical transformation, we have

$$\det J = \pm 1, \quad \text{for all} \quad z, \tau. \tag{4.3}$$

Proof (A) Let the transformation be canonical, hence it obeys the system (2.104), which we repeat here:

$$\partial'^i \left(\partial_\tau z'^j(z,\tau) \Big|_{z(z',\tau)} \right) - (i \leftrightarrow j) = 0, \tag{4.4}$$

$$\partial'^i W^{jl} - \partial'^j W^{il} = 0, \tag{4.5}$$

$$W^{ik}\omega^{jl} - W^{jk}\omega^{il} + W^{il}\omega^{jk} - W^{jl}\omega^{ik} = 0, \tag{4.6}$$

where:

$$W^{ij} \equiv \{z'^i, z'^j\}_z \Big|_{z(z',\tau)} . \tag{4.7}$$

We need to show that $W^{ij} = c\omega^{ij}$.

Equation (4.4) states that a vector field with the components $\partial_\tau z'^k|$ is curl-free. Therefore it can be presented locally as the divergence of a function N: $\partial_\tau z'^k| =$

$\partial'^k N(z', \tau)$ or, in other words, $\partial_\tau z'^k = \omega^{kl} \frac{\partial N}{\partial z'^l}|$. Below we will use a derivative of this expression

$$\frac{\partial^2 z'^k(z, \tau)}{\partial z^i \partial \tau} = \omega^{kl} \frac{\partial^2 N(z', \tau)}{\partial z'^n \partial z'^l}\bigg|_{z'(z,\tau)} \frac{\partial z'^n}{\partial z^i}. \tag{4.8}$$

Similarly, Eq. (4.5) implies $W^{jl} = 2\partial'^j G^l$ for any fixed l. Since W is antisymmetric, this equation implies the following restriction on G^l: $\partial'^j G^l = -\partial'^l G^j$ for any j, l. It allows us to rewrite the expression for W in an explicitly antisymmetric form: $W^{jl} = \partial'^j G^l - \partial'^l G^j$. The last two equations can be used to show that W does not depend on z'^i. Indeed, substitution of the last equation back into (4.5) gives the expression $-\partial'^i \partial'^l G^j + \partial'^j \partial'^l G^i = \partial'^l W^{ji} = 0$, for any i, j, l. Hence W does not depend on z' and can be a function only of τ: $W^{ij} = W^{ij}(\tau)$.

Now, contraction of Eq. (4.6) with ω_{li} gives immediately

$$W^{jk} = c(\tau)\omega^{jk}, \tag{4.9}$$

where $c(\tau) \equiv \frac{1}{n} W^{il}(\tau)\omega_{li} = \frac{1}{n} \frac{\partial z'^k}{\partial z^i} \omega^{ij} \frac{\partial z'^l}{\partial z^j}\Big|_{z(z',\tau)} \omega_{lk}$ or, equivalently, $c(\tau) = \frac{1}{n} \frac{\partial z'^k}{\partial z^i} \omega^{ij} \frac{\partial z'^l}{\partial z^j} \omega_{lk}$. The derivative of this expression with respect to τ gives, by using Eq.(4.8)

$$\begin{aligned} \frac{dc}{d\tau} = \frac{2}{n} \frac{\partial^2 z'^k}{\partial \tau \partial z^i} \omega^{ij} \frac{\partial z'^l}{\partial z^i} \omega_{lk} = \\ \frac{2}{n} \frac{\partial^2 N}{\partial z'^n \partial z'^l}\bigg| \frac{\partial z'^n}{\partial z^i} \omega^{ij} \frac{\partial z'^l}{\partial z^j}. \end{aligned} \tag{4.10}$$

This expression vanishes since it is symmetric on n, l and antisymmetric on i, j. Thus the coefficient c in Eq. (4.9) is a constant, which completes the first part of the proof.

(B) Supposing that $z'^i(z^j, \tau)$ obeys Eq. (4.2), the latter can be rewritten in an equivalent form

$$\frac{\partial z'^i}{\partial z^l}\bigg|_{z(z',\tau)} = c\omega^{ij} \frac{\partial z^k}{\partial z'^j} \omega_{kl}. \tag{4.11}$$

By using Eqs. (4.2), (2.95) and (4.11), Hamiltonian equations for the variables z' (2.96) can be written as

$$\dot{z}'^k = c\omega^{kl} \frac{\partial H(z(z', \tau))}{\partial z'^l} - c\omega^{kl} \frac{\partial z^i}{\partial z'^l} \omega_{ij} \frac{\partial z^j(z', \tau)}{\partial \tau}. \tag{4.12}$$

To confirm that they have the canonical form, it is sufficient to show that the last term can be written as $\omega^{kl} \frac{\partial}{\partial z'^l}(\ldots)$. We need the following:

Lemma 1 *Let $z^i \to z'^i = z'^i(z^j, \tau)$ be a phase-space transformation. Then the following conditions are equivalent:*

(a) the symplectic form is invariant

$$\frac{\partial z'^k}{\partial z^i}\omega^{ij}\frac{\partial z'^l}{\partial z^j} = c\omega^{kl}, \quad c = \text{const}; \tag{4.13}$$

(b) there is a function $E(z', \tau)$ such that

$$cz^j(z', \tau)\omega_{ji}\frac{\partial z^i(z', \tau)}{\partial z'^l} + \omega_{lj}z'^j = 2\frac{\partial E(z', \tau)}{\partial z'^l}. \tag{4.14}$$

Proof Suppose (a) is true. By using Eq. (4.11), it can be rewritten in an equivalent form[2] (recall that $\frac{\partial}{\partial z'_k} = \omega^{kl}\frac{\partial}{\partial z'^l}$):

$$c\frac{\partial z^j}{\partial z'_k}\omega_{jn}\frac{\partial z^n}{\partial z'_l} = -\omega^{kl}. \tag{4.15}$$

Owing to the antisymmetry of k, l this can be further rewritten

$$c\frac{\partial z^j}{\partial z'_k}\omega_{jn}\frac{\partial z^n}{\partial z'_l} - c\frac{\partial z^j}{\partial z'_l}\omega_{jn}\frac{\partial z^n}{\partial z'_k} = -2\omega^{kl}, \tag{4.16}$$

or

$$\frac{\partial}{\partial z'_k}\left(cz^j\omega_{jn}\frac{\partial z^n}{\partial z'_l} + z'^l\right) - \frac{\partial}{\partial z'_l}\left(cz^j\omega_{jn}\frac{\partial z^n}{\partial z'_k} + z'^k\right) = 0. \tag{4.17}$$

That is, the condition (4.13) of invariance of the Poisson bracket is rewritten as the conservativity condition of a vector field. This implies:

$$cz^j\omega_{jn}\frac{\partial z^n}{\partial z'_l} + z'^l = 2\frac{\partial E}{\partial z'_l} \equiv 2\omega^{ln}\frac{\partial E}{\partial z'^n}, \tag{4.18}$$

or

$$cz^j\omega_{jn}\frac{\partial z^n}{\partial z'^i} + \omega_{in}z'^n = 2\frac{\partial E}{\partial z'^i}, \tag{4.19}$$

as has been stated.

[2] The left-hand side of this expression is known as a *Lagrange bracket*.

Now suppose that (b) is true. This implies (4.17) and, since the computation can be inverted, we obtain Eq. (4.15). An equivalent form of this expression is $c\frac{\partial z^i}{\partial z'^j} = \omega^{ik}\frac{\partial z'^l}{\partial z^k}\omega_{lj}$. Using this expression in Eq. (4.15), we have the desired result (4.13).

To analyze the last term in Eq. (4.12), we need the derivative of Eq. (4.14) with respect to τ. We obtain:

$$\begin{aligned} 2\frac{\partial}{\partial z'^l}\frac{\partial E}{\partial \tau} = c\frac{\partial z^j}{\partial \tau}\omega_{ji}\frac{\partial z^i}{\partial z'^l} + cz^j\omega_{ji}\frac{\partial^2 z^i}{\partial \tau \partial z'^l} = \\ -c\frac{\partial z^i}{\partial z'^l}\omega_{ij}\frac{\partial z^j}{\partial \tau} + \frac{\partial}{\partial z'^l}\left(cz^j\omega_{ji}\frac{\partial z^i}{\partial \tau}\right) - c\frac{\partial z^j}{\partial z'^l}\omega_{ji}\frac{\partial z^i}{\partial \tau} = \\ -2c\frac{\partial z^i}{\partial z'^l}\omega_{ij}\frac{\partial z^j}{\partial \tau} + \frac{\partial}{\partial z'^l}\left(cz^j\omega_{ji}\frac{\partial z^i}{\partial \tau}\right), \end{aligned} \tag{4.20}$$

or

$$-c\frac{\partial z^i}{\partial z'^l}\omega_{ij}\frac{\partial z^j}{\partial \tau} = \frac{\partial}{\partial z'^l}\left(\frac{\partial E}{\partial \tau} - \frac{c}{2}z^i\omega_{ij}\frac{\partial z^j}{\partial \tau}\right). \tag{4.21}$$

Using this result in Eq. (4.12), it can be written in the canonical form

$$\dot{z}'^k = \omega^{kl}\frac{\partial}{\partial z'^l}\left(cH(z(z',\tau)) - \frac{c}{2}z^i\omega^{ij}\frac{\partial z^j}{\partial \tau} + \frac{\partial E(z',\tau)}{\partial \tau}\right), \tag{4.22}$$

which completes the proof.

From Eqs. (4.17) and (4.18) it follows that one can add to E an arbitrary function $e(\tau)$. Note that this does not contribute to the equations of motion (4.21).

The result obtained means that the invariance (4.2) of the Poisson bracket can be taken as a definition for the canonical transformation.

Comparing our result (4.22) with Eq. (2.97), we have an exact relationship between the original and the transformed Hamiltonians.

Consequence Let $z^i \to z'^i = z'^i(z,\tau)$ be a canonical transformation. Then there is a function E (which obeys Eq. (4.14)) such that:

$$\tilde{H}(z',\tau) = cH(z(z',\tau)) - \frac{c}{2}z^i\omega_{ij}\frac{\partial z^j}{\partial \tau} + \frac{\partial E(z',\tau)}{\partial \tau}. \tag{4.23}$$

Note that for a univalent time-independent canonical transformation the Hamiltonian transforms as a scalar function:

$$\tilde{H}(z') = H(z(z')). \tag{4.24}$$

Hence, if $H(z)$ represents the energy of a system, the same is true for $\tilde{H}(z')$.

Exercise
Show that transition to polar coordinates on two-dimensional phase space is not a canonical transformation, but slightly modified polar coordinates

$$q = \sqrt{2P}\cos S, \qquad p = \sqrt{2P}\sin S, \tag{4.25}$$

represent a univalent canonical transformation, $\{P, S\}_{q,p} = 1$.

As an example, let us consider the Shrödinger equation (2.129). We represent the wave function $\Psi = \varphi + ip$ in terms of the probability density P and the phase S as follows:

$$\varphi + ip = \sqrt{P} e^{\frac{i}{\hbar} S}. \tag{4.26}$$

This is the canonical transformation of valence $c = 2\hbar$, $\{P, S\}_{q,p} = 2\hbar$. Hence the Schrödinger system (2.130) and (2.131) acquires the form $\dot{P} = \{P, \tilde{H}\}_{P,S}$, $\dot{S} = \{S, \tilde{H}\}_{P,S}$, where:

$$\begin{aligned}\tilde{H} &= 2\hbar H(\varphi(P, S), p(P, S)) \\ &= \int d^3x\, P\left(\frac{1}{2m}\vec{\nabla}S\vec{\nabla}S + V + \frac{\hbar^2}{8m}\frac{1}{P^2}\vec{\nabla}P\vec{\nabla}P\right).\end{aligned} \tag{4.27}$$

The system reads:

$$\begin{aligned}\dot{P} + \frac{1}{m}\vec{\nabla}(P\vec{\nabla}S) = 0, \\ \dot{S} + \frac{1}{2m}\vec{\nabla}S\vec{\nabla}S + V - \frac{\hbar^2}{4m}\left(\frac{\triangle P}{P} - \frac{\vec{\nabla}P\vec{\nabla}P}{2P^2}\right) = 0.\end{aligned} \tag{4.28}$$

The Schrödinger system in this representation turns out to be the starting point for a semiclassical approximation [4] in quantum mechanics as well as forming the basis of the de Broglie-Bohm interpretation of quantum mechanics [32].

Exercise
Obtain the Eqs. (4.28) by direct substitution of (4.26) into the Schrödinger system (2.130) and (2.131).

Example Let us consider the transformation $q^a \to q'^a = q^a$, $p_a \to p'_a = p_a + A_a(q)$. We have $\{q'^a, q'^b\}_z = 0$, $\{q'^a, p'_b\}_z = \delta^a{}_b$, $\{p'_a, p'_b\}_z = -\frac{\partial A_b}{\partial q^a} + \frac{\partial A_a}{\partial q^b}$. This will be zero for curl-free vector field: $A_a = \frac{\partial A}{\partial q^b}$, then $q'^a = q^a$, $p'_a = p_a + \frac{\partial A}{\partial q^a}$ represents a canonical transformation.

Exercise
Work out an example of a canonical transformation of the form $z^i \to z'^i = z^i + B^i(z^j)$.

4.2 Infinitesimal Canonical Transformations: Hamiltonian as a Generator of Evolution

Intuitively, infinitesimal canonical transformation is not very different from the identity transformation: $\tilde{z}^i = z^i + \delta z^i$, $\delta z^i \ll 1$. Its remarkable property is that it is generated by some function through the Poisson bracket: $\delta z^i = \{z^i, \Phi\}$. As will be seen below, finite canonical transformations have a similar (but not identical) structure.

Definition 1 Consider a family of transformations that are linear on the parameter λ

$$\tilde{z}^i(z^j, \lambda) = z^i + G^i(z)\lambda. \tag{4.29}$$

They are called *infinitesimal canonical transformations*, if they obey the canonicity condition (4.2) in linear order on λ, that is

$$\{\tilde{z}^i, \tilde{z}^j\} = \omega^{ij} + O(\lambda^2), \tag{4.30}$$

or, in other words:

$$\frac{\partial G^i}{\partial z^k}\omega^{kj} + \omega^{ik}\frac{\partial G^j}{\partial z^k} = 0. \tag{4.31}$$

Example Consider a family of univalent canonical transformations parameterized by λ, which includes an identity transformation at $\lambda = 0$

$$z^i \to z'^i(z^j, \lambda), \qquad z'^i(z^j, 0) = z^i. \tag{4.32}$$

Write the Taylor expansion about $\lambda = 0$. Ignoring all the higher order terms we have:

$$z'^i(z^j, \lambda) \approx \tilde{z}^i = z^i + G^i(z)\lambda, \tag{4.33}$$

where $G^i = \partial_\lambda z'^i(z, \lambda)\big|_{\lambda=0}$. This turns out to be an infinitesimal canonical transformation. For a small value of the parameter, $\lambda \ll 1$, the second term gives a leading contribution to the complete transformation z'^i.

Generator of infinitesimal canonical transformation

Assertion $\tilde{z}^i = z^i + G^i(z)\lambda$ represents an infinitesimal canonical transformation if and only if there is a function $\Phi(z)$ such that

$$G^i = \{z^i, \Phi\} = \omega^{ij}\partial_j\Phi. \tag{4.34}$$

Accordingly, any infinitesimal canonical transformation has the form

$$\tilde{z}^i = z^i + \{z^i, \Phi\}\lambda + O(\lambda^2). \tag{4.35}$$

Hence, properties of the infinitesimal canonical transformation are determined by a unique function $\Phi(z)$. It is called a *generator of infinitesimal transformation*.

Proof Supposing that (4.29) represents an infinitesimal canonical transformation, then Eq. (4.31) is satisfied. Contracting it with $\omega_{mi}\omega_{jn}$ we obtain the equation $\partial_n(\omega_{mk}G^k) - \partial_m(\omega_{nk}G^k) = 0$. This states that $\omega_{mk}G^k$ are components of the curl-free vector field. Therefore there is a potential, $\omega_{mk}G^k = \partial_m\Phi$, which proves the statement (4.34).

Conversely, consider a transformation of the form:

$$z^i \to \tilde{z}^i = z^i + \{z^i, \Phi\}\lambda, \tag{4.36}$$

determined by the function Φ. Using the Jacobi identity we obtain:

$$\begin{aligned}\{\tilde{z}^i, \tilde{z}^j\} &= \omega^{ij} + (\{z^i, \{z^j, \Phi\}\} + \{\{z^i, \Phi\}, z^j, \})\lambda + O(\lambda^2) \\ &= \omega^{ij} + (\{\{z^i, z^j\}, \Phi\}\lambda + O(\lambda^2) = \omega^{ij} + O(\lambda^2).\end{aligned} \tag{4.37}$$

Coordinate transformations $z \to z'$ can be used to define a map on the space of phase functions $A(z)$. By definition, the function A is mapped into A' according to the rule:

$$z \to z' \quad \Rightarrow \quad A \to A', \quad \text{where} \quad A'(z') = A(z). \tag{4.38}$$

That is, a value of the transformed function A' at z' coincides with the value of A at the point z. The difference

$$\delta_f A(z) \equiv A'(z) - A(z), \tag{4.39}$$

is called the *variation in form* of the function. For the case of an infinitesimal canonical transformation, the variation in form is governed by a generator:

$$\delta_f A(z) = \{\Phi, A\}\lambda + O(\lambda^2). \tag{4.40}$$

To confirm this, let us substitute Eq. (4.33) into the definition (4.38): $A'(z + \{z^i, \Phi\}\lambda + O(\lambda^2)) = A(z)$, or

$$A'(z) + \partial_i A'(z)\{z^i, \Phi\}\lambda + O(\lambda^2) = A(z), \tag{4.41}$$

which implies

$$\begin{gathered}\delta_f A(z) = \{\Phi, A'(z)\}\lambda + O(\lambda^2) = \\ \{\Phi, A(z) - O(\lambda)\}\lambda + O(\lambda^2) = \{\Phi, A\}\lambda + O(\lambda^2).\end{gathered} \tag{4.42}$$

In the passage from the first to the second line we have used Eq. (4.41) once again.

From infinitesimal to finite canonical transformations. It is worth noting that an infinitesimal canonical transformation is generally not a canonical transformation. But it can be used to construct a canonical transformation in terms of a power series of λ.

Assertion Given an infinitesimal canonical transformation $\tilde{z}^i = z^i + G^i\lambda = z^i + \{z^i, \Phi(z)\}\lambda$, the formula:

$$z'^i = e^{\lambda\{z^k, \Phi(z)\}\frac{\partial}{\partial z^k}} z^i, \tag{4.43}$$

represents a canonical transformation.

Proof We need to confirm that z'^i obeys Eq. (4.2). Remember that z'^i obeys the equation (see Sect. 2.3):

$$\frac{\partial z'^i}{\partial \lambda} = G^i(z'^k). \tag{4.44}$$

As a consequence, the function $W^{ij}(\lambda) \equiv \{z'^i, z'^j\}$ obeys the problem:

$$\frac{\partial W^{ij}}{\partial \lambda} = \{W^{ij}, \Phi\}, \qquad W^{ij}(0) = \omega^{ij}. \tag{4.45}$$

Note that the symplectic matrix ω^{ij} solves the problem. Since the problem has a unique solution, we conclude that $W^{ij} = \omega^{ij}$, that is, $\{z'^i, z'^j\} = \omega^{ij}$.

Evolution of a system as a canonical transformation. Hamiltonian as a generator of evolution. According to Sect. 2.3, a general solution to Hamiltonian equations of a system with the Hamiltonian $H(z)$ has the form

$$z^i(\tau) = e^{\tau\{z_0^k, H(z_0)\}\frac{\partial}{\partial z_0^k}} z_0^i. \tag{4.46}$$

This formula can be considered as a family of coordinate transformations parameterized by τ that relates the initial z_0 and the final z positions of the system. Comparing (4.46) with (4.43), we conclude that the general solution represents an example of canonical transformation with the generator being the Hamiltonian of a system. For a small value of τ, an approximate solution is given by the linear term of the power expansion

$$z^i(z_0, \tau) \approx z_0^i + \{z_0^i, H(z_0)\}\tau, \tag{4.47}$$

and turns out to be an infinitesimal canonical transformation. We conclude that the evolution of a dynamical system is a canonical transformation. Moreover, the Hamiltonian turns out to be the generator of the transformation.

4.3 Generating Function of Canonical Transformation

In this section we discuss a fairly large class of transformations called *free canonical transformations*. They have the remarkable property of being generated by phase-space functions (transition functions of the free transformation appear as partial derivatives of the generating function, see Eq. (4.57) below). Intuitively, this property can be expected from the observation that any canonical transformation is related to a conservative vector field, see Eqs. (4.17) (4.18) and (4.19). The potential of the field represents, in fact, the generating function.

4.3.1 Free Canonical Transformation and Its Function $F(q', p', \tau)$

Given the canonical transformation $q^a \to q'^a = q'^a(q, p, \tau)$, $p_a \to p'_a = p'_a(q, p, \tau)$ suppose that equations for q' can be resolved with respect to p: $q'^a = q'^a(q, p, \tau) \Rightarrow p_a = p_a(q, q', \tau)$. Transformations with this property are called *free canonical transformations*. Using this last equation, we can represent the variables p, p' in terms of q, q':

$$p_a = p_a(q, q', \tau), \quad p'_a = p'_a(q, q', \tau) \equiv p'_a(q, p(q, q', \tau), \tau). \tag{4.48}$$

By construction, these expressions can in turn be resolved with respect to q', p'. So, we can work with the canonical transformation in the form (4.48), where q, q' are considered as independent variables, instead of its original form, with q, p being independent.

For later use we now rewrite the potential $E(z', \tau)$, defined by Eq. (4.14), in an equivalent, but less symmetric form. Namely, let us write parts of the system (4.14) for q' and p' separately

$$\begin{aligned} -cq^b\frac{\partial p_b}{\partial q'^a} + cp_b\frac{\partial q^b}{\partial q'^a} - p'_a &= 2\frac{\partial E(z', \tau)}{\partial q'^a}, \\ -cq^b\frac{\partial p_b}{\partial p'_a} + cp_b\frac{\partial q^b}{\partial p'_a} + q'^a &= 2\frac{\partial E(z', \tau)}{\partial p'_a}. \end{aligned} \tag{4.49}$$

We immediately note that for the function:

$$F(z', \tau) \equiv E(z', \tau) + \frac{c}{2}q^b(z', \tau)p_b(z', \tau) - \frac{1}{2}q'^b p'_b, \tag{4.50}$$

the equations acquire more simple form:

$$cp_b\frac{\partial q^b}{\partial q'^a} - p'_a = \frac{\partial F(z', \tau)}{\partial q'^a}, \qquad cp_b\frac{\partial q^b}{\partial p'_a} = \frac{\partial F(z', \tau)}{\partial p'_a}. \tag{4.51}$$

Hence the Lemma on page 130 can be formulated in terms of F: the invariance of the symplectic form (4.13) under a transformation is equivalent to the existence of the potential F, which obeys Eqs. (4.51).

As we have seen in the previous section, a general solution to the Hamiltonian equations can be identified with a canonical transformation relating the initial and final positions of a system. The generating function F of the transformation turns out to be the Hamiltonian action; see Sect. 4.8.

Exercise
Show that under a canonical transformation, the original and the transformed Hamiltonians are related as follows:

$$\tilde{H}(z', \tau) = cH(z(z', \tau)) - cp_a(z', \tau)\frac{\partial q^a(z', \tau)}{\partial \tau} + \frac{\partial F(z', \tau)}{\partial \tau}. \tag{4.52}$$

4.3.2 Generating Function $S(q, q', \tau)$

The equalities (4.51) and (4.52) acquire a remarkably simple form in the variables q, q'. Let us introduce the *generating function* $S(q, q', \tau)$ according to the rule:

$$S(q, q', \tau) = F(q', p'(q, q', \tau), \tau), \tag{4.53}$$

Using the identities $\frac{\partial q^c}{\partial p'_b}\Big| \frac{\partial p'_b(q, p', \tau)}{\partial q^a} = \delta^c{}_a$, $\frac{\partial q^c}{\partial q'^a}\Big| + \frac{\partial q^c}{\partial p'_b}\Big| \frac{\partial p'_b}{\partial q'^a} = 0$, which follow from $q^c(q', p'(q, q', \tau), \tau) \equiv q^c$ and Eq. (4.51), we can calculate:

$$\begin{aligned}
\frac{\partial S}{\partial q^a} &= \frac{\partial F(z', \tau)}{\partial p'_b}\Big|_{p'(q,q',\tau)} \frac{\partial p'_b}{\partial q^a} \\
&= cp_c \frac{\partial q^c}{\partial p'_b}\Big| \frac{\partial p'_b}{\partial q^a} = cp_a, \\
\frac{\partial S}{\partial q'^a} &= \frac{\partial F(z', \tau)}{\partial q'^a}\Big| + \frac{\partial F(z', \tau)}{\partial p'_b}\Big| \frac{\partial p'_b}{\partial q'^a} \\
= cp_c \frac{\partial q^c}{\partial q'^a}\Big| &- p'_a + cp_c \frac{\partial q^c}{\partial p'_b}\Big| \frac{\partial p'_b}{\partial q'^a} = -p'_a.
\end{aligned} \tag{4.54}$$

The Hamiltonian (4.52) in the variables q, q' acquires the form:

$$\begin{aligned}
&\tilde{H}(q', p', \tau)|_{p(q,q',\tau)} = cH(q, p(q, q', \tau), \tau) \\
&-cp_a(q', p', \tau)|_{p'} \frac{\partial q^a(q', p', \tau)}{\partial \tau}\Big|_{p'} + \frac{\partial F(q', p', \tau)}{\partial \tau}\Big|_{p'} \\
&= cH(q, p(q, q', \tau), \tau) - \\
&cp_a(q, q', \tau)\left[\frac{\partial q^a(q', p'(q, q', \tau), \tau)}{\partial \tau} - \frac{\partial q^a}{\partial p'_b}\Big| \frac{\partial p'_b}{\partial \tau}\right] \\
&+\frac{\partial F(q', p'(q, q', \tau), \tau)}{\partial \tau} - \frac{\partial F(q', p', \tau)}{\partial p'_b}\Big| \frac{\partial p'_b}{\partial \tau} \\
&= cH(q, p(q, q', \tau), \tau) + \frac{\partial S}{\partial \tau}.
\end{aligned} \tag{4.55}$$

Thus we have obtained the following

Assertion Let $q^a \to q'^a = q'^a(q, p, \tau)$, $p_a \to p'_a = p'_a(q, p, \tau)$ be a free canonical transformation. From these expressions we write:

$$p_a = p_a(q, q', \tau), \qquad p'_a = p'_a(q, q', \tau) \equiv p'_a(q, p(q, q', \tau), \tau). \tag{4.56}$$

Then

(a) there is a *generating function*, $S(q, q', \tau)$, with $\det \frac{\partial^2 S}{\partial q^a \partial q'^b} \neq 0$, such that:

$$cp_a = \frac{\partial S}{\partial q^a}, \qquad p'_a = -\frac{\partial S}{\partial q'^a}. \tag{4.57}$$

If the function $F(z', \tau)$ (4.50) is known, the generating function can be constructed as follows

$$S(q, q', \tau) = F(q', p'(q, q', \tau), \tau). \tag{4.58}$$

(b) the transformed Hamiltonian (4.52), presented as a function of q, q', has the form

$$\tilde{H}(z', \tau)\Big|_{p'(q,q',\tau)} = cH(q, p(q, q', \tau), \tau) + \frac{\partial S(q, q', \tau)}{\partial \tau}. \tag{4.59}$$

This result can be inverted, giving a simple recipe for constructing a free canonical transformation:

Assertion Let $S(q^a, q'_b, \tau)$ be some function with $\det \frac{\partial^2 S}{\partial q^a \partial q'^b} \neq 0$, for any τ. Let us solve the algebraic equations $cp_a = \frac{\partial S}{\partial q^a}$, $p'_a = -\frac{\partial S}{\partial q'^a}$ with respect to q, p. Then the solution

$$p_a = \frac{1}{c} \left.\frac{\partial S}{\partial q^a}\right|_{q(q',p',\tau)} \equiv p_a(q', p', \tau), \qquad q^a = q^a(q', p', \tau), \tag{4.60}$$

is a free canonical transformation of valence c.

Proof We need to show that the functions $z^i(z', \tau)$ satisfy the relationship $\frac{\partial z^i}{\partial z'^k}\omega^{kl}\frac{\partial z^j}{\partial z'^l} = c^{-1}\omega^{ij}$. Let us consider, for example

$$\frac{\partial q^a}{\partial q'^c}\frac{\partial p_b}{\partial p'_c} - \frac{\partial p_b}{\partial q'^c}\frac{\partial q^a}{\partial p'_c} = c^{-1}\delta^a{}_b. \tag{4.61}$$

We use below notation of the type $\left.\frac{\partial^2 S(q,q',\tau)}{\partial q'^b \partial q^a}\right|_{q(q',p',\tau)} = \left(S_{q'q}\right)_{ba}$. The identity $p'_a \equiv -\left.\frac{\partial S}{\partial q'^a}\right|_{q(q',p',\tau)}$ implies $\frac{\partial q^a}{\partial q'^c} = -\left(S^{-1}_{q'q}\right)^{ad}(S_{q'q'})_{dc}$, $\frac{\partial q^c}{\partial p'_b} = -\left(S^{-1}_{q'q}\right)^{ca}$. Further, the identity $p_a = c^{-1}\left.\frac{\partial S}{\partial q^b}\right|_{q(q',p',\tau)}$ and the previous expressions imply $\frac{\partial p_b}{\partial p'_c} = -c^{-1}(S_{qq})_{bd}\left(S^{-1}_{q'q}\right)^{dc}$, $\frac{\partial p_b}{\partial q'^c} = -c^{-1}(S_{qq})_{bd}\left(S^{-1}_{q'q}\right)^{dg}(S_{q'q'})_{gc} + c^{-1}(S_{qq'})_{bc}$. Substitution of these expressions into l.h.s. of Eq. (4.61) turn it into an identity.

There are other types of generating functions that depend on any one of three pairs of variables: (q, p'), (q', p), (p, p'). They generate a free canonical transformations written in terms of the indicated variables. For instance, in the previous chapter we have discussed the generating function in terms of (q, p')-variables. The generating functions are related by means of the Legendre transformation (see Exercise 5 on page 83). As an example, let us construct the generating function $S(q, p', \tau)$ starting from $S(q, q', \tau)$. Suppose the function $p'_a(q', \ldots) \equiv \frac{\partial}{\partial q'^a}(-S)$ has an inverse one, $q'^a = q'^a(p', \ldots)$ (here the dots stand for the variables

q, τ considered as parameters). According to the Legendre theorem, its generating function is

$$
\begin{aligned}
S(q, p', \tau) &= \left(p'_a q'^a + S(q, q', \tau)\right)\Big|_{q'(q,p',\tau)} \\
&\equiv \left(p'_a q'^a + F(q, p', \tau)\right)\Big|_{q'(q,p',\tau)}.
\end{aligned}
\tag{4.62}
$$

This implies:

$$
q'^a = \frac{\partial S(q, p', \tau)}{\partial p'_a}, \qquad p_a = \frac{\partial S(q, p', \tau)}{\partial q^a}. \tag{4.63}
$$

The Hamiltonian $\tilde{H}$ as a function of q, p' is given by:

$$
\tilde{H}(z', \tau)\Big|_{q'(q,p',\tau)} = cH(q, p(q, p', \tau)\tau) + \frac{\partial S(q, p', \tau)}{\partial \tau}. \tag{4.64}
$$

Exercises

1. Let $z^i \to z'^i(z, \tau)$ be a canonical transformation. Show that there is no generating function of the form: $z'^i = \frac{\partial S(z,\tau)}{\partial z^i}$.
2. Find the canonical transformation relating the Hamiltonian formulations obtained from two Lagrangians which differ by a total derivative term: L and $L + \frac{dN(q)}{d\tau}$.
3. Let $q^a \to q'^a(q^b)$ be a general coordinate transformation of the configuration space. Find its extension $q'^a(q^b)$, $p'_a(q^b, p_c)$, which represents a univalent time-independent canonical transformation of the phase space (this result, together with Eq. (4.59), imply that the Hamiltonian of a non-singular Lagrangian theory in generalized coordinates represents the total energy of a system). *Answer*: $q'^a = q'^a(q)$, $p'_a = \frac{\partial q^b(q')}{\partial q'^a} p_b$, that is, p_a transforms as a vector under the general coordinate transformation of q^a.

4.4 Examples of Canonical Transformations

4.4.1 Evolution as a Canonical Transformation: Invariance of Phase-Space Volume

Let $z^i = f^i(c^j, \tau)$ be the general solution to the Hamiltonian equations

$$
\frac{dz^i}{d\tau} = \omega^{ij} \frac{\partial H_0}{\partial z^j}, \tag{4.65}
$$

Given the point z'^i of the phase space, the numbers c^i can be chosen in such a way that the trajectory passes through the point at the moment $\tau = 0$: $f^i(c^j, 0) = z'^i$. The latter equation can be resolved: $c^j = c^j(z')$. Substitution of this result into the general solution gives it as a function of the initial position: $f^i(z'^j, \tau)$, $f^i(z'^j, 0) = z'^i$. Its substitution into Eq. (4.65) implies the identity

$$\frac{df^i(z', \tau)}{d\tau} \equiv \omega^{ij} \left. \frac{\partial H_0(z)}{\partial z^j} \right|_{f(z',\tau)}. \tag{4.66}$$

We can consider the function $z^i = f^i(z', \tau)$ as a transition function between the coordinate systems (z', τ) and (z, τ) of the extended phase space. Thus we have the transformation

$$\begin{pmatrix} \tau \\ z'^i \end{pmatrix} \leftrightarrow \begin{pmatrix} \tau \\ z^i = f^i(z'^j, \tau) \end{pmatrix}. \tag{4.67}$$

By construction, all points of the curve $z^i = f^i(z_1'^j, \tau)$ have the same z'-coordinate $z' = z_1'$ in the system (τ, z'). That is, the curve is presented by the vertical straight line $z'^i = z_1'^i$ in the coordinates (τ, z'). We demonstrate now the validity of the equation

$$\{f^i(z', \tau), f^j(z', \tau)\}_{z'} = \omega^{ij}, \tag{4.68}$$

Hence, the *evolution of a physical system can be identified with the univalent time-dependent canonical transformation* (4.67), see Fig. 4.1 on page 141.

Denote the l.h.s. of Eq. (4.68) as W^{ij}. We look for a differential equation for the function W. By using Eq. (4.66), we obtain immediately

$$\frac{\partial}{\partial \tau} W^{ij} = \omega^{ik} H_{kl} W^{lj} - \omega^{jk} H_{kl} W^{li}, \quad \text{for all} \quad z'^i, \tag{4.69}$$

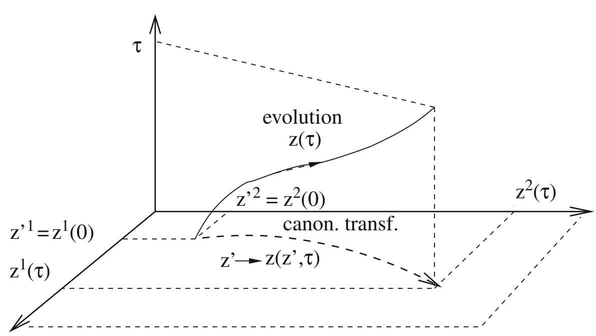

Fig. 4.1 The evolution of a system generates a canonical transformation

where $H_{kl} \equiv \partial_k \partial_l H_0(z)|_f$. Besides, since $f^i(z', 0) = z'^i$, the Eq. (4.68) implies the initial conditions $W^{ij}(0) = \omega^{ij}$. Note that $W^{ij}(\tau) = \omega^{ij}$ represents a solution with these initial conditions. It is the only solution, since the normal system (4.69) has a unique solution for given initial conditions.

Exercise
Check the validity of Eq. (4.68) up to the third order on τ by direct computations, using the Taylor expansion: $z^i = f^i(z', \tau) = f^i(z', 0) + \partial_\tau f^i|_0 \tau + \frac{1}{2}\partial_\tau^2 f^i|_0 \tau^2 + \ldots$.

According to the definition of canonical transformation, the result obtained can also be formulated as follows. Consider a dynamical system with the Hamiltonian H

$$\frac{dz^i}{d\tau} = \omega^{ij}\frac{\partial H(z)}{\partial z^j}. \tag{4.70}$$

Then the transformation inverse to (4.67), generated by the Hamiltonian flow of H_0, preserves the canonical form of Eq. (4.70): $\frac{dz'^i}{d\tau} = \omega^{ij}\frac{\partial \tilde{H}(z',\tau)}{\partial z'^j}$. In particular, the transformation inverse to (4.67) turns the system (4.65) into a system with $\tilde{H}_0 = 0$: $\frac{dz'^i}{d\tau} = 0$. It follows from the earlier observation that the curve $z^i = f^i(z', \tau)$ is presented by the vertical straight line $z' =$ const in the system (τ, z').

Exercise
Confirm this by direct computations with use of Eqs. (2.96), (2.95) and (4.68).

Hence, according to Eq. (4.59), the generating function of the transformation (4.67) obeys the equation

$$\frac{\partial S}{\partial \tau} = -H_0. \tag{4.71}$$

The univalent character of the canonical transformation has an interesting geometric interpretation. Consider a domain D' of the phase space with the volume $V' = \int_{D'} d^{2n}z'$. During the evolution, points z'^i of the domain are displaced into $z^i(z', \tau)$, and form the domain D, see Fig. 4.2 on page 143. Let us compute the volume of D. Making the change of variables $z^i(z', \tau)$ in a $2n$-dimensional integral, we obtain $V = \int_D d^{2n}z = \int_{D'} \left|\det \frac{\partial z^k}{\partial z'^i}\right| d^{2n}z' = \int_{D'} d^{2n}z' = V'$. Here Eq. (4.3) was used. Thus the volume of a phase-space domain retains a constant value during the evolution: $V = V'$.

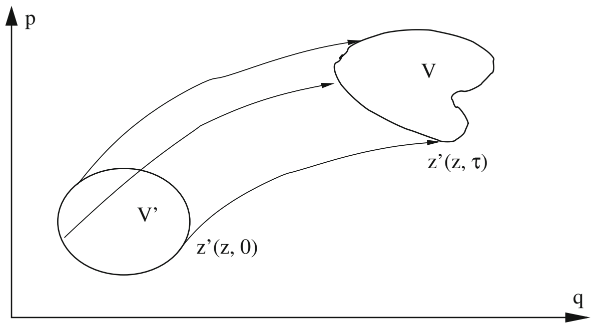

Fig. 4.2 Volume of a phase-space domain retains a constant value during the evolution: $V = V'$

4.4.2 Canonical Transformations in Perturbation Theory

Now, let us consider a dynamical system with the Hamiltonian being the sum of two terms

$$\frac{dz^i}{d\tau} = \omega^{ij} \frac{\partial (H_0(z) + H_1(z))}{\partial z^j}. \tag{4.72}$$

It is said that the initial system with H_0 is "perturbed" by H_1. Suppose that a general solution to the unperturbed system H_0 is known. Then the associated canonical transformation (4.78) turns the system (4.72) into a Hamiltonian system with the Hamiltonian H_1:

$$\frac{dz'^i}{d\tau} = \omega^{ij} \frac{\partial H_1(z(z', \tau))}{\partial z'^j}. \tag{4.73}$$

Actually, since the transformation is canonical, the new Hamiltonian is (see Eq. (4.59)) $H_0 + H_1 + \partial_\tau S$, but $\partial_\tau S = -H_0$ due to Eq. (4.71).

Exercise
Work out this result by direct computations with use of Eqs. (2.96), (2.95), (4.68).

If $z'^i(c^j, \tau)$ represents the general solution to the problem (4.73), we obtain a general solution to the problem (4.72) by taking a composition with the unperturbed solution (4.78), $z^i = z^i(z'^j(c^k, \tau), \tau)$.

We have shown, through the use of canonical transformations, how the perturbed problem (4.72) can be treated in the framework of the unperturbed one (4.66). According to the final result, $z^i = z^i(z'^j(c^k, \tau), \tau)$, perturbation in the energy of a system can be reformulated as perturbation of the initial conditions for the unperturbed problem. This observation turns out to be useful in quantum mechanics and in quantum field theory, where we can equally use either the Schrodinger, or the Heisenberg or the interaction pictures [33] to study an evolution of the quantum system.

Exercise
Apply this method to a one-dimensional problem with the Hamiltonian $H_0 + H_1 = \frac{1}{2}p^2 - \frac{e^\tau}{q-\tau p}$.

4.4.3 Coordinates Adjusted to a Surface

Consider the algebraic equation $F(q^a, p_b) = 0$. Suppose that it can be resolved with respect to one of the variables, say p_1: $p_1 = f(q^a, p_2, p_3, \ldots, p_n)$. We show here that there is a canonical transformation such that in the new coordinates the surface $F = 0$ is described by the equation $p'_1 = 0$, see Fig. 4.3.

This result appears to be interesting in the context of singular theories. In that case the system of Hamiltonian equations necessarily contains both differential equations (in the canonical form) and algebraic equations $F_\alpha = 0$ called *Dirac constraints*. So, all solutions to the equations of motion lie on a surface defined by these algebraic equations. Then it is natural to choose special coordinates such that the surface looks like a hyperplane in these coordinates: $z'_\alpha = 0$. We demonstrate that the corresponding transformation can be chosen to be canonical, that is, the canonical form of the Hamiltonian equations will not be spoiled in the new coordinates. This greatly simplifies analysis of the Hamiltonian equations and physical interpretation of a general singular theory [10].

It will be convenient to use the following notation: $z^i = (q^1, p_1, z^\alpha)$. Let us look for the new coordinates in the form

$$q'^1 = q^1, \qquad p'_1 = p_1 - f(q^1, z^\alpha), \qquad z'^\alpha = z^\alpha + h^\alpha(q^1, z^\alpha), \tag{4.74}$$

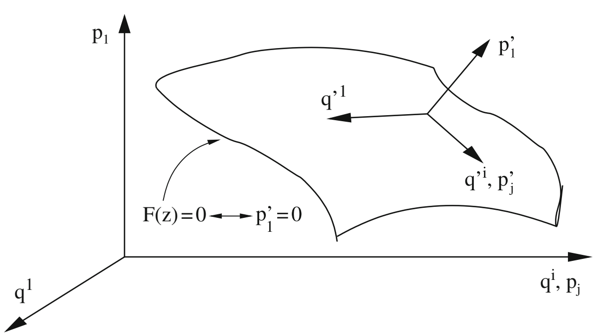

Fig. 4.3 Coordinates z'^i, adapted to the surface, can be chosen to be canonical

with undetermined functions h^α. We impose for them the following conditions:

$$h^\alpha(0, z^\alpha = 0. \tag{4.75}$$

The functions h^α can be chosen in such a way that the canonicity conditions

$$\{q'^1, q'^1\}_z = \{p'_1, p'_1\}_z = 0, \quad \{q'^1, p'_1\}_z = 1, \quad \{q'^1, z'^\alpha\}_z = 0, \tag{4.76}$$

$$\{p'_1, z'^\alpha\}_z = 0 \quad \Leftrightarrow \quad \frac{\partial h^\alpha}{\partial q^1} = \{z'^\alpha, f\}_z, \tag{4.77}$$

$$\{z'^\alpha, z'^\beta\}_z = \omega^{\alpha\beta}. \tag{4.78}$$

hold. Indeed, the Eqs. (4.76) are already satisfied. Equation (4.77) represents a first-order partial differential equation $\frac{\partial h^\alpha}{\partial q^1} + \frac{\partial f}{\partial z^\beta}\omega^{\beta\gamma}\frac{\partial h^\alpha}{\partial z^\gamma} = \omega^{\alpha\gamma}\frac{\partial f}{\partial z^\gamma}$. The Cauchy problem (4.75) for it has a unique solution $h^\alpha(q^1, z^\alpha)$; see, for example [4].

The solution automatically obeys Eq. (4.78). To confirm this, we write an equation for the function $\{z'^\alpha, z'^\beta\}$ by differentiating the commutator with respect to q^1. Using Eq. (4.77) and the Jacobi identity, we obtain

$$\frac{\partial}{\partial q^1}\{z'^\alpha, z'^\beta\} = \{\{z'^\alpha, z'^\beta\}, f\} \quad \text{for any fixed } \alpha, \beta, \tag{4.79}$$

while Eq. (4.75) implies the boundary condition:

$$\{z'^\alpha, z'^\beta\}\big|_{q^1=0} = \omega^{\alpha\beta}. \tag{4.80}$$

Note that the matrix $\omega^{\alpha\beta}$ obeys Eqs. (4.79) and (4.80). Since, as before, the problem has a unique solution, one concludes that (4.78) holds.

4.5 Transformation Properties of the Hamiltonian Action

According to Sect. 2.9, Hamiltonian equations (2.88) can be obtained by application of the principle of least action to the Hamiltonian action

$$S_H = \int d\tau (p_a \dot{q}^a - H(q, p)), \tag{4.81}$$

while for canonically transformed variables q', p' the corresponding equations follow from a similar expression with the Hamiltonian given by (4.52). It is interesting to see the deformation of the integrand in (4.81) after the substitution of $z(z', \tau)$. By direct substitution, we obtain

$$\left(p_a\dot{q}^a - H(z)\right)\big|_{z(z',\tau)} = c^{-1}\left(p'_a\dot{q}'^a - \left(cH(z(z',\tau)) - cp_a(z',\tau)\frac{\partial q^a(z',\tau)}{\partial\tau} + \frac{\partial F}{\partial\tau}\right) + \frac{dF}{d\tau}\right). \quad (4.82)$$

where $F(z',\tau)$ is precisely the function specified in Eq. (4.51). Note that the transformed Hamiltonian (4.52) appears on r.h.s. of the integrand. For a univalent transformation we write

$$p_a\dot{q}^a - H(q,p) = p'_a\dot{q}'^a - \tilde{H}(q',p',\tau) + \frac{dF(q',p',\tau)}{d\tau}. \quad (4.83)$$

The relationship among the integrands is a consequence of the previously obtained properties of canonical transformations. *Assuming* that the Eq. (4.83) holds for a canonical transformation, we easily find most of their properties, see [1, 12, 13].

Exercise
Does the Hamiltonian equations (4.22) follow from the principle of least action applied to Hamiltonian action with the integrand (4.82)? Consider also the cases of univalent and univalent time-independent canonical transformations.

4.6 Summary: Equivalent Definitions for Canonical Transformation

In Sect. 2.7 canonical transformations were defined as those preserving the standard form of Hamiltonian equations. In subsequent sections we have found a number of equivalent definitions. For convenience, we present here the resulting list:

Let $z^i \to z'^i(z^j,\tau)$ be a phase-space transformation. Then the following statements are equivalent and any one of them can be taken as a definition for the canonical transformation:

1. The transformation preserves the canonical form of Hamiltonian equations for any Hamiltonian system:

$$\dot{z}^i = \omega^{ij}\frac{\partial H}{\partial z^j} \overset{z\to z'}{\longrightarrow} \dot{z}'^i = \omega^{ij}\frac{\partial\tilde{H}(z',\tau)}{\partial z'^j}, \quad \text{any} \quad H, \quad \text{some} \quad \tilde{H}. \quad (4.84)$$

2. The transformation leaves the symplectic matrix invariant (disregarding a constant c)

$$\frac{\partial z'^k}{\partial z^i}\omega^{ij}\frac{\partial z'^l}{\partial z^j} = c\omega^{kl}, \quad \textit{or} \quad \{z'^k, z'^l\}_z = c\omega^{kl}, \quad c = \text{const}. \quad (4.85)$$

3. There is a function $E(z', \tau)$ such that

$$cz^j(z', \tau)\omega_{ji}\frac{\partial z^i(z', \tau)}{\partial z'^l} + \omega_{lj}z'^j = 2\frac{\partial E(z', \tau)}{\partial z'^l}. \tag{4.86}$$

4. There is a function $F(z', \tau)$ such that

$$cp_b\frac{\partial q^b}{\partial q'^a} - p'_a = \frac{\partial F(z', \tau)}{\partial q'^a}, \qquad cp_b\frac{\partial q^b}{\partial p'_a} = \frac{\partial F(z', \tau)}{\partial p'_a}. \tag{4.87}$$

   If the function $E(z', \tau)$ (4.86) is known, $F(z', \tau)$ can be constructed as follows

$$F(z', \tau) \equiv E(z', \tau) + \frac{c}{2}q^b(z', \tau)p_b(z', \tau) - \frac{1}{2}q'^b p'_b, \tag{4.88}$$

5. For the free transformation, there is a *generating function* $S(q, q', \tau)$, with $\det \frac{\partial^2 S}{\partial q^a \partial q'^b} \neq 0$, such that

$$cp_a(q, q', \tau) = \frac{\partial S}{\partial q^a}, \qquad p'_a(q, q', \tau) = -\frac{\partial S}{\partial q'^a}. \tag{4.89}$$

   If the function $F(z', \tau)$ (4.87) is known, the generating function can be constructed as follows

$$S(q, q', \tau) = F(q', p', \tau)\Big|_{p'(q,q',\tau)}. \tag{4.90}$$

4.7 Hamilton–Jacobi Equation

According to Eq. (4.52), the Hamiltonian has a nontrivial transformation law under the time-dependent canonical transformation. The transformed Hamiltonian $\tilde{H}$ depends on the function $F(z', \tau)$, which determines the transformation according to Eqs. (4.87) and (4.89). We can look for the F that makes $\tilde{H}$ as simple as possible, which implies an interesting method to look for a general solution to Hamiltonian equations

$$\dot{z}^i = \omega^{ij}\frac{\partial H}{\partial z^j}. \tag{4.91}$$

Performing the univalent canonical transformation

$$z^i \to z'^i = z'^i(z, \tau), \tag{4.92}$$

we have equations of motion for the new variables, $z' = \{z', \tilde{H}\}$, namely

$$\dot{z}'^i = \omega^{ij}\frac{\partial}{\partial z^j}\left(H(z(z',\tau)) - p_a(z',\tau)\frac{\partial q^a(z',\tau)}{\partial\tau} + \frac{\partial F(z',\tau)}{\partial\tau}\right) \tag{4.93}$$

Suppose we have found the transformation (4.92) which annihilates $\tilde{H}$:

$$\frac{\partial F(z',\tau)}{\partial\tau} - \left(p_a\partial_\tau q^a - H(z)\right)\big|_{z(z',\tau)} = 0, \tag{4.94}$$

then (4.93) acquires the form $\dot{z}'^i = 0$, and can be immediately solved: $z'^i = c^i =$ const. In the new coordinates the system is at rest. Now let us return to the initial variables: one substitutes this result into l.h.s. of Eq. (4.92) and solves them in relation to z:

$$z^i = z^i(\tau, c^j). \tag{4.95}$$

By construction, this gives the general solution to equations of motion (4.91).

The geometric interpretation of this procedure is very simple: we search for a special coordinate system (z'^i, τ) of the extended phase space, such that trajectories of the dynamical system look like vertical straight lines at these coordinates, see Fig. 4.4 on page 148.

According to this scheme, the problem (4.91) is replaced by the problem (4.94), which contains $2n+1$ unknown functions $z^i(z'^j,\tau)$, $F(z'^i,\tau)$. By construction they obey $2n$ Eqs. (4.87). Supposing that the transformation under investigation is free, the system (4.94) and (4.87) can be equally analyzed in terms of independent variables q, q' instead of q', p'. According to Eqs. (4.59) and (4.58), in the variables q, q' the system acquires the form[3]

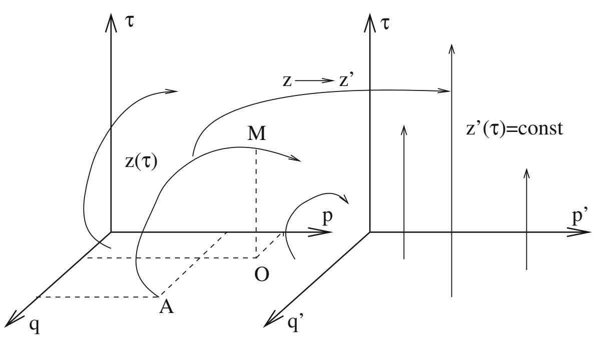

Fig. 4.4 Geometric interpretation of the Hamilton–Jacobi method: while coordinates of the point M in the system z are defined by projection along MO, in the system z' they are defined by projection along MA

[3] Since the general solution (4.95) determines the canonical transformation (4.92), the eq. (4.97) state that we search for the generating function of the evolution.

$$\frac{\partial S(q, q', \tau)}{\partial \tau} + H(q, p(q, q', \tau)) = 0, \tag{4.96}$$

$$p_a = \frac{\partial S(q, q', \tau)}{\partial q^a}, \qquad p'_a = -\frac{\partial S(q, q', \tau)}{\partial q'^a}. \tag{4.97}$$

Using the first equation from (4.97) in Eq. (4.96), the equation for S can be separated from those for p, p'. The closed equation for $S(q^a, \tau)$ is

$$\frac{\partial S(q^a, \tau)}{\partial \tau} + H\left(q^a, \frac{\partial S(q^a, \tau)}{\partial q^b}\right) = 0, \tag{4.98}$$

where q'^b have been omitted since they enter into the resulting equation as parameters. This partial differential equation is known as a *Hamilton–Jacobi equation*. Remind that solutions to partial differential equations generally depend on arbitrary functions. In particular, we can look for the so called *complete solution* that depends on n arbitrary numbers q'^b. Let $S(q^a, q'^b, \tau)$, with det $\frac{\partial S}{\partial q^a \partial q'^b} \neq 0$ be such a solution. Then Eq. (4.97) determines the free canonical transformation (4.92) which annihilates the Hamiltonian $\tilde{H}$. According to the previous analysis, solving the algebraic Eqs. (4.97) for $z^i = q^a, p_b$, we obtain the general solution $z^i = z^i(z'^j, \tau)$, to the Hamiltonian equations (4.91).

In short, the Hamilton–Jacobi method for solving Hamiltonian equations (4.91) can be formulated as follows:

1. Find the solution $S(q^a, q'^b, \tau)$, det $\frac{\partial^2 S}{\partial q^a \partial q'^b} \neq 0$ to the Hamilton–Jacobi equation

$$\frac{\partial S(q^a, \tau)}{\partial \tau} + H\left(q^a, \frac{\partial S(q^a, \tau)}{\partial q^b}\right) = 0, \tag{4.99}$$

which depends on n arbitrary numbers q'^b.

2. Write the expressions

$$p_a = \frac{\partial S(q, q', \tau)}{\partial q^a}, \qquad p'_a = \frac{\partial S(q, q', \tau)}{\partial q'^a}, \tag{4.100}$$

and resolve them in relation to q, p:

$$q^a = q^a(q', p', \tau), \qquad p_a = p_a(q', p', \tau). \tag{4.101}$$

These functions represent the general solution to the Eqs. (4.91), with $2n$ integration constants q'^a, p'_a.

Summing up, the problem to find a general solution to $2n$ ordinary differential Eqs. (4.91) can be replaced by the problem to find the solution S of partial differential Eq. (4.99) which depends on n arbitrary constants.

Example Using the method of separation of variables, a one-dimensional Hamilton–Jacobi equation can be solved for an arbitrary time-independent potential. Consider a particle on a straight line in the potential $U(x)$, $L = \frac{1}{2m}\dot{x}^2 - U$. Its Hamiltonian reads $H = \frac{1}{2m}p^2 + U$, hence the Hamilton–Jacobi equation is:

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + U = 0. \tag{4.102}$$

We look for a solution to the form $S(t, x) = S_1(t) + S_2(x)$. After substitution of this expression into (4.102) the variables t and x separate:

$$-\frac{\partial S_1(t)}{\partial t} = \frac{1}{2m}\left(\frac{\partial S_2(x)}{\partial x}\right)^2 + U(x). \tag{4.103}$$

This implies $\frac{\partial S_1}{\partial t} = -x'$, $\frac{1}{2m}\left(\frac{\partial S_2}{\partial x}\right)^2 + U(x) = x'$, where x' stands for a number. This represents the total energy of the system, as we can see from comparison of the last equation with the Hamiltonian. The equations can be immediately integrated out, we obtain

$$S = S_1 + S_2 = -x't + \int dx\sqrt{2m(x' - U)}. \tag{4.104}$$

Then Eq. (4.100) reads:

$$p = \sqrt{2m(x' - U)}, \qquad p' = -t + m\int \frac{dx}{\sqrt{2m(x' - U)}}. \tag{4.105}$$

It gives the general solution $x(t, x', p')$, $p(t, x', p')$ written in an implicit form. For example, for the free particle, $U = 0$, we obtain:

$$p = \sqrt{2mx'}, \qquad x = \frac{\sqrt{2mx'}}{m}t + \frac{\sqrt{2mx'}\,p'}{m}, \tag{4.106}$$

which is the expected expression $x = \frac{c_1}{m}t + c_2$, where $c_1 \equiv \sqrt{2mx'}$ represents the initial momentum and $c_2 \equiv \sqrt{\frac{2x'}{m}}\,p'$ is the initial position of the particle.

Let us consider also the harmonic oscillator, $U = \frac{k}{2}x^2$. Equation (4.105) acquires the form

$$p = \sqrt{2mx' - kmx^2},$$
$$p' = -t + m\int \frac{dx}{\sqrt{2m(x' - \frac{k}{2}x^2)}} = -t + \sqrt{\frac{m}{k}}\arcsin\sqrt{\frac{k}{2x'}}x. \quad (4.107)$$

Solving these equalities with respect to x, p we obtain the general solution

$$x(t) = \sqrt{\frac{2x'}{k}}\sin(\omega t + \delta) = x_0\cos\omega t + \frac{p_0}{\sqrt{km}}\sin\omega t,$$
$$p(t) = \sqrt{2mx'}\cos(\omega t + \delta) = p_0\cos\omega t - \sqrt{km}x_0\sin\omega t, \quad (4.108)$$

where $\omega \equiv \sqrt{\frac{k}{m}}$, $\delta \equiv \sqrt{\frac{k}{m}}p'$, and the initial position and momentum are given by $x_0 = \sqrt{\frac{2x'}{k}}\sin\delta$, $p_0 = \sqrt{2mx'}\cos\delta$.

4.8 Action Functional as a Generating Function of Evolution

In the Sect. 4.4.1 we have seen that a general solution to Hamiltonian equations can be identified with a canonical transformation. It was demonstrated in Sect. 4.2 that the Hamiltonian represents the generator of the corresponding infinitesimal transformation. Here we construct the generating function of the finite transformation. When the general solution to the Hamiltonian equations is known, it is possible to construct the complete solution S of the Hamilton–Jacobi equation in closed form in terms of the Hamiltonian action S_H, see Eq. (4.113) below. According to the previous section, this S represents the generating function of the evolution.

Let

$$z = z(z', \tau), \qquad z(z', \tau_0) = z', \quad (4.109)$$

be a general solution to the Hamiltonian equations as a function of initial values z' (see discussion at the end of Sect. 4.2). Let us substitute (4.109) into Eq. (4.94)

$$\frac{\partial F(z', \tau)}{\partial \tau} = \left(p_a \partial_\tau q^a - H(z)\right)\Big|_{z(z',\tau)}. \quad (4.110)$$

This is the Hamilton–Jacobi equation written in terms of F. Equation (4.87), taken at $\tau = 0$, implies that $F(q', p', \tau_0) = \text{const}$. Since F is defined up to a constant, we omit it in the following:

$$F(q', p', \tau_0) = 0. \quad (4.111)$$

Since the r.h.s. of Eq. (4.110) now represents the known function of τ, we immediately find its solution subject to the initial condition (4.111)

$$F(q', p', \tau) = \int_{\tau_0}^{\tau} d\tau \; (p\dot{q} - H)|_{z(z',\tau)} \, . \tag{4.112}$$

The r.h.s. is just the Hamiltonian action written as a function of the initial position and momentum z' of the system. Using the relationship (4.90) between F and the generating function, we obtain:

$$\begin{aligned} S(q, q', \tau) = & \; F(q', p', \tau)\big|_{p'(q,q',\tau)} \\ = & \left[\int_{\tau_0}^{\tau} d\tau \; (p\dot{q} - H)|_{z(z',\tau)} \right] \Bigg|_{p'(q,q',\tau)} \, . \end{aligned} \tag{4.113}$$

Being a complete solution to the Hamilton–Jacobi equation, S represents the generating function of the canonical transformation $z' \to z$ associated with the general solution. The expression on the r.h.s. is the Hamiltonian action represented as a function of initial and final position. Hence it can be said that the *Hamiltonian action is the generating function of canonical transformation along true trajectories: it transforms the system coordinates from one time to another*.

According to the previous section, knowledge of the complete integral of the Hamilton–Jacobi equation allows us to construct the general solution to the Hamiltonian equations of motion. Here this result has been inverted. Hence, we have the mathematically notable fact that searching for a complete solution to the partial differential equation $\frac{\partial S}{\partial \tau} + H\left(q^a, \frac{\partial S}{\partial q^b}\right) = 0$ is equivalent to searching for a general solution to the system of ordinary differential equations $\dot{q}^a = \frac{\partial H}{\partial p_a}$, $\dot{p}_a = -\frac{\partial H}{\partial q^a}$.

Example As an illustration, we use a general solution to the harmonic oscillator equations of motion and the formula (4.113) to construct the corresponding generating function. Substitution of Eq. (4.113) into the expression $\int_0^t (p\dot{x} - H) = \int_0^t (\frac{1}{2m} p^2 - \frac{k}{2} x^2)$, it reads:

$$\begin{aligned} & \int_0^t \left[(\frac{1}{2m} p_0^2 - \frac{k}{2} x_0^2) \cos 2\omega t - \omega x_0 p_0 \sin 2\omega t \right] \\ = & \frac{1}{\omega} (\frac{1}{2m} p_0^2 - \frac{k}{2} x_0^2) \sin \omega t \cos \omega t - x_0 p_0 \sin^2 \omega t . \end{aligned} \tag{4.114}$$

From Eq. (4.108) we find the initial momentum p_0 presented as a function of initial and final position, $p_0 = \sqrt{km}(x - x_0 \cos \omega t) \sin^{-1} \omega t$. According to (4.113), the generating function is obtained substituting p_0 into Eq. (4.114). After some computations we arrive at:

$$S(t, x_0, p_0) = \frac{\sqrt{km}}{2}(x^2 - x_0^2)\arctan\omega t - \sqrt{km}xx_0\sin^{-1}\omega t. \quad (4.115)$$

By direct substitution we verify that it obeys the Hamilton–Jacobi equation (4.102).

Exercise. Show by direct substitution that S of Eq. (4.113) obeys the Hamilton–Jacobi equation.

Chapter 5
Integral Invariants

This chapter is devoted to the discussion of the theory of integral invariants, which reveals an interesting structure of the general solution to Hamiltonian equations. We discuss the basic Poincaré-Cartan and Poincaré integral invariants that represent line integrals of a special vector field defined on extended phase space. The integrals retain the same value for any closed contour taken on a given two-dimensional surface formed by solutions to the Hamiltonian equations. As will be discussed in Sect. 5.1.3, this property could be taken as a basic principle of mechanics, instead of the principle of least action. Besides their applications in mechanics, integral invariants are widely used in the theory of differential equations, see [1, 4].

5.1 Poincaré-Cartan Integral Invariant

5.1.1 Preliminaries

We recall here some facts related to the description of a surface and a curve in Euclidean space. Consider the space $\mathbb{R}^{2n+1}$ parameterized by the coordinates $(z^i, \tau) \equiv (q^a, p_b, \tau)$, $i = 1, 2, \ldots, 2n$, $a, b = 1, 2, \ldots n$. Let S be the two-dimensional cylindrical surface embedded in $\mathbb{R}^{2n+1}$ (see Fig. 5.1 on page 156) Henceforth this will be called a *tube*. Let $\beta, \alpha, \alpha \subset [0, l]$ be the coordinates of a coordinate system established on S. Then the points $M(\tau, z^i)$ of the surface have the corresponding coordinates β, α. This implies the parametric equations that describe the embedding of the surface in $\mathbb{R}^{2n+1}$

$$S: \quad \begin{cases} z^i = z^i(\beta, \alpha), \\ \tau = \tau(\beta, \alpha). \end{cases} \tag{5.1}$$

By construction, we have $\tau(\beta, 0) = \tau(\beta, l)$, $z^i(\beta, 0) = z^i(\beta, l)$.

Suppose that the curve C on S can be described by the equation $\beta = \beta(\alpha)$. Then the parametric equations

A. Deriglazov, *Classical Mechanics*, DOI 10.1007/978-3-642-14037-2_5,

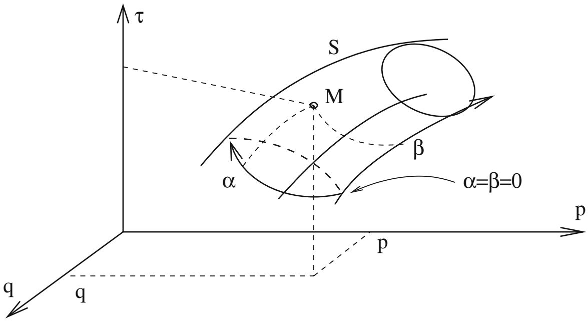

Fig. 5.1 Point $M(\tau, z^i)$ on the tube has the coordinates β, α. This implies parametric equations of the tube $z^i = z^i(\beta, \alpha)$, $\tau = \tau(\beta, \alpha)$. If τ is taken as one of the coordinates: $\beta = \tau$, we have the parametric equations $z^i = z^i(\tau, \alpha)$, $\tau = \tau$

$$C: \quad \begin{cases} z^i = z^i(\beta(\alpha), \alpha) \equiv z^i(\alpha), \\ \tau = \tau(\beta(\alpha), \alpha) \equiv \tau(\alpha), \end{cases} \tag{5.2}$$

describe its embedding in $\mathbb{R}^{2n+1}$.

We will be interested in the surfaces formed by a one-parameter family $z^i(\tau, \alpha)$ of solutions to the first-order system[1]

$$\dot{q}^a = Q^a(q, p), \qquad \dot{p}_a = P_a(q, p), \tag{5.3}$$

where Q, P are given functions. Remind that the Hamiltonian system is a particular case of (5.3), when there is a function $H(q, p)$, such that

$$Q^a = \frac{\partial H}{\partial p_a}, \qquad P_a = -\frac{\partial H}{\partial q^a}. \tag{5.4}$$

Comment To construct an example of the family, suppose that the general solution $z^i(\tau, c^j)$, $z^i(0, c^j) = c^j$ of (5.3) is known. Let $c^i = f^i(\alpha)$, $\tau = 0$ be the parametric equations of a closed curve in $\mathbb{R}^{2n+1}$. Then $z^i(\tau, \alpha) \equiv z^i(\tau, c^j(\alpha))$ represents an example of the one-parameter family; see Fig. 5.2 on page 157.

For the tube formed by solutions to Eq. (5.3), one can take τ as one of the coordinates on the surface. That is, the coordinate system on S is now $\tau, \alpha, \alpha \subset [0, l]$. Then the parametric equations of the surface are:

$$S: \quad \begin{cases} z^i = z^i(\tau, \alpha), \\ \tau = \tau. \end{cases} \tag{5.5}$$

By construction:

[1] All the results of this section remain true for the functions Q, P with the manifest dependence on time.

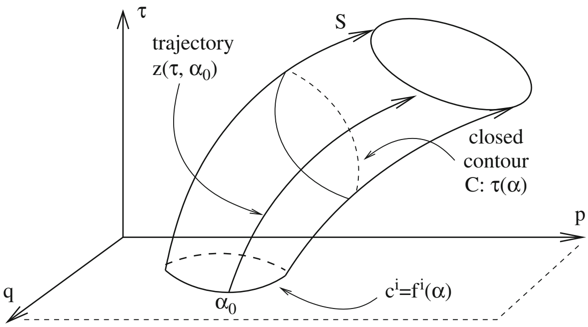

Fig. 5.2 The tube of trajectories can be constructed starting from the "initial-value" curve $c^i = f^i(\alpha)$. The Poincaré-Cartan integral is defined by using an arbitrary closed contour $C \subset S$

$$z^i(\tau, 0) = z^i(\tau, l), \tag{5.6}$$

and for any fixed $\alpha = \alpha_0$, the curve $z^i(\tau, \alpha_0)$ represents a solution to Eq. (5.3).

Let a curve $C \subset S$ be described by the equation $\tau = \tau(\alpha)$. Then the corresponding parametric equations of its embedding into $\mathbb{R}^{2n+1}$ are:

$$C: \quad \begin{cases} z^i = z^i(\tau(\alpha), \alpha) \equiv z^i(\alpha), \\ \tau = \tau(\alpha). \end{cases} \tag{5.7}$$

5.1.2 *Line Integral of a Vector Field, Hamiltonian Action, Poincaré-Cartan and Poincaré Integral Invariants*

Consider the vector field $\vec{V}(z^i, \tau)$=($v_a(z^i, \tau), u^b(z^i, \tau), v(z^i, \tau)$) defined on the extended phase space $\mathbb{R}^{2n+1}$. Then we define the *line integral of the vector field* along the oriented curve $C_{M\tilde{M}}$, see Fig. 5.3 on page 158:

$$\int_C \vec{V} d\vec{r} = \int_C v_a dq^a + u^b dp_b + v d\tau$$
$$\equiv \lim_{N\to\infty} \sum_{\nu=1}^{N} (\overrightarrow{V(M_\nu)}, \overrightarrow{\triangle r_\nu}), \tag{5.8}$$

where $(\vec{V}, \overrightarrow{\triangle r_\nu})$ is the scalar product $v_a \triangle q^a + u^b \triangle p_b + v \triangle \tau$. If C is given in the parametric form $z^i = z^i(\gamma)$, $\tau = \tau(\gamma)$, the line integral can be presented in terms of the definite integral as follows:

$$\int_C \vec{V} d\vec{r} = \int_{\gamma_1}^{\gamma_2} (\ v_a(\gamma)\frac{dq^a}{d\gamma} + u^b(\gamma)\frac{dp_b}{d\gamma} + v(\gamma)\frac{d\tau}{d\gamma}\ )d\gamma, \tag{5.9}$$

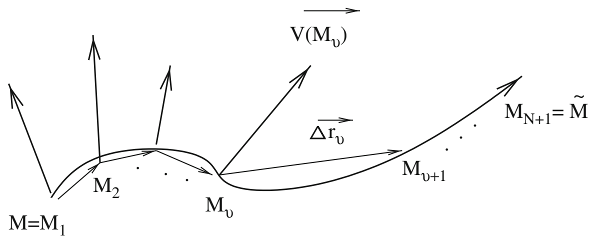

Fig. 5.3 To define the line integral of the vector field $\vec{V}$, we replace the oriented curve by a sequence of the following displacement vectors: $C_{M\tilde{M}} \to \bigcup_{\nu=1}^{N} \vec{\triangle r_\nu}$

where $\vec{V}(\gamma) = \vec{V}(z^i(\gamma), \tau(\gamma))$.

Let C: $\tau = \tau(\alpha)$ be a closed contour on the tube of solutions (5.5). The line integral (5.8) is called the *first-order integral invariant*, if its value does not depend on the choice of the contour on the tube. If C: $\tau = const$ is the closed contour composed of simultaneous points of the tube, the integral reduces to

$$\oint_C V_i dz^i = \oint_C v_a dq^a + u^b dp_b, \tag{5.10}$$

and is called the *first-order universal integral invariant*.

We will be mainly interested in a rather particular case of the vector field given by the expression

$$\vec{V}(q^a, p_b, \tau) = (\, p_a, 0, -H(q^a, p_b)\,), \tag{5.11}$$

where H is the Hamiltonian of the system (5.3) and (5.4). Note that $\vec{V}$ is orthogonal to any p-axis. The line integral acquires the form

$$\int_C p_a dq^a - H d\tau. \tag{5.12}$$

For the curve allowed as a trajectory of the system (5.3) and (5.4), the line integral (5.12) can be identified with the Hamiltonian action. Indeed, consider the curve that can be described in the parametric form as follows: $z^i = z^i(\tau)$, $\tau_1 \leq \tau \leq \tau_2$. Then Eq. (5.12) acquires the form

$$\int_{\tau_1}^{\tau_2} (p_a \dot{q}^a - H) d\tau \equiv S_H. \tag{5.13}$$

Now, consider a curve that corresponds to the closed contour. The line integral (5.12) along the closed contour is called the *Poincaré-Cartan integral invariant*

$$I = \oint_C p_a dq^a - H d\tau. \tag{5.14}$$

Note that, unlike the previous case, the closed contour cannot be the allowed trajectory of the system (5.3) and (5.4). For the simultaneous contour C: $\tau = \text{const}$, the integral reduces to $I_1 = \oint_C p_a dq^a$ and is called the *Poincaré (universal) integral invariant*.

We specify the expression of the Poincaré-Cartan integral for the closed contour that lies on the tube of trajectories S (5.5). Let C: $\tau = \tau(\alpha)$ be the equation of the contour in the coordinate system established on S. Then the corresponding parametric equations are (5.7), and the Poincaré-Cartan integral is represented by the definite integral

$$I = \int_0^l d\alpha \left(p(\tau(\alpha), \alpha) \frac{dq(\tau(\alpha), \alpha)}{d\alpha} - H(z(\tau(\alpha), \alpha)) \frac{d\tau}{d\alpha} \right). \tag{5.15}$$

Summing up, we have seen that *a line integral of the vector field (5.11), being computed along the proper classes of curves, reduces either to the Hamiltonian action, or to the Poincaré-Cartan integral, or to the Poincaré universal integral.*

By construction, the Poincaré-Cartan integral could depend on a choice of the contour C: $I = I_C$. Remarkably enough, it happens to be contour-independent: I does not change its value in the case of an arbitrary displacement (with deformation) of the contour along the tube. This is one of the results discussed in the next section.

5.1.3 Invariance of the Poincaré–Cartan Integral

Here we demonstrate that I has the same value for any contour taken on a given tube of trajectories of the Hamiltonian system. Conversely, the contour-independence of I (constructed with the help of a function H) on a tube of solutions to the system (5.3) implies that the system is the Hamiltonian one. More exactly, we have

Assertion For the system

$$\dot{q}^a = Q^a(q, p), \qquad \dot{p}_a = P_a(q, p), \tag{5.16}$$

let $z^i(\tau, \alpha)$, $\alpha \subset [0, l]$ be a one-parameter family of solutions which form a tube: $z^i(\tau, 0) = z^i(\tau, l)$. Then the following statements are equivalent:

(a) The system is a Hamiltonian one: there is a function $H(z)$ such that

$$Q^a = \frac{\partial H}{\partial p_a}, \qquad P_a = -\frac{\partial H}{\partial q^a}. \tag{5.17}$$

(b) There is a function $H(z)$ such that a value of the Poincaré–Cartan integral

$$I = \oint_C p_a dq^a - H d\tau. \tag{5.18}$$

does not depend on the choice of the closed contour C on the tube.

Proof Henceforth we will frequently use the following notation: $\dot{z}^i \equiv \frac{\partial z^i(\tau,\alpha)}{\partial \tau}$, $z'^i \equiv \frac{\partial z^i(\tau,\alpha)}{\partial \alpha}$.

(A) Let the system (5.16) be a Hamiltonian one. The invariance of I turns out to be closely related to the properties of the Hamiltonian action in the passage from one trajectory to another. Consider two closed contours C_1: $\tau_1(\alpha)$ and C_2: $\tau_2(\alpha)$ on the tube S, see Fig. 5.4 on page 160. For any fixed α, we write the line integral (5.12) along the solution $z^i(\tau,\alpha)$ to the system (5.16)

$$S_H(\alpha) = \int_{\tau_1(\alpha)}^{\tau_2(\alpha)} \left(p_a(\tau,\alpha)\frac{\partial q^a(\tau,\alpha)}{\partial \tau} - H(\,z(\tau,\alpha)\,)\right) d\tau. \tag{5.19}$$

This gives the value of the Hamiltonian action for the trajectory. So, the function $S_H(\alpha)$ describes the variation of the Hamiltonian action in the passage from one trajectory to another. Since the values $\alpha = 0, l$ correspond to the same trajectory, we have

$$S_H(l) = S_H(0), \tag{5.20}$$

from this it follows:

$$\int_0^l \frac{dS_H(\alpha)}{d\alpha} d\alpha = 0. \tag{5.21}$$

Let us compute the variation rate

$$\frac{dS_H(\alpha)}{d\alpha} = \left(p\frac{\partial q}{d\tau} - H \right)\Bigg|_{\tau_2(\alpha)} \frac{d\tau_2}{d\alpha} - (\tau_2 \to \tau_1) +$$
$$\int_{\tau_1(\alpha)}^{\tau_2(\alpha)} \left(p'\dot{q} + p\frac{\partial}{\partial \tau}q' - \frac{\partial H}{\partial p}p' - \frac{\partial H}{\partial q}q' \right) d\tau. \tag{5.22}$$

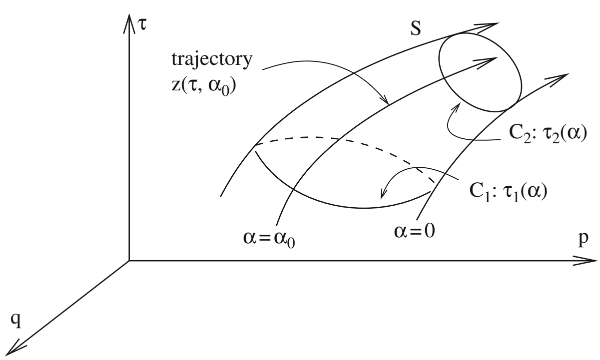

Fig. 5.4 For any fixed $\alpha = \alpha_0$, the function $S_H(\alpha_0)$ is the Hamiltonian action computed along the solution $z(\tau,\alpha_0)$ of Eq. (5.16) between the points $\alpha_1(\tau_0)$ and $\alpha_2(\tau_0)$. The Poincaré-Cartan integral has the remarkable property: $I_{C_1} = I_{C_2}$

Integration by parts of the second term gives

$$pq'\Big|_{\tau_1(\alpha)}^{\tau_2(\alpha)} - \int_{\tau_1(\alpha)}^{\tau_2(\alpha)} q'\dot{p}d\tau. \tag{5.23}$$

Then the variation rate is

$$\begin{aligned}\frac{dS_H(\alpha)}{d\alpha} = p(\tau_2(\alpha), \alpha)\left(\dot{q}|_{\tau_2(\alpha)}\frac{d\tau_2}{d\alpha} + q'\big|_{\tau_2(\alpha)}\right)\\ -H(z(\tau_2(\alpha), \alpha)\,)\frac{d\tau_2}{d\alpha} - (\tau_2 \to \tau_1)\\ +\int_{\tau_1(\alpha)}^{\tau_2(\alpha)}\left(p'\left(\dot{q} - \frac{\partial H}{\partial p}\right) - q'\left(\dot{p} + \frac{\partial H}{\partial q}\right)\right)d\tau.\end{aligned} \tag{5.24}$$

The last line vanishes due to Eqs. (5.16) and (5.17), while the first line is equal to $p\frac{dq}{d\alpha}$. Thus we have

$$\begin{aligned}\frac{dS_H(\alpha)}{d\alpha} = p(\tau_2(\alpha), \alpha)\frac{dq(\tau_2(\alpha), \alpha)}{d\alpha} - H(z(\tau_2(\alpha), \alpha))\frac{d\tau_2}{d\alpha}\\ -(\tau_2 \to \tau_1).\end{aligned} \tag{5.25}$$

Note that the r.h.s. of (5.25) coincides with the integrand of the Poincaré-Cartan integral (5.15). Inserting this expression into Eq. (5.21), we obtain the desired result: $I_{C_1} = I_{C_2}$ for any closed contours C_i on S.

(B) Suppose that the integral (5.18) with a given function H is contour-independent on the tube of the system (5.16). Let C': $\tau'(\alpha)$ be an arbitrary closed contour near C: $\tau(\alpha)$, and let us denote $\tau'(\alpha) - \tau(\alpha) \equiv \delta\tau(\alpha)$. Due to the contour independence, we have $I_{C'} - I_C = 0$. This implies that the variation vanishes: $\delta I = (I_{C'} - I_C)|_{\text{linear part on }\delta\tau} = 0$. On the other hand, it can be computed directly, performing an expansion of $I_{C'}$ around the point $\tau(\alpha)$ (below, the notation $|$ means the substitution of $\tau(\alpha)$). Using $z(\tau(\alpha)+\delta\tau, \alpha) = z(\tau(\alpha), \alpha) + \frac{\partial z(\tau,\alpha)}{\partial\tau}\Big|\delta\tau + O^2(\delta\tau)$, we obtain:

$$\begin{aligned}\delta I = \int_0^l d\alpha\left(\dot{p}|\,\frac{dq(\tau(\alpha), \alpha)}{d\alpha}\delta\tau +\right.\\ p(\tau(\alpha), \alpha)\frac{d}{d\alpha}\left(\dot{q}|\,\delta\tau\right) - H\frac{d\delta\tau}{d\alpha} -\\ \left.\frac{\partial H}{\partial q}\dot{q}\Big|\frac{d\tau}{d\alpha}\delta\tau - \frac{\partial H}{\partial p}\dot{p}\Big|\frac{d\tau}{d\alpha}\delta\tau\right).\end{aligned} \tag{5.26}$$

The first line can be presented as

$$\dot{p}q'\big|\,\delta\tau + \dot{p}\dot{q}|\,\frac{d\tau}{d\alpha}\delta\tau, \tag{5.27}$$

while integration by parts in the second line leads to the expressions

$$\int_0^l \left(-\, p'\dot{q} \big| \, \delta\tau - \, \dot{p}\dot{q} \big| \, \frac{d\tau}{d\alpha}\delta\tau \right) d\alpha + \, p\dot{q} \big| \, \delta\tau \big|_0^l \, , \tag{5.28}$$

$$\int_0^l \left(\frac{\partial H}{\partial q} \left(\dot{q}\frac{d\tau}{d\alpha} + q' \right) \bigg| \, \delta\tau + \right.$$
$$\left. \frac{\partial H}{\partial p} \left(\dot{p}\frac{d\tau}{d\alpha} + p' \right) \bigg| \, \delta\tau \right) d\alpha - \, H\delta\tau \big|_0^l \, . \tag{5.29}$$

Since the values $\alpha = 0, l$ correspond to the same point on the tube, the last terms in Eqs. (5.28) and (5.29) vanish. Bringing together the remaining terms, we obtain

$$\delta I = \int_0^l d\alpha \left[q' \left(\dot{p} + \frac{\partial H}{\partial q} \right) \bigg| - p' \left(\dot{q} - \frac{\partial H}{\partial p} \right) \bigg| \right] \delta\tau(\alpha). \tag{5.30}$$

Using Eq. (5.16) and the contour independence $\delta I = 0$, we have:

$$\int_0^l d\alpha \left[q' \big| \left(P + \frac{\partial H}{\partial q} \right) \bigg|_{z(\tau(\alpha),\alpha)} - \right.$$
$$\left. p' \big| \left(Q - \frac{\partial H}{\partial p} \right) \bigg|_{z(\tau(\alpha),\alpha)} \right] \delta\tau(\alpha) = 0. \tag{5.31}$$

Being true for any $\delta\tau(\alpha)$ as well as for any contour $\tau(\alpha)$, this equality implies Eq. (5.17).

The affirmation demonstrated means, in particular, that for a given Poincaré–Cartan integral there is a unique system of differential equations that admits this integral as the integral invariant. This statement could be taken as the basic principle of mechanics, instead of the principle of least action.

5.2 Universal Integral Invariant of Poincaré

Let us consider a particular case of the Poincaré-Cartan invariant $I = \oint p_a dq^a - H d\tau$, when the closed contour C is formed by *simultaneous points* of the system. It corresponds to the points of intersection of the tube with the hyperplane $\tau = \tau_0 = \text{const}$, see Fig. 5.5 on page 163. We have $d\tau = 0$ and the Poincaré–Cartan invariant acquires the form

$$I_1 = \oint_C p_a dq^a. \tag{5.32}$$

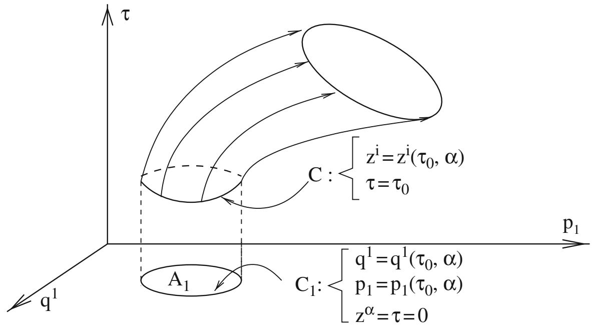

Fig. 5.5 Closed contour of simultaneous points C and its projection C_1 on the plane (q^1, p_1)

This is called the *Poincaré (universal) integral invariant*. The equation of the contour is $\tau = \tau_0$, and the corresponding parametric equations are $\tau = \tau_0$, $z^i = z^i(\tau_0, \alpha)$. This implies the following expression for I_1 in terms of the definite integral

$$I_1 = \int_0^l p_a(\tau_0, \alpha) \frac{\partial q^a(\tau_0, \alpha)}{\partial \alpha} d\alpha. \tag{5.33}$$

Being a particular case of I, the Poincaré integral invariant has similar properties. In particular, the Assertion of the previous section can be reformulated for I_1 as follows:

Assertion For the system

$$\dot{q}^a = Q^a(q, p), \qquad \dot{p}_a = P_a(q, p), \tag{5.34}$$

let $z^i(\tau, \alpha)$, $\alpha \subset [0, l]$ be a one-parameter family of solutions which form a tube: $z^i(\tau, 0) = z^i(\tau, l)$. Then the following statements are equivalent:

(a) The system is a Hamiltonian one: there is a function $H(z)$ such that

$$Q^a = \frac{\partial H}{\partial p_a}, \qquad P_a = -\frac{\partial H}{\partial q^a}. \tag{5.35}$$

(b) The value of the Poincaré integral

$$I_1 = \oint_C p_a dq^a \tau. \tag{5.36}$$

does not depend on the choice of the simultaneous closed contour C on the tube.

Since H is not presented in the expression for I_1, it is an invariant of any Hamiltonian system. This explains the terminology "universal".

It should be stressed that I_1 is invariant under the replacement $C \to C'$, where both contours are simultaneous. So, C' can be considered as the result of the evolution of the points of C for the same time interval. Recall that the evolution is an example of a canonical transformation. This explains why the Poincaré integral invariant can be used to study properties of canonical transformations; see, for example [14]. Note also, that according to Eq. (5.33), the invariance of I_1 implies $I_1(\tau_0) = I_1(\tau')$. That is, I_1 does not depend on time.

The Poincaré integral invariant has an interesting geometric interpretation. To discuss this interpretation, let us recall that the following line integral

$$A = \oint_D pdq, \tag{5.37}$$

on the two-dimensional plane parameterized by q and p, gives an area of the region limited by the closed contour D. In the extended phase space, let us consider the contour C_1: $q^1 = q^1(\tau_0, \alpha)$, $p^1 = p^1(\tau_0, \alpha)$, $z^\alpha = 0$, $\tau = 0$. This is the projection of the integration contour C: $z^i = z^i(\tau_0, \alpha)$, $\tau = \tau_0$ of the Poincaré integral invariant (5.32) on the (q^1, p_1)-plane (see Fig. 5.5 on page 163). According to Eq. (5.37), the area inside C_1 can be computed as

$$A_1 = \oint_{C_1} p_1 dq^1 = \int_0^l p_1(\tau_0, \alpha)\frac{\partial q^1(\tau_0, \alpha)}{\partial \alpha} d\alpha. \tag{5.38}$$

Comparing the expressions (5.33) and (5.38), we conclude that the Poincaré integral invariant represents a sum of the areas A_a

$$I_1 = \oint_C p_a dq^a = \sum A_a. \tag{5.39}$$

While the contours C, C_a and their areas can vary during an evolution, the sum (5.39) of the areas A_a, being equal to the invariant I_1, remains unaltered. This gives the geometric interpretation of the Poincaré integral invariant.

Let us enquire about the structure of the universal invariant of the general form (5.10). In other words, we are interested in finding the most general form of the vector field that implies the invariance of the integral.

Assertion Let the line integral $\tilde{I}_1$ of the vector field $V_i(z^j, \tau) = (v_a, u^a)$ be the universal integral invariant. Then

1. The field has the form

$$V_i = \frac{1}{2}cz^j\omega_{ji} + \partial_i\Phi(z^j, \tau), \quad \text{ou} \quad \begin{cases} v^a = \frac{1}{2}cp_a + \frac{\partial\Phi}{\partial q^a}, \\ u_a = -\frac{1}{2}cq^a + \frac{\partial\Phi}{\partial p_a}, \end{cases} \tag{5.40}$$

where ω_{ij} is the symplectic matrix and Φ is a function.

2. The integral $\tilde{I}_1$ is proportional to the Poincaré integral invariant

$$\tilde{I}_1 \equiv \oint_C v_a dq^a + u^b dp_b = cI_1 \equiv c\oint_C p_a dq^a, \quad c = \text{const}. \tag{5.41}$$

This last statement means that the Poincaré integral invariant is essentially a unique universal integral invariant.

Proof Let C: $\tau = const$ be the closed simultaneous contour on the tube (5.5) of solutions to the Hamiltonian system H. Using the parametric equations of the contour $z^i = z^i(\tau, \alpha)$, $\tau = \text{const}$, $\tilde{I}_1$ can be presented as the definite integral

$$\tilde{I}_1(\tau) = \int_0^l \left(v_a(z(\tau, \alpha), \tau) q'^a(\tau, \alpha) + u^a(z(\tau, \alpha) p'_a(\tau, \alpha)\right) d\alpha$$
$$\equiv \int_0^l V_i(z(\tau, \alpha), \tau) z'^i d\alpha \tag{5.42}$$

Owing to the invariance $\tilde{I}_1(\tau) = \tilde{I}_1(\tau')$, we have $\frac{d\tilde{I}_1}{d\tau} = 0$. Direct computation of the derivative gives the definite integral corresponding to the line integral

$$\oint_C G_i(z, \tau) dz^i = 0, \tag{5.43}$$

where

$$G_i \equiv -W_{ij}\omega^{jk}\partial_k H + \frac{\partial V_i}{\partial \tau},$$
$$W_{ij} \equiv \partial_i V_j - \partial_j V_i. \tag{5.44}$$

Since the previous analysis was carried out for an arbitrary tube, the integral (5.43) vanishes for any contour on the hyperplane $\tau = $const, in other words, it is contour-independent. That is, G_i is the conservative field, $\partial_l G_i - \partial_i G_l = 0$. The explicit form of this expression is

$$(\partial_j W_{li})\omega^{jk}\partial_k H + \frac{\partial}{\partial \tau} W_{li} + W_{lj}\omega^{jk}\partial_k\partial_i H - W_{ij}\omega^{jk}\partial_k\partial_l H = 0. \tag{5.45}$$

Due to the universality of $\tilde{I}_1$, Eq. (5.45) is true for any H, therefore

$$\partial_j W_{li} = \frac{\partial}{\partial \tau} W_{li} = 0,$$
$$W_{lj}\omega^{jk}\partial_k\partial_i H - W_{ij}\omega^{jk}\partial_k\partial_l H = 0. \tag{5.46}$$

The first line implies that W is the numeric matrix.

Exercise
Verify that the second line implies $W_{ij} = c\omega_{ij}$, where $c = \text{const}$.

Accordingly, we write $\partial_i V_j - \partial_j V_i = c\omega_{ij}$, or, equivalently $\partial_i(V_j - \frac{1}{2}cz^k\omega_{kj}) - \partial_j(V_i - \frac{1}{2}cz^k\omega_{ki}) = 0$. In turn, it implies $V_i - \frac{1}{2}cz^k\omega_{ki} = \partial_i\Phi$. That is, V_i has the form $V_i = \frac{1}{2}cz^k\omega_{ki} + \partial_i\Phi$, as has been stated.

Integration of the vector field along a contour gives $\tilde{I}_1 = \oint_C V_i dz^i = \frac{1}{2}c\oint_C p_a dq^a - q^a dp_a + \oint_C \partial_i \Phi dz^i = c\oint_C p_a dq^a$. All the integrated terms vanish due to the closeness of the contour. Thus, any universal integral invariant differs from the Poincaré one only by a numeric factor.

Chapter 6
Potential Motion in a Geometric Setting

6.1 Analysis of Trajectories and the Principle of Maupertuis

The Maupertuis variational principle is the oldest least-action principle of classical mechanics. Its precise formulation was given by Euler and Lagrange; for its history, see [34]. However, the traditional formulation (as a variational problem subject to the constraint that only the motions with fixed total energy are considered), remained problematic, as emphasized by V. Arnold (double citation): "In his Lectures on Dynamics (1842–1843), C. Jacobi commented: "In almost all textbooks, even the best, this Principle is presented in such a way that it is impossible to understand". I do not choose to break with tradition" [2].

In this section we present the principle of Maupertuis as an unconstrained variational problem. We discuss a conservative Lagrangian system. Its equations of motion can be obtained according to the least action principle. In a full analogy, the *principle of Maupertuis* can be formulated as the least action principle leading to the equations for *trajectories* of the system.

In greater detail, the strategy of this section is as follows.

Given the solution $q^a = q^a(\tau)$ of the Lagrangian equations, we can exclude the time τ, thus obtaining the functions $q^\alpha(q^1)$ describing a trajectory of motion. We are interested in an analysis of the trajectories. It is possible to write a system of differential equations describing the trajectories. The system shows a number of very interesting properties, that form in fact the contents of the principle of Maupertuis. On the phase space, equations for the trajectories form a Hamiltonian system (with q^1 playing the role of an "evolution parameter"). So, we are able to find the corresponding Hamiltonian action. Equations for trajectories can be obtained as stationarity conditions of the Hamiltonian action. This variational problem for trajectories is precisely the principle of Maupertuis. Further, from the Hamiltonian formulation we can restore a Lagrangian one, thus obtaining the Lagrangian version of the principle of Maupertuis. The Hamiltonian and Lagrangian for the trajectories $q^\alpha(q^1)$ will be found below in terms of the initial Hamiltonian and Lagrangian for $q^a(\tau)$.

This formalism will be further applied to describe a conservative mechanical system in geometric terms.

A. Deriglazov, *Classical Mechanics*, DOI 10.1007/978-3-642-14037-2_6,

6.1.1 Trajectory: Separation of Kinematics from Dynamics

Let $q^a = q^a(\tau)$ denote any solution to the Lagrangian equations (2.16), which we write here in the normal form

$$\ddot{q}^a = K^a(q^a, \dot{q}^a), \qquad K^a \equiv \tilde{M}^{ab} K_b. \tag{6.1}$$

The Lagrangian has no manifest time-dependence.

Geometrically, the solution $q^a = q^a(\tau)$ is a set of points in the configuration space together with a given parametrization. The set itself (which is an image of the interval $[\tau_1, \tau_2]$ in the configuration space) is called a *trajectory*. At least locally, it can be described without mentioning the evolution parameter. Suppose that one of the functions describing the solution, say $q^1 = q^1(\tau)$, can be resolved with respect to τ: $\tau = \tau(q^1)$. The substitution of this function into the remaining ones gives expressions determining the *trajectory*

$$q^\alpha(q^1) \equiv q^\alpha(\tau(q^1)), \qquad \alpha = 2, 3, \ldots, n. \tag{6.2}$$

By construction, we have the following representation of the functions $q^\alpha(\tau)$ in terms of $q^\alpha(q^1)$ defined by Eq. (6.2)

$$q^\alpha(\tau) = q^\alpha(q^1)\Big|_{q^1 \to q^1(\tau)} \equiv q^\alpha(q^1)\Big|. \tag{6.3}$$

Besides the notation $\dot{q}^a \equiv \frac{dq^a(\tau)}{d\tau}$, in this section we will use also the notation $q^{,a} \equiv \frac{dq^a(q^1)}{dq^1}$. From Eqs. (6.3) and (6.1) we can write

$$\dot{q}^\alpha = q^{,\alpha}|\dot{q}^1, \qquad \ddot{q}^\alpha = q^{,,\alpha}|(\dot{q}^1)^2 + q^{,\alpha}|K^1(q^a, \dot{q}^1, q^{,\alpha}\dot{q}^1). \tag{6.4}$$

Starting from the system (6.1), it is possible to write closed equations determining the trajectory (6.2). This possibility is due to the fact that the order of any system admitting a first integral can be reduced by two units. Let us describe the procedure for the case of the total energy first integral. The energy conservation law

$$\frac{\partial L}{\partial \dot{q}^a}\dot{q}^a - L = h = \text{const}, \tag{6.5}$$

being a consequence of the system, can be added to this. Using the first equation from (6.4), we write Eq. (6.5) in terms of $q^{,\alpha}|$, $\dot{q}^1$ and resolve it algebraically for the symbol $\dot{q}^1$:

$$\dot{q}^1 = \dot{q}^1(q^1, q^\alpha, q^{,\alpha}, h)|. \tag{6.6}$$

Replacement of $\dot{q}^a$, $\ddot{q}^\alpha$ in the equations $\ddot{q}^\alpha = K^\alpha$ according to Eqs. (6.4) and (6.6), with subsequent substitution of the function $\tau(q^1)$ in place of τ, gives the desired equations for variables $q^\alpha(q^1)$ that do not contain the variable τ

$$q^{,,\alpha}(q^1) = \frac{1}{(\dot{q}^1)^2}\left(K^\alpha - q^{,\alpha}K^1\right). \tag{6.7}$$

Solving these equations, we substitute the result into Eq. (6.6), which leads to a closed equation for the function $\tau(q^1)$

$$\frac{d\tau}{dq^1} = \frac{1}{\dot{q}^1(q^1, h)}. \tag{6.8}$$

This can be immediately integrated out, giving the time interval for the transition from q_1^a to q_2^a

$$\triangle\tau = \int_{q_1^1}^{q_2^1} \frac{dq^1}{\dot{q}^1(q^1, h)}. \tag{6.9}$$

It determines time evolution of the configuration-space particle $q^a(\tau)$.

In short, for the time-independent Lagrangian, the Lagrangian equations admit a *separation of variables*: we are able to write a closed system of $n-1$ equations for the trajectory $q^\alpha(q^1)$. charge.

Let us specify these results for the case of a particle moving in a potential. From the action

$$S = \int d\tau \left(\frac{1}{2}(\dot{q}^a)^2 - U(q^a)\right), \tag{6.10}$$

we have the equations of motion

$$\ddot{q}^a + \partial_a U = 0. \tag{6.11}$$

Eqs. (6.5) and (6.6) acquire the form

$$(\dot{q}^1)^2\frac{1}{2}(q^{,a})^2 + U(q) = h \quad \Rightarrow \quad \frac{1}{\dot{q}^1} = \frac{d\tau}{dq^1} = \sqrt{\frac{(q^{,a})^2}{2(h-U)}}. \tag{6.12}$$

By using Eq. (6.4), we obtain equations for the trajectory of the particle

$$q^{,,\alpha} + \frac{1+(q^{,\beta})^2}{2(h-U)}\left(\frac{\partial U}{\partial q^\alpha} - q^{,\alpha}\frac{\partial U}{\partial q^1}\right) = 0. \tag{6.13}$$

Equations (6.11) are equivalent to the system (6.12) and (6.13). In this system the description of the dynamics is separated from the description of the trajectory (that is, of the kinematics).

The solutions $q^a(\tau)$ give an extremum of the variational problem (2.34). In the next sections we will show that the trajectories $q^\alpha(q^1)$ can also be obtained from the unconstrained variational problem known as the *principle of Maupertuis*.

6.1.2 Equations for Trajectory in the Hamiltonian Formulation

The previous discussion can be repeated in the Hamiltonian formulation, leading to the conclusion that equations for the trajectory $q^\alpha(q^1)$, $p_\alpha(q^1)$, $\alpha = 2, 3, \ldots, n$ form a Hamiltonian system. We find here the corresponding Hamiltonian.

Let $H(q^a, p_a)$ stand for a Hamiltonian of the system $L(q, \dot{q})$. Conservation of energy, being a consequence of Hamiltonian equations, can be added to this:

$$\dot{z}^a = \omega^{ab}\frac{\partial H}{\partial z^b}, \tag{6.14}$$

$$H(q^1, q^\alpha, p_1, p_\alpha) = h. \tag{6.15}$$

Let us resolve the Eq. (6.15) algebraically with respect to p_1. The solution is given as

$$p_1 = -\tilde{H}(q^1, q^\alpha, p_\alpha, h). \tag{6.16}$$

By construction, we have the identity:

$$H(q^1, q^\alpha, -\tilde{H}(q^1, q^\alpha, p_\alpha, h), p_\alpha, h) \equiv h, \tag{6.17}$$

which implies:

$$\begin{aligned} \frac{\partial \tilde{H}}{\partial q^\alpha} &= \left(\left.\frac{\partial H}{\partial p_1}\right|_{p_1=-\tilde{H}}\right)^{-1} \left.\frac{\partial H}{\partial q^\alpha}\right|_{p_1=-\tilde{H}}, \\ \frac{\partial \tilde{H}}{\partial p_\alpha} &= \left(\left.\frac{\partial H}{\partial p_1}\right|_{p_1=-\tilde{H}}\right)^{-1} \left.\frac{\partial H}{\partial p_\alpha}\right|_{p_1=-\tilde{H}}. \end{aligned} \tag{6.18}$$

Exercise
Using the above identities, show that the equation $\dot{p}_1 = -\frac{\partial H}{\partial q^1}$ is a consequence of other equations of the system (6.14) and (6.15) and thus can be omitted.

Similarly to the previous section, let $z^a(\tau) = (q^a(\tau), p_a(\tau))$ be a solution to Eq. (6.14), and $z^\alpha(q^1)$, $p_1(q^1)$ be the corresponding phase-space trajectory. Then Eq. (6.14) implies

$$\dot{q}^\alpha = q'^\alpha \dot{q}^1 = q'^\alpha \frac{\partial H}{\partial p_1}, \qquad \dot{p}_\alpha = p'_\alpha \dot{q}^1 = p'_\alpha \frac{\partial H}{\partial p_1}. \tag{6.19}$$

This allows us to write equations for $z^\alpha(q^1)$ (which are known as *Whittaker equations*). They arise from Eq. (6.14) by using Eqs. (6.19), (6.16) and (6.18)

$$\frac{dq^\alpha}{dq^1} = \frac{\partial \tilde{H}}{\partial p_\alpha}, \qquad \frac{dp_\alpha}{dq^1} = -\frac{\partial \tilde{H}}{\partial q^\alpha}. \tag{6.20}$$

Together with the equations

$$p_1 = -\tilde{H}(q^1, q^\alpha, p_\alpha, h), \qquad \dot{q}^1 = \frac{\partial H}{\partial p_1}, \tag{6.21}$$

they form an equivalent to the Eq. (6.14) system. So, the Eqs. (6.20) give trajectories of (2.16) with the fixed energy h. Notice that the Eqs. (6.20) form a Hamiltonian system, with the Hamiltonian $\tilde{H}$ obtained as a solution to Eq. (6.17). Integrating $2n-2$ Whittaker's equations (6.20), one substitutes the result into the first equation from (6.21), thus obtaining the expressions $q^\alpha(q^1)$, $p_a(q^1)$ for the phase-space trajectory. Substitution of these functions into the second equation from (6.21) gives the equation

$$\frac{d\tau}{dq^1} = \left(\frac{\partial H}{\partial p_1}\right)^{-1}, \tag{6.22}$$

for the function $\tau(q^1)$, which can immediately be integrated out.

6.1.3 The Principle of Maupertuis for Trajectories

As we have seen in the previous section, the trajectory $q^\alpha(q^1)$, $p_\alpha(q^1)$, obeys the Hamiltonian equations (6.20). According to Sect. 2.9, the equations can be obtained from the variational problem for the Hamiltonian action

$$\begin{aligned} \tilde{S}_{\tilde{H}} &= \int dq^1 \left[p_\alpha q^{\cdot\alpha} - \tilde{H}(q^\alpha, p_\alpha, q^1, h) \right] \\ &\equiv \int p_\alpha dq^\alpha - \tilde{H} dq^1 \equiv \int p_a dq^a. \end{aligned} \tag{6.23}$$

Here $\tilde{H}$ is constructed according to Eqs. (6.15) and (6.16). The problem is known as the *principle of Maupertuis* (the notation $p_a dq^a$ is due to (6.16)). This states that among all the phase-space trajectories with a given energy $H(q^a, p_a) = h$, a system follows the trajectory that supplies an extremum to the functional (6.23). Notice that the construction of $\tilde{H}$, which enters into Eq. (6.23), implies knowledge of the initial system Hamiltonian.

Due to the Hamiltonian character of equations for trajectory, in any particular example one can restore the corresponding Lagrangian $\tilde{L}(q^\alpha, q^{\cdot\alpha}, q^1, h)$) by applying the procedure described in Sect. 2.1.4. We now discuss how this $\tilde{L}$ can be constructed from the initial Lagrangian L in the general case[1].

6.1.4 Lagrangian Action for Trajectories

Consider the action with the time-independent Lagrangian

$$S = \int d\tau L(q^a, \dot{q}^a). \tag{6.24}$$

Construction Let us write the energy conservation law in terms of $q^{\cdot\alpha}$, $\dot{q}^1$, see Eq. (6.4)

$$\left(\frac{\partial L}{\partial \dot{q}^1} + \frac{\partial L}{\partial \dot{q}^\alpha} q^{\cdot\alpha}\right) \dot{q}^1 - L(q^a, \dot{q}^1, q^{\cdot\alpha}\dot{q}^1) = h. \tag{6.25}$$

Resolving this equation algebraically for the symbol $\dot{q}^1$

$$\dot{q}^1 = \dot{q}^1(q^a, q^{\cdot\alpha}, h) \equiv \dot{q}^1(q^{\cdot\alpha}), \tag{6.26}$$

we construct the following action for the variables $q^\alpha(q^1)$:

$$\begin{aligned} \tilde{S} &= \int dq^1 \frac{1}{\dot{q}^1}(L(q^a, \dot{q}^a) + h)|_{\dot{q}^a \to q^{\cdot a}\dot{q}^1}|_{\dot{q}^1 = \dot{q}^1(q^{\cdot\alpha})} \\ &\equiv \int dq^1 \tilde{L}(q^\alpha, q^{\cdot\alpha}, q^1, h) \end{aligned} \tag{6.27}$$

In the rest of this section, the symbol | denotes the double substitution indicated in this equation.

We demonstrate now that the trajectories $q^\alpha(q^1, h)$ of the theory $\tilde{L}$ with a given energy h can be obtained as solutions to the variational problem (6.27). For this purpose, it is sufficient to confirm that a Hamiltonian of the theory $\tilde{L}$ obeys the Eq. (6.17); see the discussion after Eq. (6.21).

Consider the Hamiltonian formulation for the theory (6.27). We write equations for the momenta and the corresponding solutions: $\tilde{p}_\alpha = \frac{\partial \tilde{L}}{\partial q^{\cdot\alpha}} \Rightarrow q^{\cdot\alpha} = \tilde{v}^\alpha(q^a, \tilde{p}_\alpha)$. Using Eqs. (6.27) and (6.25), the derivative of $\tilde{L}$ can be written in terms of L as follows: $\frac{\partial \tilde{L}}{\partial q^{\cdot\alpha}} \equiv \left.\frac{\partial L}{\partial \dot{q}^\alpha}\right|$. Combining these equations, the momenta can be presented in terms of L

[1] Equation (6.7) can also be taken as a starting point for obtaining $\tilde{L}$.

$$\frac{\partial L}{\partial \dot{q}^\alpha}\bigg|\bigg|_{q^{\cdot}=\tilde{v}} \equiv \tilde{p}_\alpha. \tag{6.28}$$

The Hamiltonian of the system (6.27) can also be presented in terms of L:

$$\begin{aligned}\tilde{H}(q^a, \tilde{p}_\alpha, h) &= (\tilde{p}_\alpha q^{\cdot\alpha} - \tilde{L})\Big|_{q^{\cdot}=\tilde{v}} \\ &= -\frac{\partial L}{\partial \dot{q}^1}\bigg|\bigg|_{q^{\cdot}=\tilde{v}}.\end{aligned} \tag{6.29}$$

The Hamiltonian corresponding to L is given by:

$$H(q^a, p_a) = p_a v^a - L(q^a, v^a). \tag{6.30}$$

Here $v^a(q^b, p_b)$ represents a solution to the equation $p_a = \frac{\partial L}{\partial \dot{q}^a}$, so:

$$v^a\left(q^b, \frac{\partial L}{\partial \dot{q}^b}\right) \equiv \dot{q}^a. \tag{6.31}$$

Now, let us substitute the function $-\tilde{H}(\tilde{p}_\alpha)$ into the expression for H in place of p_1. The substitution reads

$$\frac{\partial L}{\partial \dot{q}^1}\bigg| v^1\left(\frac{\partial L}{\partial \dot{q}^1}\bigg|, p_\alpha\right) + p_\alpha v^\alpha\left(\frac{\partial L}{\partial \dot{q}^1}\bigg|, p_\alpha\right) - L\left(q^a, v^a\left(\frac{\partial L}{\partial \dot{q}^1}\bigg|, p_\alpha\right)\right). \tag{6.32}$$

To see that it is equal to the constant h, we express the symbol p_α according to the identity (6.28) and use Eq. (6.31), thus obtaining:

$$(\frac{\partial L}{\partial \dot{q}^a}\dot{q}^a - L(q^a, \dot{q}^a))|\,|_{q^{\cdot}=\tilde{v}}. \tag{6.33}$$

From Eqs. (6.25) and (6.26), the expression in brackets is equal to h. Thus we have obtained the Lagrangian version of the *principle of Maupertuis*: trajectories $q^\alpha(q^1, h)$ of the system $L(q^a(\tau), \dot{q}^a(\tau))$ with a given energy h can be obtained as solutions to the variational problem (6.27).

In particular, application of the Eqs. (6.25), (6.26) and (6.27) to a particle moving in a potential, see Eq. (6.10), gives the following action for the trajectory

$$\begin{aligned}\tilde{S}(q^\alpha) = \int dq^1\sqrt{2(h-U)\delta_{ab}q^{\cdot a}q^{\cdot b}} = \\ \int dq^1\sqrt{2(h-U)(1+(q^{\cdot\alpha})^2)}.\end{aligned} \tag{6.34}$$

Corresponding equations of the trajectory can be presented in the form

$$q^{\prime\prime\alpha} + \hat{\Gamma}^{\alpha}{}_{bc} q^{\prime b} q^{\prime c} = 0, \tag{6.35}$$

where

$$\hat{\Gamma}^{\alpha}{}_{bc} = \Gamma^{\alpha}{}_{bc} - q^{\prime\alpha} \Gamma^{1}{}_{bc},$$
$$\Gamma^{a}{}_{bc} = -\frac{1}{2(h-U)} (\delta^{a}{}_{c} \partial_b U + \delta^{a}{}_{b} \partial_c U - \delta^{b}{}_{c} \partial_a U). \tag{6.36}$$

This form of equation will be useful in the next section. Substitution of the coefficients $\Gamma^{a}{}_{bc}$ into Eq. (6.35) gives (6.13).

Exercise
Obtain the Eq. (6.35) from the variational problem.

To sum up, we have discussed trajectories of the conservative Lagrangian system. Equations for the trajectories (6.7) also form the Lagrangian system, so they can be obtained from a variational problem with an appropriately chosen Lagrangian, which has been constructed in a closed form in terms of the initial Lagrangian, see Eq. (6.27).

The principle of Maupertuis represents the Hamiltonian version of this variational problem: phase-space equations for trajectories form the Hamiltonian system (6.20), and hence can be obtained as stationarity conditions of the Hamiltonian action functional (6.23).

6.2 Description of a Potential Motion in Terms of a Pair of Riemann Spaces

The trajectory of a free moving particle is a straight line. From the geometric point of view, straight lines in Euclidean space represent a very special class, being the shortest lines between two points. So, one can say in a geometric interpretation of free motion, that among all the trajectories, the particle chooses the shortest one. Generally, the trajectory is not a straight line for a potential system. Nevertheless, there is an interesting possibility of a similar geometric treatment for that case.

In differential geometry there is a class of so called *Riemann spaces* with the metrical and parallel transport properties generally different from those of the Euclidean space. Instead of the standard expression for infinitesimal distance: $ds = \sqrt{\delta_{ab} dq^a dq^b}$, in the Riemann space we have the expression $ds = \sqrt{g_{ab} dq^a dq^b}$. The metrical properties are determined by the set of functions $g_{ab}(q^a)$ known as the *metric* of the space. Since the metric depends on q^a, the metrical properties change from one point to another as well as differing in various directions from a given point. The analogy of a straight line (that is, the analogy of free motion) in the Riemann space is a *geodesic line* representing the curve of a minimal length between two points.[2] Parallel transport is defined by an independent quantity called the affine connection.

From the previous section we know that the description of a trajectory of motion can be separated from the description of the dynamics along it; see Eqs. (6.12) and (6.13) for the case of potential motion. Here we show that the configuration space of the system can be equipped with a metric (constructed with the help of the potential) in such a way that trajectories turn out to be the shortest lines of this metric. On the kinematic level, potential motion in flat Euclidean space is thus equivalent to free motion in curved Riemann space. Intuitively, presence of a potential can be treated as leading to the deformation of the metrical properties of the initially flat configuration space. This in turn causes deviation of the trajectory from the straight line, keeping it the shortest line with respect to the metric. Moreover, the dynamics (time of evolution) also has an invariant geometric meaning. Potential motion in this framework looks like an inertial motion, in a full (but formal) analogy with the inertial motion in the presence of gravity in Einstein's general theory of relativity.

Here we give only an elementary discussion of the problem. The subsequent sections are devoted to a more detailed discussion. For our present purposes it is sufficient to bear in mind that the shortest line in Riemann space can be described in the parametric form, $q^a(\tau)$, by so-called *geodesic equations in canonical parametrization* (for an elementary demonstration of this fact see a sequence of exercises at the end of this section)

$$\ddot{q}^a + \Gamma^a{}_{bc}\dot{q}^b\dot{q}^c = 0, \tag{6.37}$$

subject to the conditions

$$q^a(\tau_1) = q_1^a, \qquad q^a(\tau_2) = q_2^a. \tag{6.38}$$

The conditions mean that we look for the path from the point q_1^a to q_2^a. The functions $\Gamma^a{}_{bc}$ are constructed from the metric as follows

$$\Gamma^a{}_{bc} = \frac{1}{2}g^{ad}(\partial_b g_{dc} + \partial_c g_{bd} - \partial_d g_{bc}). \tag{6.39}$$

These are known as a *Riemann connection* or *Christoffel symbols*.

Now we proceed as in Sect. 6.1.1, rewriting the system (6.37) in an equivalent form, with the kinematics separated from the dynamics.

[2] More exactly, the notions of a geodesic and a minimal length line coincide only in the Riemann space with a Riemann connection. Here we do not distinguish these notions. They are discussed in more detail in the following sections.

Exercise
Verify the conservation of the charge:

$$\frac{d}{d\tau}\left(g_{ab}\dot{q}^a\dot{q}^b\right) = 0, \tag{6.40}$$

for solutions to the problem (6.37).

For solutions with the charge equal to v^2 this implies

$$v\frac{d\tau}{dq^1} = \sqrt{g_{ab}q'^a q'^b}. \tag{6.41}$$

Using Eqs. (6.37), (6.4) and (6.41), we obtain the following equation for the trajectory $q^\alpha(q^1)$:

$$q''^\alpha + \hat{\Gamma}^\alpha{}_{bc}q'^b q'^c = 0, \qquad \hat{\Gamma}^\alpha{}_{bc} \equiv \Gamma^\alpha{}_{bc} - q'^\alpha\Gamma^1{}_{bc}. \tag{6.42}$$

Under the boundary conditions $q^\alpha(q_1^1) = q_1^\alpha$, $q^\alpha(q_2^1) = q_2^\alpha$, it has a unique solution $q^\alpha(q^1)$.

Now we are ready to compare these equations with those of potential motion (6.12) and (6.35), which we write here once again

$$\frac{d\tau}{dq^1} = \sqrt{\frac{\delta_{ab}}{2(h-U)}q'^a q'^b}, \tag{6.43}$$

$$q''^\alpha + \hat{\Gamma}^\alpha{}_{bc}q'^b q'^c = 0. \tag{6.44}$$

The coefficients $\hat{\Gamma}$ are given in terms of the potential U by Eq. (6.36). It is sufficient to specify the metric as follows

$$g_{ab} = 2(h-U)\delta_{ab}. \tag{6.45}$$

Then the geodesic Eqs. (6.42) coincide with Eq. (6.44) for the potential motion with total energy h. That is, the trajectory of potential motion represents the shortest line of this metric. Complete equations are not equivalent, as is clear from comparison of Eqs. (6.41) and (6.43). So, an interesting task would be to find the equations of geometric origin equivalent to the complete problem (6.11). This will be done in Sect. 6.8. Here we only point out that the dynamics (6.43) also has an invariant geometric meaning. Let us write the solution $q^\alpha(q^1)$ in the parametric form $q^1 = \tau$, $q^\alpha = q^\alpha(\tau)$, and compute a length of the curve with respect to the metric

$$G_{ab} = \frac{\delta_{ab}}{2(h-U)}, \tag{6.46}$$

constructed with help of an inverse potential. Comparing an expression for the length

$$l = \int_{\tau=q_1^1}^{\tau=q_2^1} d\tau \sqrt{G_{ab}\dot{q}^a\dot{q}^b} = \int_{q_1^1}^{q_2^1} dq^1 \sqrt{\frac{\delta_{ab}}{2(h-U)} q^{,a} q^{,b}}, \tag{6.47}$$

with the Eq. (6.43), one concludes that the propagation time can be identified with the length of the trajectory in space with the metric (6.46).

In this way, one can speak of a formal geometrization of mechanics: with a system of total energy h propagating in a given potential $U(q^a)$, we associate a pair of Riemann spaces. The influence of the potential on the motion is encoded in the metrical properties of these spaces. The trajectory of the potential motion in the Euclidean configuration space coincides with the shortest line in Riemann space $g_{ab} = 2(h-U)\delta_{ab}$, while the time of motion coincides with the length of this line in Riemann space $G_{ab} = \frac{\delta_{ab}}{2(h-U)}$.

This geometrization has been called formal since the metric has no relationship with the physical space-time metric that appears in the general relativity for the description of gravity. Besides, the metric associated with a given system depends on its total energy. That is, the configuration-space particles with different total energy probe different geometries.

Exercises

1. Verify that we have a similar situation in a more general case of the action

$$S = \int d\tau (\frac{1}{2} c_{ab}(q)\dot{q}^a\dot{q}^b - U(q^a)). \tag{6.48}$$

Here the corresponding Riemann spaces are $g_{ab} = 2(h-U)c_{ab}$ and $G_{ab} = \frac{c_{ab}}{2(h-U)}$.

2. Show that the geodesic equations in canonical parametrization (6.37) follow from the action functional

$$S = \int d\tau \frac{1}{2} g_{ab}\dot{q}^a\dot{q}^b. \tag{6.49}$$

3. Show that the corresponding action (6.27) for trajectories is given by

$$\tilde{S} = \int dq^1 \sqrt{g_{ab} q^{,a} q^{,b}} \tag{6.50}$$

Note a geometric interpretation of $\tilde{S}$: since its integrand $\sqrt{g_{ab}q'^a q'^b}$ gives a distance between nearby closed points of the line $q^\alpha(q^1)$, $\tilde{S}$ itself represents a length of the line. Therefore, solving the variational problem (6.27), we look for the shortest path.

4. Show that the Eqs. (6.42) for a trajectory follow from this action functional.

6.3 Some Notions of Riemann Geometry

Here we briefly describe some basic notions of differential geometry of Riemann space, with the aim of giving a more systematic description of a potential motion as free motion in Riemann space. Besides, our purpose is to discuss certain concepts such as covariance and coordinate independence, which are important for a proper understanding of modern physical theories, including gravity and gauge fields.

6.3.1 Riemann Space

Without going into details,[3] an n-dimensional manifold $\mathbb{M}$ is a set equipped with a *(local) coordinate system* in the vicinity of any point $q_0 \in \mathbb{M}$; that is, a continuous bijective map of this vicinity on a neighborhood of zero of $\mathbb{R}^n$: $\varphi\colon q \rightarrow q^a = \varphi^a(q)$, $a = 1, 2, \ldots, n$, is established. We immediately note that the coordinate system is not unique: if $f^a(q^b)$, with $\det \frac{\partial f^a}{\partial q^b} \neq 0$, are given functions, then the map $\psi \equiv f \circ \varphi\colon q \rightarrow q'^a = f^a(\varphi^b(q))$, also represents a coordinate system. In abuse of notation, we write $f^a(q^b) \equiv q'^a(q^b)$. The transition from one description to another

$$
q^a \rightarrow q'^a = q'^a(q^b), \tag{6.51}
$$

is called a *transformation of coordinates*. The function $q'^a(q^b)$ is known as a *transition function*. For the given coordinate systems φ, ψ, the transition function is defined by $f \equiv \psi \circ \varphi^{-1}$, see Fig. 6.1 on page 179. There is no preferable coordinate system on $\mathbb{M}$; all of them are considered on equal footing. Accordingly, a well-defined construction on $\mathbb{M}$ is one that is defined in relation to each coordinate system or, in other words, it must be specified in all the systems simultaneously. In this section, it is taken for granted that the reader is familiar with the definition and basic properties of tensor fields on $\mathbb{M}$; see Chap. 4 in [15]. For later use, we recall the definition of a *contravariant vector field* or $(1, 0)$*-rank tensor*). It is said that the contravariant vector field $\xi(q)$ is defined on $\mathbb{M}$, if in any coordinate system q^a the set of functions $\xi^a(q^b)$ is given, such that the coordinate transformation (6.51) implies

---

[3] Detailed discussion of the coordinate formulation of Riemann geometry for non-mathematicians can be found, for example, in [15]. For the coordinate-free formulation, see [35].

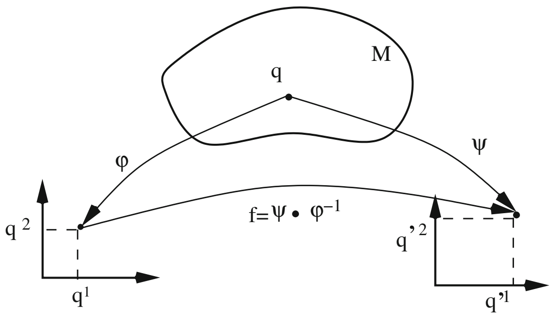

Fig. 6.1 Local coordinate systems and the transition function

$$\xi'^a(q'^c) = \frac{\partial q'^a}{\partial q^b}\xi^b(q^c). \tag{6.52}$$

So, by the vector (tensor) field in the coordinate formulation of differential geometry, we mean a totality of the sets ξ^a, symbolically

$$\xi(q) = \{\xi^a(q^b), \xi'^a(q'^b), \ldots\}. \tag{6.53}$$

For the *covariant vector* (or (0, 1) -rank tensor) one has

$$\xi'_a(q'^c) = \frac{\partial q^b}{\partial q'^a}\xi_b(q^c). \tag{6.54}$$

As compared with a general transformation of the form $\xi'^a(q'^c) = F^a(\xi^b, q^c)$), the tensor transformation law (6.52) is linear and homogeneous with respect to ξ^a. Owing to this fact, tensor fields form a linear space: a linear combination of tensors of the same rank is a tensor. In particular, vector fields form a (infinite-dimensional) linear space, while vectors at a given point q form an n-dimensional linear space $T_q(\mathbb{M})$ known as the *tangent space* to $\mathbb{M}$. Besides, one can define the product and contraction of the tensor fields. In practical computations it is often convenient to use the notation

$$q^{a'} \equiv q'^a, \qquad \xi^{a'} \equiv \xi'^a, \tag{6.55}$$

and so on. This allows us to better control the position of the indices in various equations. In particular, Eq. (6.52) in this notation acquires the form

$$\xi^{a'} = \frac{\partial q^{a'}}{\partial q^a}\xi^a, \tag{6.56}$$

where a and a' stand for different indices.

Example (and exercises)

1. According to Eq. (6.52), if all the components of a vector vanish in a coordinate system, they also vanish in any other system. That is, the set (6.53) of columns composed of zeros is an example of a vector.
2. Consider a *scalar function* defined by the transformation law $h'(q'^a) \equiv h(q^a(q'^b))$. Check that the set $\partial h(q) = \{\partial_a h(q^b), \partial_{a'} h'(q'^b), \ldots\}$ is a (0, 1) -rank tensor.
3. Let us define the set $\triangle q = \{\triangle q^a, \triangle q'^a, \ldots\}$, where $\triangle q^a = q_2^a - q_1^a$ is the difference of coordinates of the points q_1, q_2 in the system q^a, and so on. Show that the set does not transform as a vector. Hence, in contrast to the Euclidean case, the vectors cannot be identified with ordered pairs of points (oriented intervals) of $\mathbb{M}$, and do not belong to $\mathbb{M}$.
4. Consider the vector field $\xi(q)$. Check that the set $\partial\xi = \{\partial_a \xi^b(q^c), \partial_{a'} \xi'^b(q'^c), \ldots\}$ is not a tensor. Hence the partial derivative ∂_a is not an operation defined on the space of tensor fields.
5. Differentiate components of the field $\xi(q)$ in the system q^a: $\partial_a \xi^b \equiv \xi_a^b$. Starting from these functions, define

$$\xi_{a'}^{b'} \equiv \frac{\partial q^a}{\partial q^{a'}} \frac{\partial q^{b'}}{\partial q^b} \xi_a^b. \tag{6.57}$$

in the system q'^a, and so on. By construction, the set $\{\xi_a^b, \xi_{a'}^{b'}, \ldots\}$ determines the (1, 1) rank tensor field. However, it would not be natural to identify it with a derivative of the vector field $\xi(q)$ since if we start from the components $\xi'^b(q'^c)$ instead of $\xi^b(q^c)$ and repeat the construction, we obtain a tensor field which is different from the previous one.

These two examples show that the usual derivative is not a very useful notion on the space of tensors.

Construction Generalizing the last example, it is easy to construct examples of vector (tensor) fields. Starting from the given functions $\xi^a(q^b)$ referring to the system q^a, one defines the functions $\xi'^a(q'^b)$ referring to the system q'^b according to the expression consistent with Eq. (6.52)

$$\xi'^a(q'^c) \equiv \frac{\partial q'^a}{\partial q^b} \xi^b(q^c)\Big|_{q^a(q'^b)}, \tag{6.58}$$

and so on. Below we do not write explicitly the substitution of $q^a(q'^b)$. The resulting set $\xi(q) = \{\xi^a(q^b), \xi'^a(q'^b), \ldots\}$ determines the vector field.

It is worth noting that applying the above construction to the components $\xi^a(q^b)$ of the given vector field $\xi(q)$, we obtain the field itself.

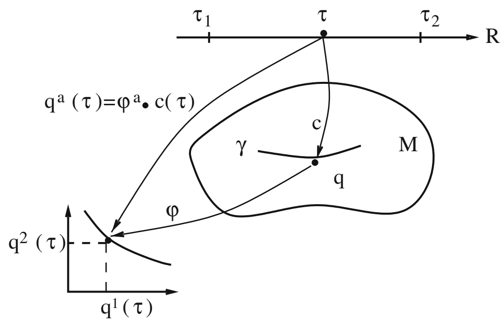

Fig. 6.2 In the local coordinates, the curve $c(\tau)$ is described by n functions $q^a(\tau)$. The line γ is the image of the interval (τ_1, τ_2) applying the map c

Let τ be a variable of the interval $(\tau_1, \tau_2) \in \mathbb{R}$. The *curve* on $\mathbb{M}$ is an injective continuous map[4]

$$c : (\tau_1, \tau_2) \to \mathbb{M}, \qquad \tau \to q = c(\tau). \tag{6.59}$$

If q^a are local coordinates, the map c induces a map $\mathbb{R} \to \mathbb{R}^n$ defined as $q^a = \varphi^a(c(\tau)) \equiv q^a(\tau)$. It gives an analytic description of the curve in the local coordinates in terms of n functions $q^a = q^a(\tau)$; see Fig. 6.2 on page 181. In the system q'^a, we obtain the functions $q'^a = \psi^a(c(\tau)) \equiv q'^a(\tau)$. They are related according to $q'^a(\tau) = \psi^a(c(\tau)) = f^a(\varphi^b(c(\tau))) = f^a(q^b(\tau))$, or, as should be the case,

$$q'^a(\tau) = q'^a(q^b(\tau)), \tag{6.60}$$

where $q'^a(q^b)$ are the transition functions (6.51); see Fig. 6.3 on page 182.

According to the above definition, the curve is a set of points in $\mathbb{M}$ together with a given parametrization. The set itself is called[5] *line* γ. In other words, a set $\gamma \in \mathbb{M}$ is the line, if there is a curve $c(\tau)$ such that γ =Image c. Various curves can determine the same line and sometimes are called its parameterizations. In particular, if the curve $q^a(\tau)$ parameterizes γ and $\tau(\tau')$ is a given function, the curve $q^a(\tau(\tau'))$ also represents a parametrization of γ. It is easy to see that any two parameterizations of γ are related in this way. Now, let the curves $q^a = \varphi^a(c(\tau))$, and $y^a = \varphi^a(c'(\tau'))$ correspond to the same line. Then

$$y^a(\tau') = \varphi^a(c \circ c^{-1} c'(\tau')) = \varphi^a(c(\tau(\tau'))) = y^a(\tau(\tau')), \tag{6.61}$$

where $\tau(\tau') \equiv c^{-1} \circ c'(\tau')$.

[4] Note that this definition does not mention coordinates, representing an example of the coordinate-free definition of differential geometry.

[5] The line was called the trajectory in Sects. 6.1 and 6.2.

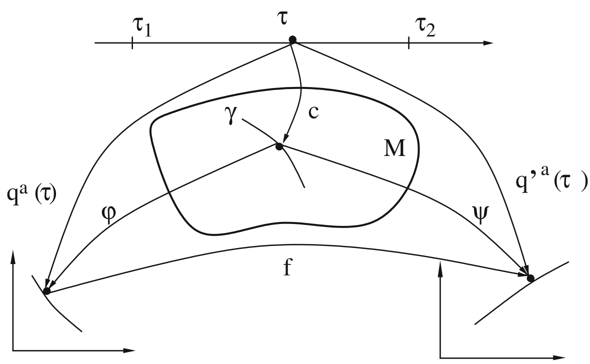

Fig. 6.3 Two parameterizations $q^a(\tau)$ and $q'^a(\tau)$ of the line γ. Since $q = \varphi \circ c$ and $q' = \psi \circ c$, they are related by the transition function f

A curve determines a *tangent vector* ξ to the curve at each point $q^a(\tau)$ according to the rule[6]

$$\xi^a = \frac{dq^a}{d\tau} \equiv \dot{q}^a. \tag{6.62}$$

Exercise
Verify that the components ξ^a transform according to Eq. (6.52).

It can be shown that any vector of the tangent space $T_q(\mathbb{M})$ can be considered as the tangent vector to a curve.

Two basic quantities defined on a manifold are the metric and the affine connection. The *metric* on $\mathbb{M}$ is a symmetric (2, 0) -rank tensor

$$g_{ab}(q) \to g_{a'b'}(q') = \frac{\partial q^a}{\partial q^{a'}} \frac{\partial q^b}{\partial q^{b'}} g_{ab}(q), \tag{6.63}$$

which is non-degenerated: $\det g_{ab} \neq 0$, and positively defined:

$$g_{ab}\xi^a \xi^b \geq 0, \quad \text{for all} \quad \xi \neq 0. \tag{6.64}$$

The inverse tensor is denoted as g^{ab}: $g^{ab} g_{bc} = \delta^a{}_c$. The manifold endowed with the metric is called the *Riemann space*. The metric determines a *scalar product* on the space $T_q(\mathbb{M})$

6 Accordingly, any vector proportional to ξ is called a tangent vector to the line determined by the curve.

$$g(\xi, \eta) = g_{ab}\xi^a \eta^b. \tag{6.65}$$

Exercise
Verify that the scalar product is invariant under the change of coordinates (6.51): $g'(\xi', \eta') = g(\xi, \eta)$.

With the metric in hand, we define the length of a line and, finally, introduce the notion of a distance between points of the Riemann space. We return to this task in Sect. 6.7.

Affine connection[7] on a manifold is a set of functions $\Gamma^a{}_{bc}$, $\Gamma^a{}_{bc} = \Gamma^a{}_{cb}$, given in each coordinate system, with the non-tensor transformation law

$$\begin{aligned}\Gamma^{a'}{}_{b'c'} &= \frac{\partial q^{a'}}{\partial q^a}\frac{\partial q^b}{\partial q^{b'}}\frac{\partial q^c}{\partial q^{c'}}\Gamma^a{}_{bc} - \frac{\partial q^a}{\partial q^{b'}}\frac{\partial q^b}{\partial q^{c'}}\frac{\partial^2 q^{a'}}{\partial q^a \partial q^b} \\ &\equiv \frac{\partial q^{a'}}{\partial q^a}\frac{\partial q^b}{\partial q^{b'}}\frac{\partial q^c}{\partial q^{c'}}\Gamma^a{}_{bc} + \frac{\partial^2 q^a}{\partial q^{b'}\partial q^{c'}}\frac{\partial q^{a'}}{\partial q^a}.\end{aligned} \tag{6.66}$$

The last equality follows from the differentiation of the identity $\delta^{a'}{}_{b'} = \frac{\partial q^a}{\partial q^{b'}}\frac{\partial q^{a'}}{\partial q^a}$ with respect to $q^{c'}$.

6.3.2 Covariant Derivative and Riemann Connection

According to Exercises 4 and 5 of the previous section, the quantity $\partial_a \xi^b$ cannot be considered as a reasonable notion of derivative on tensor space (the same is true for the derivative along the curve $q^a(\tau)$: $\frac{d\xi^a(q(\tau))}{d\tau}$). The proper generalization is as follows. Starting from the vector field $\xi(q)$, let us construct the set

$$D\xi = \{D_b \xi^a, D'_b \xi'^a, \ldots\}, \tag{6.67}$$

where

$$D_b \xi^a \equiv \partial_b \xi^a + \Gamma^a{}_{bc}\xi^c, \tag{6.68}$$

and so on. The set turns out to be a $(1, 1)$-rank tensor, that is Eq. (6.68) defines the map D of $(1, 0)$-rank tensor space in the space of $(1, 1)$-rank tensors. It is called the *covariant derivative* of the vector field ξ. The tensor transformation law of $D_b \xi^a$ is supplied by the non-tensor transformation law (6.66) of the affine connection. Construction of the covariant derivative for an arbitrary rank tensor is clear from the following example

[7] We consider only torsion-free affine connections.

$$D_d A^{ab}{}_c \equiv \partial_d A^{ab}{}_c + \Gamma^a{}_{dk} A^{kb}{}_c + \Gamma^b{}_{dk} A^{ak}{}_c - \Gamma^k{}_{dc} A^{ab}{}_k. \tag{6.69}$$

In particular, for the scalar function $h(q^a)$, the covariant derivative coincides with the usual one (see also Exercise 2 of the previous section)

$$D_a h = \partial_a h. \tag{6.70}$$

Exercise
Show that Eq. (6.69) implies the *Leibnitz rule*, for example $D_a(A^{bc}B_d) = (D_a A^{bc})B_d + A^{bc} D_a B_d$ as well as a commutativity with contractions, for example $D_a(A^{bc}B_c) = (D_a A^{bc})B_c + A^{bc} D_a B_c$.

So, the covariant derivative (6.69) is a map that takes (k, m) -rank tensors to tensors of $(k, m+1)$ -rank, and has the usual properties of a derivative: it is a linear map that obeys the Leibnitz rule. Besides, it commutes with the contractions.

The vector (tensor) field is called a *covariantly-constant field* if it obeys the equation

$$D_b \xi^a = 0. \tag{6.71}$$

A covariantly constant field in Riemann space is an analogy[8] of a constant field in Euclidean space.

Affine connection on a manifold is not unique (any set of functions $\Gamma^a{}_{bc}$, given in the system q^a, can be used to create an affine connection (6.66) using the construction described in the previous section). In Riemann space we can fix the connection from the requirement that it must respect the metrical properties. In Euclidean space the scalar product of the constant fields has the same value at any point; that is, a derivative of the scalar product vanishes: $\partial_c(\xi, \eta) = 0$. So, for the covariantly constant fields in Riemann space it is natural to demand the same condition

$$\partial_c g(\xi, \eta) = (D_c g_{ab})\xi^a \eta^b + g_{ab}(D_c \xi^a)\eta^b + g_{ab}\xi^a D_c \eta^b = 0, \tag{6.72}$$

which is equivalent to the covariant constancy of the metric

$$D_c g_{ab} = 0. \tag{6.73}$$

These equations can be treated as determining Γ in terms of a given metric. They can be resolved algebraically. Equation (6.73) implies

$$D_c g_{ab} + D_a g_{bc} - D_b g_{ca} = 0. \tag{6.74}$$

[8] Parallel transport of the covariantly constant field along any line takes it into itself, see below.

For the case of symmetric affine connection this is equivalent to

$$\Gamma^a{}_{bc}(g) = \frac{1}{2} g^{ad}(\partial_b g_{dc} + \partial_c g_{bd} - \partial_d g_{bc}). \tag{6.75}$$

Exercise
Verify that the connection transforms according to Eq. (6.66).

If the affine connection is not an independent quantity and has been chosen according to Eq. (6.75), it is called a *Riemann connection*. Note that for a given metric, the symmetric Riemann connection is unique.

For a given vector field η^a, the *covariant derivative along the field* is defined as:

$$D_\eta \xi^a \equiv \eta^b D_b \xi^a. \tag{6.76}$$

For a given curve $q^a(\tau)$, the *covariant derivative along the curve* is defined as:

$$D\xi^a \equiv \frac{d\xi^a(q(\tau))}{d\tau} + \Gamma^a{}_{bc}(q(\tau))\dot{q}^b \xi^c(q(\tau)) = \dot{q}^b \left.(D_b \xi^a)\right|_{q(\tau)}. \tag{6.77}$$

6.3.3 Parallel Transport: Notions of Covariance and Coordinate Independence

Vectors of the tangent space $T_q(\mathbb{M})$ form a linear space, hence two vectors can be compared by comparing their coordinates. In the Euclidean case, it is possible to connect tangent spaces at different points introducing a natural notion of parallel transport. This possibility is based on two properties: (a) the vector can be identified with the ordered pair of points of $\mathbb{E}$, (b) the only straight line parallel to a given one passes through a point of $\mathbb{E}$. Since the transport is defined in a unique way, we have the possibility of comparing vectors taken at different points. While Riemann space does not admit such properties, a useful notion of parallel transport along a line can also be established. However, it does not resemble all the properties of the Euclidean case (the transport generally depends on the line; therefore, it does not imply a way to compare tangent spaces at different points).

Parallel transport of the vector ξ_1 given at the point q_1 along the line γ can be defined as follows. Consider the manifold $\mathbb{M}$ with an affine connection. Let γ be a line between q_1 and q_2; c: $(\tau_1, \tau_2) \to \mathbb{M}$ represents the corresponding curve and $q^a(\tau)$ is its expression in local coordinates. Let ξ_1 be a vector at q_1.

Definition The set

$$\xi = \{\xi^a(\tau), \xi'^a(\tau), \ldots\}, \tag{6.78}$$

composed by solutions to the equation

$$D\xi^a \equiv \frac{d\xi^a}{d\tau} + \Gamma^a{}_{bc}\dot{q}^b\xi^c = 0, \tag{6.79}$$

with the initial condition

$$\xi^a(\tau_1) = \xi^a_1, \tag{6.80}$$

determines a vector field along the line.[9] It is called the *parallel transport* of ξ_1. Sometimes we write $D(q(\tau))$ instead of D to emphasize that the problem (6.79) is formulated for a particular parametrization $q^a(\tau)$ of the line.

The vector field $\xi^a(q)$, given along the line $q^a(\tau)$, is called *parallel* if it obeys the Eq. (6.79). In this case the parallel transport of the vector $\xi^a(q_1)$ to the point q_2 gives the vector that coincides with $\xi^a(q_2)$.

We need to verify the consistency of the definition: whether the set (6.78) really does give the vector field, as well as its independence from the parametrization implied in the definition.

Let us confirm that the set (6.78) actually determines a vector field. Let the functions $\xi^a(\tau)$ obey the problem (6.79), (6.80) in the coordinates q^a (note that the problem has a unique solution since the Eqs. (6.79) form the normal system), while $\xi'^a(\tau)$ obey the problem

$$D'\xi'^a = \frac{d\xi'^a}{d\tau} + \Gamma'^a{}_{bc}\dot{q}'^b\xi'^c = 0, \qquad \xi'^a(\tau_1) = \xi'^a_1, \tag{6.81}$$

in the coordinates $q'^a = q'^a(q^b)$. Using Eqs. (6.60) and (6.66), the parallel transport Eq. (6.79) can be *identically* rewritten in the form

$$0 = D\xi^a = \frac{\partial q^a}{\partial q'^b} D' \left(\frac{\partial q'^b}{\partial q^c}\xi^c \right), \tag{6.82}$$

where D' is just the covariant derivative (6.81) in the primed system. Since $\det \frac{\partial q^a}{\partial q'^b} \neq 0$, the equation $D\xi = 0$ turns out to be equivalent to $D'\left(\frac{\partial q'}{\partial q}\xi\right) = 0$. That is, if ξ obeys (6.79), the quantity $\frac{\partial q'}{\partial q}\xi$ obeys Eq. (6.81). Since its solution is unique, one concludes

$$\frac{\partial q'^b}{\partial q^c}\xi^c = \xi'^b. \tag{6.83}$$

Hence the problem (6.79) and (6.80) actually determines the vector field.

[9] Components $\xi^a(q^b)$ at the point $q^b = q^b(\tau)$ are defined as $\xi^a(q^b) \equiv \xi^a(\tau)$.

Reparametrization independence. Consider the solutions $\xi^a(\tau)$ and $\eta^a(\tau')$ of the problem (6.79) and (6.80) in the parameterizations $q^a(\tau)$ and $y^a(\tau')$

$$D(q(\tau))\xi^a = 0 \Rightarrow \xi^a(\tau), \tag{6.84}$$
$$D(y(\tau'))\eta^a = 0 \Rightarrow \eta^a(\tau'). \tag{6.85}$$

We show that the corresponding vector fields ξ and η coincide. Let the point q_0 corresponds to values of the parameters beingτ_0 and τ'_0. We have $\tau_0 = \tau(\tau'_0)$, where the function $\tau(\tau')$ has been defined in Eq. (6.61). We need to show that the fields ξ and η coincide at q_0: $\xi^a(\tau_0) = \eta^a(\tau'_0)$. Starting from the given functions $\xi^a(\tau)$, $\tau(\tau')$, let us construct the following function of τ': $\xi^a(\tau') \equiv \xi^a(\tau(\tau'))$. It obeys the Eq. (6.85)

$$\begin{aligned}\frac{d\xi^a(\tau')}{d\tau'} &= \left.\frac{d\xi^a(\tau)}{d\tau}\right|_{\tau(\tau')}\frac{d\tau}{d\tau'} = -\,\Gamma^a{}_{bc}\left.\frac{dq^b}{d\tau}\xi^c(\tau)\right|_{\tau(\tau')}\frac{d\tau}{d\tau'}\\ &= -\Gamma^a{}_{bc}\frac{dq^b(\tau(\tau'))}{d\tau'}\xi^c(\tau(\tau')) = -\Gamma^a{}_{bc}\frac{dy^b(\tau')}{d\tau'}\xi^c(\tau').\end{aligned} \tag{6.86}$$

Here Eqs. (6.84) and (6.61) have been used. Since the problem (6.85) has a unique solution, we conclude

$$\eta^a(\tau') = \xi^a(\tau(\tau')), \tag{6.87}$$

in particular $\eta^a(\tau'_0) = \xi^a(\tau(\tau'_0))$. Since $\tau(\tau'_0) = \tau_0$, one finally has $\eta^a(\tau'_0) = \xi^a(\tau_0)$.

In short, parallel transport of a vector can be performed according to Eq. (6.79) using any coordinate system and parametrization.

Exercise

Verify that for the case of the Riemann connection, parallel transport preserves the scalar product of the transported vectors

$$\frac{d}{d\tau}g(\xi, \eta) = 0. \tag{6.88}$$

Therefore, both the length of the vector and the angle between the vectors are preserved.

Comments 1. Covariance of equations and coordinate independence. Let us mention a slightly different treatment frequently implicit in coordinate constructions of the type (6.79). We can find the solution $\xi^a(\tau)$ in the coordinate system q^a, then construct the functions ξ'^a (6.52) in the system q'^a, and so on, according to the procedure described on page 180. By construction, it gives a vector field. In this case we need to confirm that the resulting field does not depend on the choice of the particular coordinate system q^a used in its construction. We now demonstrate that this coordinate independence is guaranteed by a property of the defining equation (6.79) known as its covariance.

For the present treatment of the problem, all the constituents of Eq. (6.79) have the well-established transformation properties (6.52), (6.66) and (6.60) under the coordinate transformations (6.51). Using them, the equation can be *identically* rewritten in terms of the quantities related with the system q'^a. We obtain

$$D\xi^a = \frac{\partial q^a}{\partial q'^b} D'\xi'^b = 0, \tag{6.89}$$

where

$$D'\xi'^b = \frac{d\xi'^b}{d\tau} + \Gamma'^b{}_{cd}\dot{q}'^c\xi'^d = 0, \tag{6.90}$$

which is precisely Eq. (6.79) in the primed system. Further, since $\det \frac{\partial q^a}{\partial q'^b} \neq 0$, the equation $D\xi = 0$ is equivalent to $D'\xi' = 0$. Hence, the equation of parallel transport preserves its form when we pass from one system to another. This property is called *covariance of the equation under coordinate transformations*. Note that covariance is neither a general nor a self-evident fact. For example, the equation of a circle $x^2 + y^2 = 1$ in polar coordinates acquires the form $r = 1$ instead of $r^2 + \theta^2 = 1$, and hence is non-covariant.

Now we are ready to prove the coordinate independence. Suppose that the parallel field has been constructed starting from the system q'^a: $D'\eta^a = 0$, $\eta(\tau_1) = \xi'^a_1$. Note that the functions $\xi'^a(\tau)$ obey this problem due to the covariance property (6.89): $D'\xi'^a \sim D\xi^a = 0$, then $\eta^a = \xi'^a$. Hence $\eta^a(\tau)$ determines the same field $\xi(\tau)$, which shows its coordinate independence.

2. Parallel and covariantly constant fields. According to the known theorem, parallel transport turns out to be line-independent in Riemann space with curvature tensor equal to zero. In the general case, parallel transport depends on the line, as the defining Eq. (6.79) contains $\dot{q}(\tau)$. Nevertheless it can happen that parallel transport of a particular vector is line-independent. As an example, let us consider the covariantly constant field $\xi^a(q)$. Due to Eq. (6.71), it turns out to be parallel along any curve

$$D\xi^a = \dot{q}^b \left[D_b\xi^a\right]\big|_{q(\tau)} = 0. \tag{6.91}$$

Accordingly, parallel transport of $\xi(q_1)$ to a point q_2 along any line, gives a vector of the field itself at q_2. Hence the transport of $\xi(q_1)$ does not depend on the line chosen.[10]

Let us finish this section with two illustrative examples.

[10] Let us point out that Eq. (6.79) itself cannot be rewritten in terms of D_b.

Second law of Newton in curvilinear coordinates. Consider Euclidean space parameterized by the Cartesian coordinates $x^{a'}$. We admit now an arbitrary non-degenerated transformation of the coordinates $x^{a'} \to q^a = q^a(x^{b'})$. The Euclidean scalar product, rewritten in the curvilinear coordinates q^a, acquires the form $(\vec{v}, \vec{w}) = \delta_{a'b'} v^{a'} w^{b'} = \delta_{a'b'} \partial_a x^{a'} \partial_b x^{b'} v^a w^b$. Hence the metric components in the system q^a are given by $g_{ab} = \delta_{a'b'} \partial_a x^{b'} \partial_b x^{b'}$. They are built starting from the matrix $\delta_{a'b'}$ according to the construction described on page 180. In turn, the derivative of the vector field $\vec{v}$ reads

$$\frac{\partial v^{a'}}{\partial x^{b'}} = \frac{\partial q^b}{\partial x^{b'}} \frac{\partial x^{a'}}{\partial q^a} \left(\frac{\partial v^a}{\partial q^b} + \frac{\partial q^a}{\partial x^{d'}} \frac{\partial^2 x^{d'}}{\partial q^b \partial q^c} v^c \right). \tag{6.92}$$

and is identified with the covariant derivative (6.68). The connection $\Gamma^a{}_{bc} = \partial_{c'} q^a \partial^2_{bc} x^{c'}$ is built from $\Gamma^{a'}{}_{b'c'} = 0$ according to the same construction. One can verify that it coincides with the Riemann connection (6.75) of the metric constructed above. Following the same lines, the second law of Newton $\ddot{x}^{a'} = -\partial_{a'} U(x')$ can be written in the form $D\dot{q}^a = -g^{ab} \partial_b U(q)$. On the left-hand side the covariant derivative (6.79) of the vector $\dot{q}^a$ appears. In Sect. 6.8 we show how the potential U can be incorporated into the connection coefficients, thus obtaining the interpretation of the law of Newton in terms of parallel transport.

Free motion on a sphere. Riemann geometry naturally arises in the description of dynamical systems with kinematical constraints. Consider the unit-mass particle constrained to move on the sphere $(x^i)^2 = 1$. We choose x^a, $a = 1, 2$ as the local coordinates on the upper half sphere, then its parametric equations are $x^1 = x^1$, $x^2 = x^2$, $x^3 = \sqrt{1 - (x^a)^2}$.

The variables x^a can be taken as the configuration-space coordinates of the particle. As we have seen in Sect. 1.8, the Lagrangian action for $x^a(\tau)$ is obtained from the free particle action $S = \int d\tau \frac{1}{2} (\dot{x}^i)^2$ by substitution of the constraint $x^3 = \sqrt{1 - (x^a)^2}$ into the integrand. The substitution reads $S = \int d\tau \frac{1}{2} g_{ab} \dot{x}^a \dot{x}^b$, where $g_{ab} = \delta_{ab} + \frac{x_a x_b}{1-(x^a)^2}$ is the metric on the sphere induced by the Euclidean scalar product of the environment space. The corresponding Riemann connection can be computed according to Eq. (6.75); the result is $\Gamma^a{}_{bc} = x^a g_{bc}(x)$. The variation of the action leads to the equations of motion $\ddot{x}^a + \Gamma^a{}_{bc} \dot{x}^b \dot{x}^c = 0$, where the Riemann connection coefficients appear automatically in the course of the variation. According to Eq. (6.79), the equations of motion mean that the velocity $\dot{x}^a$ is the parallel vector field along the particle trajectory. It implies that the trajectory is the shortest line on the sphere; see Sect. 6.7.

6.4 Definition of Covariant Derivative Through Parallel Transport: Formal Solution to the Parallel Transport Equation

The covariant derivative (6.77) of a given vector field $\xi^a(q)$ along the curve $q^a(\tau)$ can be written as follows:

$$D\xi^a = \frac{d\xi^a}{d\tau} + \Gamma^a{}_{bc}\dot{q}^b\xi^c =$$
$$\lim_{\triangle\tau\to 0}\frac{\xi^a(\tau+\triangle\tau) + \Gamma^a{}_{bc}\dot{q}^b\xi^c|_\tau\triangle\tau - \xi^a(\tau)}{\triangle\tau}. \tag{6.93}$$

Let us carry out parallel transport of the vector $\xi^a(\tau+\triangle\tau)$ to the point τ. That is, we solve the problem:

$$\frac{d\eta^a}{d\tau} + \Gamma^a{}_{bc}\dot{q}^b\eta^c = 0, \qquad \eta^a(\tau+\triangle\tau) = \xi^a(\tau+\triangle\tau). \tag{6.94}$$

Expanding the resulting vector $\eta(\tau)$ in a Taylor series at $\tau+\triangle\tau$ we have, disregarding $\triangle\tau^2$ terms

$$\eta^a(\tau) = \eta^a(\tau+\triangle\tau-\triangle\tau) = \eta^a(\tau+\triangle\tau) - \dot{\eta}^a|_{\tau+\triangle\tau}\triangle\tau + \ldots$$
$$= \xi^a(\tau+\triangle\tau) + \Gamma^a{}_{bc}\dot{q}^b\xi^c|_\tau\triangle\tau + \ldots. \tag{6.95}$$

Comparing this expression with Eq. (6.93) we conclude that the covariant derivative can be defined through parallel transport according to the formula

$$D\xi^a = \lim_{\triangle\tau\to 0}\frac{\eta^a(\tau) - \xi^a(\tau)}{\triangle\tau}, \tag{6.96}$$

where $\eta^a(\tau)$ represents the result of parallel transport of the vector $\xi^a(\tau+\triangle\tau)$ to the point τ. It also implies an approximate expression for the transported vector in terms of the initial one:

$$\xi^a(\tau+\triangle\tau)\big|_{\text{parallel transported at }\tau} = \xi^a(\tau) + \triangle\tau D\xi^a(\tau) + \ldots. \tag{6.97}$$

Consider the vector ξ^a_0 at the point $q^a(0)$ of the curve $q^a(\tau)$. We present a generalization of the previous formula that gives a formal solution to the parallel transport equation:

$$D\eta^a = 0, \qquad \eta^a|_{\tau=0} = \xi^a_0. \tag{6.98}$$

Let $\xi^a(\tau)$, $\xi^a(0) = \xi^a_0$ be a vector field along the curve. Then the field

$$\eta^a(\tau) = e^{-\tau D}\xi^a(\tau) \equiv \xi^a - \tau D\xi^a + \frac{1}{2}\tau^2 DD\xi^a + \ldots, \tag{6.99}$$

obeys the transport equation, as can be verified by direct substitution.

6.5 The Geodesic Line and Its Reparametrization Covariant Equation

A straight line in Euclidean space can be characterized by any one of the following properties: (a) a tangent vector to the straight line remains a tangent in the course of its parallel transport along the line; (b) among all the lines between two points the straight line has the minimal length. In the Riemann case, the first property is taken as a basis for the notion of a geodesic line, while the second one defines the shortest line. Since the metrical and parallel transport properties are determined by two independent quantities (by the metric tensor and by affine connection), the lines are different, unless a Riemann connection in Riemann space is chosen. As will be seen in Sect. 6.8, classical mechanics prefers Riemann space with special affine connection.

Reparametrization covariant equation of the geodesic line.

Definition The line $\gamma \in \mathbb{M}$ is called a *geodesic line* if its tangent vector remains a tangent under parallel transport along the line.

Let us obtain a differential equation determining the geodesic line. Let $q^a(\tau)$, $q^a(\tau_1) = q_1^a$ be a parametrization of the geodesic line, and $\xi^a(\tau)$ is the parallel vector field obtained by the transport of a vector $\xi_1(q_1)$ tangent to the geodesic line. According to the above definition we can write

$$\alpha(\tau)\xi^a(\tau) = \dot{q}^a(\tau), \tag{6.100}$$

where $\alpha(\tau)$ is some function. This allows us to rewrite equations of parallel transport (6.79) in terms of q^a and α

$$\ddot{q}^a + \Gamma^a{}_{bc}\dot{q}^b\dot{q}^c - \frac{\dot{\alpha}}{\alpha}\dot{q}^a = 0. \tag{6.101}$$

They are accompanied by the initial conditions

$$q^a(\tau_1) = q_1^a, \qquad \dot{q}^a(\tau_1) = \alpha_1\xi_1^a, \tag{6.102}$$

where $\alpha_1 = \alpha(\tau_1)$. Hence, if γ is the geodesic line, any parametrization $q^a(\tau)$ of it obeys this equation for a certain $\alpha(\tau)$. Conversely, if the functions $q^a(\tau), \alpha(\tau)$ obey the problem (6.101), (6.102), the tangent field $\xi^a \equiv \frac{1}{\alpha}\dot{q}^a$ obeys Eq. (6.79) and thus is a parallel field along the curve $q^a(\tau)$. Hence the curve parameterizes a geodesic line.

The Eqs. (6.101) represent a system of n second order equations for $n + 1$ unknown functions $q^a(\tau)$, $\alpha(\tau)$. This implies that the solution to the problem (6.101), (6.102) is not unique. This ambiguity is not surprising, since the geodesic line, being a set of points in Riemann space, can be parameterized in various ways. We show now that this ambiguity is exclusively due to the reparametrizations. Besides, any particular parametrization is specified by the choice of the function α.

To start with, we show that any solution to the problem (6.101), (6.102) determines the same geodesic line. In other words, a unique geodesic line passes through a given point in a given direction. We also show that, given two solutions to the problem, $q^a(\tau), \alpha(\tau)$ and $y^a(\tau'), \beta(\tau')$, the functions $q^a(\tau)$ and $y^a(\tau')$ represent parameterizations of the same line.

Starting from the solution $q^a(\tau), \alpha(\tau)$, consider the functions $q^\alpha(q^1)$ describing the corresponding line. Substitution of Eq. (6.4) into the geodesic equation leads to the equations for the trajectory $q^\alpha(q^1)$

$$q^{,,\alpha} + \hat{\Gamma}^\alpha{}_{bc} q^{,b} q^{,c} = 0, \qquad \hat{\Gamma}^\alpha{}_{bc} \equiv \Gamma^\alpha{}_{bc} - q^{,\alpha} \Gamma^1{}_{bc}, \tag{6.103}$$

which do not contain α. Subject to the initial conditions[11] following from (6.102), $q^\alpha(q_1^1) = q_1^\alpha$, $q^{,\alpha}(q_1^1) = \frac{\xi_1^\alpha}{\xi_1^1}$, it has a unique solution. So, all solutions to the problem (6.101), (6.102) give the same line $q^\alpha(q^1)$.

Now we demonstrate that the set of solutions $\{q^\alpha(\tau), \alpha(\tau)\}$ is in one-to-one correspondence with the set $\{q^a(\tau)\}$ of all possible parameterizations of the geodesic line.

Let $q^a(\tau)$, $\alpha(\tau)$ and $y^a(\tau')$, $\beta(\tau')$ be two solutions. Since the functions $q^a(\tau)$, $y^a(\tau')$ parameterize the same line, they are related according to

$$y^a(\tau') = q^a(\tau(\tau')), \tag{6.104}$$

where $\tau(\tau')$ is some function, see Eq. (6.61). We substitute this expression into the equation for y

$$\ddot{y}^a + \Gamma^a{}_{bc} \dot{y}^b \dot{y}^c - \frac{\dot{\beta}}{\beta} \dot{y}^a = 0, \tag{6.105}$$

and compare the result with Eq. (6.101). This gives the relationship between α and β

$$\beta(\tau') = \frac{d\tau}{d\tau'} \alpha(\tau(\tau')). \tag{6.106}$$

So, any two solutions are related by Eqs. (6.104) and (6.106) with a function $\tau(\tau')$. This implies that different functions $\alpha \neq \beta$ lead to different parameterizations $q^a \neq y^a$. We can say that the set of solutions is "parameterized" by an arbitrary function $\alpha(\tau)$.

Combining the results, the problem (6.101), (6.102) is ambiguous, the complete set of solutions being composed of the pairs $q^a(\tau)$, $\alpha(\tau)$, where $\alpha(\tau)$ is an arbitrary function and $q^a(\tau)$ is the unique solution to the problem with α substituted into Eq. (6.101). The set is in one-to-one correspondence with the set $\{q^a(\tau)\}$ of parameterizations of the same geodesic line. The fixation of the function α in the

[11] Note that they do not depend on α or on the length of ξ^a.

geodesic Eq. (6.101) is thus equivalent to a choice of a particular parametrization of the geodesic line.

Let us stress that, being of geometric origin, problem (6.101) itself has no dynamical content: while it determines the geodesic line, it does not imply any definite dependence on the parameter τ. For each given parametrization $q^a(\tau)$ of the line, there is $\alpha(\tau)$ such that the pair q^a, α obeys the problem. The dynamics can be "created" by hand, and we do this below specifying the function $\alpha(\tau)$.

Equation (6.100) shows that $\dot{q}^a(\tau)$ is not a parallel field unless $\alpha \neq$ const. The same conclusion follows from the comparison of Eqs. (6.79) and (6.101).

Geodesic equation in canonical parametrization. According to the previous analysis, the function α in the geodesic Eq. (6.101) acquires any desired form after an appropriate choice of parametrization. In particular, there is a parametrization such that $\alpha = 1$. Then Eq. (6.101) acquires a more simple form

$$\ddot{q}^a + \Gamma^a{}_{bc}\dot{q}^b\dot{q}^c = 0. \tag{6.107}$$

Comparing it with the parallel transport Eq. (6.79) we conclude that in this parametrization the tangent vector $\dot{q}^a$ to the curve $q^a(\tau)$ turns out to be a parallel field. The parametrization is known as the *canonical parametrization* of the geodesic line. In contrast to Eq. (6.101), the Eq. (6.107) is not covariant under reparametrizations.

Recall that parallel transport preserves the length of the transported vector: $g(\dot{q}\dot{q}) = v^2 = const$ for all τ. Then the curve $\tilde{q}^a \equiv q^a\left(\frac{1}{v}\tau\right)$ represents the canonical parametrization with the unit tangent vector. It is known as the *natural parametrization*.

6.6 Example: A Surface Embedded in Euclidean Space

The surface S in the three-dimensional Euclidean space $\mathbb{E}$ can be naturally endowed with the Riemann space structure, which is induced by the Euclidean geometry of $\mathbb{E}$. Then various Riemann space constructions acquire a simple geometric interpretation in terms of the Euclidean geometry of the environment space.

Let $\vec{r}(q^a)$ be the parametric equation of the surface. The notation we use is $\vec{r} = x^i\vec{e}_i$, where x^i, $i = 1, 2, 3$, stands for the cartesian coordinates of $\mathbb{E}$, and q^a, $a = 1, 2$, are the local coordinates on the surface. Let $q^a(\tau)$ be a curve on S. The equation $\vec{r}(\tau) \equiv \vec{r}(q^a(\tau))$ describes its embedding into $\mathbb{E}$. Consider the tangent vector to the curve. Being a vector of $\mathbb{E}$, it has the components $V^i = \frac{dx^i(\tau)}{d\tau}$. Being the tangent vector to the surface, it has the coordinates $v^a = \frac{dq^a}{d\tau}$ in the local system q^a. They are related by

$$\vec{V} = \frac{\partial\vec{r}}{\partial q^a}v^a. \tag{6.108}$$

The surface can be endowed with a Riemann space structure as follows. The Euclidean scalar product $(\vec{V}, \vec{W})$ defines the *induced metric* g_{ab} on the surface according to the rule

$$(\vec{V}, \vec{W}) = \frac{\partial x^i}{\partial q^a}\frac{\partial x^i}{\partial q^b} v^a v^b \equiv g_{ab} v^a v^b. \tag{6.109}$$

This can be used to construct the Riemann connection (6.75), the covariant derivative (6.68), and so on.

At each point M of the surface we construct a basis of $\mathbb{E}$ adapted to the surface. The coordinate curves $q^1 = \tau$, $q^2 = q^2_M$ and $q^1 = q^1_M$, $q^2 = \tau$ in the vicinity of M determine the tangent vectors $\vec{\omega}_a$ at the point M

$$\vec{\omega}_1 = \left.\frac{d\vec{r}(q^1, q^2_M)}{dq^1}\right|_{q^1_M}, \qquad \vec{\omega}_2 = \left.\frac{d\vec{r}(q^1_M, q^2)}{dq^2}\right|_{q^2_M}. \tag{6.110}$$

They form a basis of the tangent space $T_M(S)$ called *coordinate basis*. This can be completed up to a basis of $\mathbb{E}$ by addition of the unit normal vector to the surface, which is constructed with the help of the vector product: $\vec{n} = \frac{[\vec{\omega}_1, \vec{\omega}_2]}{|[\vec{\omega}_1, \vec{\omega}_2]|}$. The construction of the adapted basis is illustrated in Fig. 6.4 on page 194.

Now we are ready to rewrite the Riemann space quantities in terms of the Euclidean space basis $(\vec{\omega}_a, \vec{n})$.

Metric. This is presented through the Euclidean scalar product of $\vec{\omega}_a$ as follows:

$$g_{ab} = (\vec{\omega}_a, \vec{\omega}_b). \tag{6.111}$$

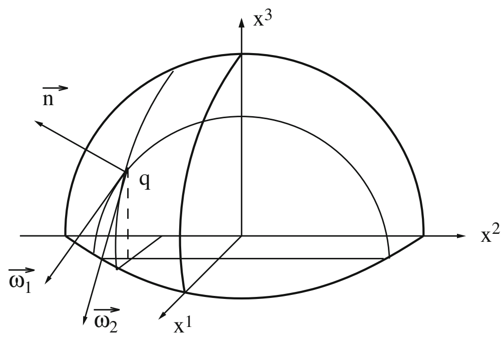

Fig. 6.4 For the semisphere $x^3 = \sqrt{1-(x^1)^2-(x^2)^2}$ we take x^1, x^2 as the local coordinates q^1, q^2. Then the coordinate lines through the point q are obtained by intersection of the sphere with the planes parallel to the coordinate planes (x^2, x^3), (x^1, x^3). The induced metric is given by the Euclidean scalar product $g_{ab} = (\vec{\omega}_a, \vec{\omega}_b)$

Connection. The variation rate of $\vec{\omega}_a$ along the coordinate lines is given by the Euclidean vectors $\frac{\partial\vec{\omega}_a}{\partial q^b}$. They can be decomposed in relation to the adapted basis; we write

$$\frac{\partial\vec{\omega}_a}{\partial q^b} = \Gamma^c{}_{ab}\vec{\omega}_c + N_{ab}\vec{n}. \tag{6.112}$$

Computing the scalar product of this expression with $\vec{\omega}_d$ we obtain

$$\Gamma^c{}_{ab} = g^{cd}(\vec{\omega}_d, \partial_b\vec{\omega}_a) \equiv \frac{1}{2}g^{cd}(\partial_a g_{db} + \partial_b g_{ad} - \partial_d g_{ab}). \tag{6.113}$$

That is the coefficients Γ appeared in Eq. (6.112) represent the Riemann connection. Hence, Eq. (6.112) states that the Riemann connection determines the tangential part of the coordinate basis variation.[12]

Covariant derivative. Let $\vec{V} = \frac{\partial\vec{r}}{\partial q^a}v^a$ be the vector field defined along a curve $q^a(\tau)$. Using Eqs. (6.110), (6.111) and (6.112), its derivative along the curve can be written as follows:

$$\frac{d}{d\tau}\vec{V} = Dv^a\vec{\omega}_a + N_{ab}v^a\dot{q}^b\vec{n}, \tag{6.114}$$

where Dv^a stands for the covariant derivative (6.77). This means that the covariant derivative determines the tangential part of the variation rate of the vector $\vec{V}$ along the curve, see Fig. 6.5 on page 194

Parallel transport. According to Eqs. (6.109) and (6.88), parallel transport along a curve of S preserves the (Euclidean) angle between any pair of transported vectors

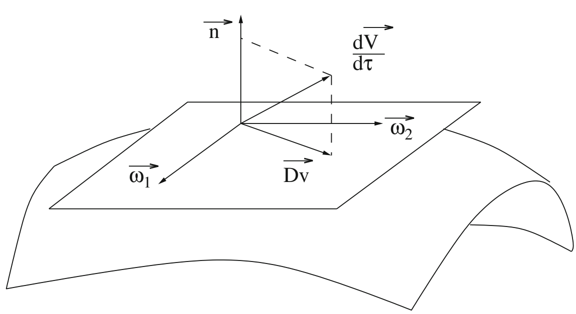

Fig. 6.5 Tangent spaces at distinct point of the surface can have distinct orientations, so the derivative $\left.\frac{d\vec{V}}{d\tau}\right|_q$ of the vector field $\vec{V}$ does not generally lie on the tangent space $T_S(q)$. The covariant derivative $\vec{Dv}$ is the tangential part of $\frac{d\vec{V}}{d\tau}$

[12] N_{ab} are known as the coefficients of *second quadratic form* of the surface.

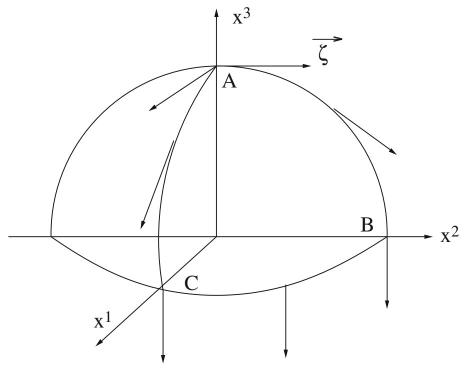

Fig. 6.6 The figure shows the vector field obtained by parallel transport of the vector $\vec{\xi}$ along the closed contour on a sphere. The contour is formed by the geodesic lines: $AB \cup BC \cup CA$

$\vec{V}$ and $\vec{W}$. Let the curve be the geodesic line and $\vec{W}$ represent its tangent vector. Then parallel transport preserves the angle between $\vec{V}$ and the line; see Fig. 6.6 on page 196.

6.7 Shortest Line and Geodesic Line: One More Example of a Singular Action

A metric allows us to define the length of a line and also to introduce the notion of the distance between points in Riemann space.

Consider the Riemann space $\mathbb{M}$ with the metric g_{ab}. Let $q^a(\tau)$ be a parametrization of the line γ. The *length of the line* is the number

$$S = \int_{\tau_1}^{\tau_2} d\tau \sqrt{g_{ab}(q)\dot{q}^a\dot{q}^b}, \tag{6.115}$$

which is a value of the functional S: $\{q^a(\tau)\} \to \mathbb{R}$ computed for any curve $q^a(\tau)$ corresponding to the line. Let us confirm that Eq. (6.115) actually associates a unique number to the given line. First, S does not depend on the choice of coordinates, since the scalar product $g(\dot{q}, \dot{q})$ is invariant under the coordinate transformations. Second, S does not depend on the parametrization of the line: if $y^a(\tau') = q^a(\tau(\tau'))$ is another parametrization (see Eq. (6.61)), we obtain

$$\begin{aligned} S(y^a(\tau')) &= \int_{\tau_1'}^{\tau_2'} d\tau' \left(g_{ab}(y(\tau')) \frac{dy^a}{d\tau'} \frac{dy^b}{d\tau'} \right)^{\frac{1}{2}} \\ &= \int_{\tau_1'}^{\tau_2'} d\tau' \left(g_{ab}(q(\tau(\tau'))) \frac{dq^a(\tau(\tau'))}{d\tau'} \frac{dq^b(\tau(\tau'))}{d\tau'} \right)^{\frac{1}{2}} \end{aligned}$$

$$= \int_{\tau_1'}^{\tau_2'} d\tau' \frac{d\tau}{d\tau'} \left(g_{ab}(q(\tau)) \frac{dq^a(\tau)}{d\tau'} \frac{dq^b(\tau)}{d\tau'} \bigg|_{\tau(\tau')} \right)^{\frac{1}{2}}$$

$$= \int_{\tau_1}^{\tau_2} d\tau \left(g_{ab}\dot{q}^a\dot{q}^b \right)^{\frac{1}{2}} = S(q^a(\tau)), \tag{6.116}$$

that is:

$$S(y^a(\tau')) = S(q^a(\tau)). \tag{6.117}$$

In physical applications, this property is known as the *reparametrization invariance of the functional S*.

Let γ be the shortest line connecting the points q_1, q_2 (that is $S(\gamma) \leq S(\beta)$, where β is any other line between the points). Then the length of γ is called the *distance* between q_1 and q_2: $d(q_1, q_2) = S(\gamma)$. Accordingly, to find d one looks for the function $q(\tau)$ that gives a minimum of the functional S. That is, we need to solve the variational problem (6.115) with fixed ends. Owing to the reparametrization invariance of S, solution of the variational problem is not unique, as it is clear from Eq. (6.117). To analyze the ambiguity, let us find the equations determining the shortest line. The variation of the functional (6.115) gives

$$\Pi^a{}_b \left[\ddot{q}^b + \Gamma^b{}_{cd}(g)\dot{q}^c\dot{q}^d \right] = 0, \tag{6.118}$$

where

$$\Pi^a{}_b = \delta^a{}_b - \frac{\dot{q}^a\dot{q}^c g_{cb}}{g(\dot{q}, \dot{q})} \equiv \delta^a{}_b - \Lambda^a{}_b, \tag{6.119}$$

and $\Gamma^b{}_{cd}(g)$ is the Riemann connection (6.75), which arises automatically in the course of the variation.

Exercise
Obtain these equations.

Some important relativistic models (particle, string, membrane) are usually formulated in terms of the reparametrization invariant action functionals (in this formulation a relativistic invariance turns out to be manifest, see Sects. 1.12.4 and 7.4). So, the functional (6.115) represents a good laboratory for discussion of this kind of model. If we try to treat Eq. (6.115) as an action functional of a mechanical system with the Lagrangian $L = \sqrt{g(\dot{q}, \dot{q})}$, we find $\frac{\partial^2 L}{\partial \dot{q}^a \partial \dot{q}^b} \sim \Pi^a{}_b$. From Eq. (6.119) it follows that the matrix Π possesses the null vector $\dot{q}^b$:

$$\Pi^a{}_b\dot{q}^b = 0, \tag{6.120}$$

So $\det \Pi = \det \frac{\partial^2 L}{\partial\dot{q}\partial\dot{q}} = 0$. Hence Eqs. (6.115) and (6.118) represent an example of singular Lagrangian theory. It can be shown that the singularity is a direct consequence of the reparametrization invariance of the functional. The reparametrization invariance (6.117) clearly shows that the Eqs. (6.118) do not specify any definite law for propagation of the "particle" q^a along the line. This means that in the reparametrization invariant Lagrangian theory the parameter τ cannot be considered to be a parameter of evolution.

The geodesic line Eq. (6.101) has similar properties; see page 192. Moreover, comparing Eqs. (6.107) and (6.118), we conclude that any solution to the geodesic equation in canonical parametrization obeys (6.118). We now demonstrate that this is not merely a coincidence.

Consider a Riemann space with the metric g and the Riemann connection $\Gamma(g)$. Then we can write both the geodesic Eq. (6.101) and the shortest line Eq. (6.118).

We show an equivalence of the problems, establishing a one-to-one correspondence between their solutions.

(A) According to Eq. (6.120), any solution $q^a = f^a(\tau)$ of Eq. (6.101) obeys Eq. (6.118)

$$\Pi^a{}_b(\ddot{f}^b + \Gamma^b{}_{cd}\dot{f}^c\dot{f}^d) = \Pi^a{}_b(\dot{f}^b\frac{\dot{\alpha}}{\alpha}) = 0. \tag{6.121}$$

(B) Let $q^a = f^a(\tau)$ be a solution to Eq. (6.118). Denoting

$$\ddot{f} + \Gamma\dot{f}\dot{f} = t, \tag{6.122}$$

we have $\Pi t = 0$, or, according to Eq. (6.119) $t = \Lambda t = \dot{f}R$, where $R \equiv \frac{g(\dot{f},t)}{g(\dot{f},\dot{f})}$. Then Eq. (6.122) reads $\ddot{f} + \Gamma\dot{f}\dot{f} - \dot{f}R \equiv 0$. That is, f obeys Eq. (6.101) with $\alpha = \exp\int d\tau R$.

Thus, in Riemann space equipped with the Riemann connection, a geodesic line from q_1 to q_2 turns out to be the shortest line between these points. This means, in particular, that we can use Eq. (6.101) instead of Eq. (6.118) to analyze the shortest line, just as we did in Sect. 6.2.

Exercise

Equations of the free relativistic particle $\left(\frac{\dot{x}^\mu}{\sqrt{(\dot{x}^\mu)^2}}\right)^{\cdot} = 0$ represents a particular case of (6.118), so they are equivalent to the system $\ddot{x}^\mu - \frac{\dot{\alpha}}{\alpha}\dot{x}^\mu = 0$. Show that any solution of the system has the form $x^\mu(\tau) = b^\mu f(\tau) + x_0^\mu$, where b^μ, x_0^μ are constants, and $f(\tau)$ is arbitrary function.

Projectors. Ambiguity in solutions to Eq. (6.118) is related to pure algebraic properties of the matrices Π, Λ, which have a simple geometric interpretation. By construction, they have the properties

$$\Lambda^2 = \Lambda, \qquad \Pi^2 = \Pi, \tag{6.123}$$
$$\Pi\Lambda = 0, \tag{6.124}$$
$$1 = \Pi + \Lambda, \tag{6.125}$$

Matrices with these properties are called *projectors*. Equation (6.125) implies decomposition of an arbitrary vector into two parts

$$\xi = (\Pi + \Lambda)\xi = \Pi\xi + \Lambda\xi \equiv \xi_\perp + \xi_\|. \tag{6.126}$$

Given $\dot{q}^a$ that specifies the projectors (6.119), we have

$$\xi_\|^a = (\Lambda\xi)^a = \frac{g(\dot{q}, \xi)}{g(\dot{q}, \dot{q})}\dot{q}^a, \quad \text{or} \quad \xi_\|^a \sim \dot{q}^a, \tag{6.127}$$
$$g(\dot{q}, \xi_\perp) \equiv 0. \tag{6.128}$$

Hence, an arbitrary vector can be decomposed into a sum of its longitudinal and transverse parts with respect to $\dot{q}$: according to Eq. (6.127), $\xi_\|$ is a projection of ξ on the direction of $\dot{q}$, while Eq. (6.128) shows that $\xi_\perp$ is a projection of ξ on the orthogonal to $\dot{q}$ subspace.

In particular, let us decompose the vector t from Eq. (6.122): $t = t_\| + t_\perp$. Then the shortest-line Eq. (6.118) reads $t_\perp = 0$, giving a restriction on the transverse part only. The longitudinal part of t can be arbitrary, which gives an algebraic explanation of the ambiguity presented in solutions to the problem (6.118).

It has already been mentioned that $\det \Pi = 0$. Let us demonstrate

$$\text{rank}\, \Pi = n - 1. \tag{6.129}$$

From $\det \Pi = 0$ it follows that $rank\ \Pi \le n - 1$. Suppose $rank\ \Pi < n - 1$. Then Π has at least one more independent null vector $\eta \ne c\dot{q}$, $c = \text{const}$. Equations (6.125) and (6.127) then lead to the contradiction: $\eta = \Pi\eta + \Lambda\eta = \Lambda\eta \sim \dot{q}$.

Owing to the non-invertibility of Π, the system (6.118) does not have the normal form. Equation (6.129) implies that Π has an invertible $(n-1)\times(n-1)$ -block. Hence the system (6.118) is equivalent to a normal system of $n - 1$ equations for n variables.

Exercise

Supposing invertibility of the block $\Pi^\alpha{}_\beta$, $\alpha, , \beta = 2, 3, \ldots, n$, obtain equations for the trajectory (6.42) from (6.118).

6.8 Formal Geometrization of Mechanics

According to Sect. 6.2, equations for the *trajectory* of a potential motion can be identified with the geodesic line equations in canonical parametrization. As we have seen, in this parametrization the evolution parameter τ does not correspond to the physical (classical mechanical) time. Our purpose now is to find the geometric condition that picks out physical time among all the possible parameterizations of the geodesic line. First we look for the equations of geometric origin that could describe the complete problem. We demonstrate that equations of motion in a given potential can be identified with the geodesic equation in a special parametrization on a manifold with the affine connection specified by the potential. Further, the manifold can be equipped with an appropriate metric, which is also specified by the potential. Then, the special parametrization can be fixed from the pure geometric condition that the tangent vector to the geodesic curve has a unit length in this metric. That is we have geodesic motion with unit speed. In this way, we arrive at the fully geometric treatment of the potential motion problem.

Consider the action

$$S = \int d\tau (\frac{1}{2} c_{ab}(q)\dot{q}^a \dot{q}^b - U(q^a)), \tag{6.130}$$

in the generalized coordinates $q^a(\tau)$. Here $\frac{1}{2}c_{ab}(q)\dot{q}^a\dot{q}^b$ is the kinetic energy and $U(q^a)$ is a potential. This leads to the equations of motion

$$\ddot{q}^a + \Gamma^a{}_{bc}(c)\dot{q}^b\dot{q}^c + c^{ab}\partial_b U = 0, \tag{6.131}$$

where the coefficients $\Gamma(c)$ are given by Eq. (6.75). We wish to hide the potential term in the connection coefficients. So, let us write $\Gamma^a{}_{bc}(c_{de}) = \Gamma^a{}_{bc}(\frac{1}{\phi}g_{de})$, where $g_{ab} \equiv \phi c_{de}$, and try to choose the function $\phi(U)$ that allows us to identify the equations of motion with the geodesic equations. In this notation Eq. (6.131) acquires the form

$$\ddot{q}^a + \Gamma^a{}_{bc}(g)\dot{q}^b\dot{q}^c - \frac{\dot{\phi}}{\phi}\dot{q}^a + g^{ab}\left(\frac{1}{2\phi}g(\dot{q},\dot{q})\partial_b\phi + \phi\partial_b U\right) = 0. \tag{6.132}$$

Except for the last term, it is similar to the geodesic equation. The last term depends on the velocity and cannot be generally canceled by the choice of ϕ. But it can be achieved for solutions with a fixed total energy. Indeed, let $q^a(\tau)$ be a solution with the total energy h

$$\frac{1}{2\phi}g(\dot{q},\dot{q}) + U(q) = h. \tag{6.133}$$

If we substitute the solution into Eq. (6.132), the last term of the equation acquires the form $\partial_c\phi(h-U) - \phi\partial_c(h-U)$. It vanishes if we take $\phi \sim (h-U)$. The conventional choice is $\phi = 2(h - U)$. Then any solution to (6.131) with fixed energy h obeys the equation

$$\ddot{q}^a + \Gamma^a{}_{bc}(2(h - U)c)\dot{q}^b\dot{q}^c + \frac{\dot{U}}{h - U}\dot{q}^a = 0. \tag{6.134}$$

This is identical to the geodesic Eq. (6.101) in the parametrization (the factor 2 below is also conventional)

$$\alpha = 2(h - U), \tag{6.135}$$

on a manifold with the affine connection $\Gamma^a{}_{bc}(2(h-U)c)$ (so far without a metric!). By analogy with Eq. (6.107), the Eq. (6.134) can be called the *geodesic equation in dynamical parametrization*.[13]

The equation obtained is not yet of geometric origin, since the particular parametrization has little sense from the geometric point of view. We now improve it by pointing out a geometric condition equivalent to the parametrization. In the above construction we dealt with n differential equations for n unknown functions $q^a(\tau)$. The geometric condition can be formulated in a slightly different context of $n + 1$ equations for $n + 1$ variables. We return to the reparametrization-covariant Eq. (6.101) with an arbitrary function $\alpha(\tau)$, and add one more equation that implies fixation of α according to (6.135). For the equation, the natural candidate is the constant energy condition (6.133), which we write in the form

$$G_{ab}\dot{q}^a\dot{q}^b = 1, \qquad G_{ab} \equiv \frac{c_{ab}}{2(h - U)}. \tag{6.136}$$

Let us equip the manifold with the metric G_{ab}. Then the equation states that the vector $\dot{q}^a$ is of unit length. Now, on the Riemann manifold with the metric G and the affine connection $\Gamma(2(h - U)c)$ let us consider the system

$$\ddot{q}^a + \Gamma^a{}_{bc}(2(h - U)c)\dot{q}^b\dot{q}^c + \frac{\dot{\alpha}}{\alpha}\dot{q}^a = 0, \tag{6.137}$$

$$G_{ab}\dot{q}^a\dot{q}^b = 1. \tag{6.138}$$

This problem turns out to be equivalent to the potential motion problems (6.130) and (6.131). To confirm this statement, let $q^a(\tau), \alpha(\tau)$ be a solution to the problem. The Eq. (6.137) implies that the vector field $\xi^a = \frac{1}{\alpha}\dot{q}^a$ is parallel along the geodesic line; see Sect. 6.5. Since our affine connection looks like the Riemann connection constructed on the base of tensor $g_{ab} = 2(h - U)c_{ab}$, the vector ξ obeys (see Eq. (6.88)) $g(\xi, \xi) = v^2 = \text{const}$, or, equivalently, $\frac{2(h-U)c_{ab}}{\alpha^2}\dot{q}^a\dot{q}^b = v^2$. Using

[13] For the case of the Riemann connection, dynamical parametrization is precisely the natural parametrization, see page 193.

Eq. (6.138), we conclude that $\alpha = \frac{2(h-U)}{v}$. With this α, the Eq. (6.137) coincides with the eq. (6.134). Therefore the functions $q^a(\tau)$ describe the potential motion.

To sum up, the potential motion problem (6.130), (6.131) can be described in geometric terms as follows. The configuration space is endowed with the Riemann space structure introducing the metric $G_{ab} = \frac{c_{ab}}{2(h-U)}$ and the affine connection (6.75) constructed on the base of the tensor $g_{ab} = 2(h-U)c_{ab}$. Then the configuration space particle $q^a(\tau)$ with total energy h moves along the geodesic line with unit speed computed with respect to the metric G. The motion can be described by equations of geometric origin (6.137) and (6.138). Equation (6.137) states that the particle chooses the geodesic line as the trajectory of motion. Equation (6.138) means that among all the parameterizations of the geodesic line, the particle chooses the one that implies its unit speed with respect to the metric G.

Chapter 7
Transformations, Symmetries and Noether Theorem

It was mentioned in Sect. 2.5 that conservation laws play an important role in the analysis of classical and quantum systems. This chapter is mainly devoted to discussion of the Noether theorem, which gives the relationship between the existence of conservation laws for the system in question, and symmetries of the associated action functional. The symmetries usually have a certain physical interpretation; in particular, they may reflect some fundamental properties assumed for our space-time: homogeneity, isotropy, ... In this case, the Noether theorem states that conservation laws are consequences of these properties. For example, symmetry under spatial translations implies the conservation of the total momentum of a system.

To demonstrate an idea of the Noether theorem, let us consider the following special situation. Starting from any given trajectory $q^a(\tau)$, let $q'^a(\tau) = q^a(\tau) + R^a(q(\tau))\omega$ be a family of trajectories parameterized by the parameter ω. Here $R^a(q)$ is a given function. Suppose that the Lagrangian action is invariant under the substitution $q \to q'$, that is $S[q'] = S[q]$, for any given $q(\tau)$ and ω. In particular, the variation of the action must also be zero: $\delta S = S[q']\big|_{\text{linear on }\omega\,\text{part}} - S[q] = 0$. On other hand, the variation is given by the well-known expression $\delta S = \frac{\delta S}{\delta q}R + (\frac{\partial L}{\partial \dot{q}}R)^{\cdot}$. Due to invariance, we obtain $(\frac{\partial L}{\partial \dot{q}}R)^{\cdot} = -\frac{\delta S}{\delta q}R$. The identity holds for any $q(\tau)$. In particular, if $q(\tau)$ is a solution to the equations of motion: $\frac{\delta S}{\delta q} = 0$, the identity implies $(\frac{\partial L}{\partial \dot{q}}R)^{\cdot} = 0$. That is the quantity $\frac{\partial L}{\partial \dot{q}^a}R^a$ is a constant throughout any solution.

Besides the Noether theorem, we discuss some closely-related topics: the notion of symmetry for equations of motion, its relationship with the symmetry of an action, the relationship between the Lagrangian and Hamiltonian symmetries, Galileo and Poincaré symmetry groups and so on. The reader who is interested only in the Noether theorem can skip to the corresponding section after reading the first one.

7.1 The Notion of Invariant Action Functional

Here we discuss the intuitive notion of invariant action with simple examples. Exact definitions will be given in the next section. Consider a free particle action functional

A. Deriglazov, *Classical Mechanics*, DOI 10.1007/978-3-642-14037-2_7,

$$S = \frac{1}{2}\int dt\dot{x}^a\dot{x}^a. \tag{7.1}$$

Given the numeric matrix R_{ab}, let us make the following formal substitution

$$x^a \to R_{ab}x^b, \tag{7.2}$$

in Eq. (7.1). It gives a functional that is generally different from (7.1), $\frac{1}{2}\int dt(R^T R)_{ab}\dot{x}^a\dot{x}^b$. But for the orthogonal matrix, $R^T R = 1$, the substitution does not change the Lagrangian as well as the action functional

$$\frac{1}{2}\int dt(R^T R)_{ab}\dot{x}^a\dot{x}^b = \frac{1}{2}\int dt\dot{x}^a\dot{x}^a, \text{ that is } S[Rx] = S[x]. \tag{7.3}$$

In this case it is said that the action is *invariant*, and the corresponding substitution is called a *symmetry transformation of the action*.

It should be noted that an action functional is an operation defined on functions $f^a(\tau)$ instead of coordinates x^a. So consistent treatment of Eq. (7.3) implies that we work with the function $x^a = f^a(\tau)$ and assign to it the function $x^a = f'^a(\tau) \equiv R_{ab}f^b(\tau)$ induced by the substitution (7.2). Bearing in mind this correction, the above-mentioned terminology is reasonable.

To show the meaning of invariance property (7.3) we discuss its two physical interpretations.

(A) The symmetry turns solutions to equations of motion into other solutions. So in practice it can be used to construct new solutions from known ones.

To confirm this, take a trajectory $q^a = f^a(t)$, and construct another one, $q^a = f'^a(t) \equiv R_{ab}f^b(t)$, induced by the substitution (7.2). Invariance implies that the action has the same value on these trajectories, $S[f'] = S[f]$. Consider now a set of trajectories $\{f^a, g^a, \ldots\}$ with the same initial and final points, and let f represent the true trajectory, that is $S[f] \le S[g]$ for all g of the set. All the transformed trajectories $\{f', g', \ldots\}$ have the same initial and final points. Due to the invariance, one obtains $S[f'] \le S[g']$.

(B) Let us identify the substitution (7.2) with a transformation relating cartesian coordinates x^a and x'^a used by two observers, O and O' (remembering that an orthogonal matrix corresponds to a rotation of the cartesian axis)

$$x^a \longrightarrow x'^a = R_{ab}x^b, \tag{7.4}$$

Now x and x' stand for coordinates of the *same* point of configuration space. If O found parametrization of a trajectory to be $q^a = f^a(t)$, the observer O' will describe this curve by the function $q'^a = f'^a(t) \equiv R_{ab}f^b(t)$. According to Sect. 1.6, to study the particle motion, O' can use an action obtained from (7.1) by the change of variables, $x^a = \tilde{R}_{ab}x'^b$, where $\tilde{R}$ is the inverse matrix for R. It reads

$$S'[x'] = \frac{1}{2}\int dt(\tilde{R}^T\tilde{R})_{ab}\dot{x}'^a\dot{x}'^b = \frac{1}{2}\int dt\dot{x}'^a\dot{x}'^a = S[x']. \tag{7.5}$$

The second equality is due to the invariance. That is an invariance guarantees that O', describing the system, can take *the same* action as O, simply using his own coordinates x' instead of x. This implies an identical form of equations of motion in the two systems. If $F(x, \dot{x}) = 0$ stands for an equation obtained by O, then $F(x', \dot{x}') = 0$ with *the same* F represents this equation in the system O'. One can say that physical laws have an identical form in the coordinate systems O and O'.

Disregarding the total derivative term, the action (7.1) is also invariant under the Galileo boost $t \to t$, $x^a \to x'^a = x^a + v^a t$, $v^a = \text{const}$

$$S = \frac{1}{2}\int dt(\dot{x}^a + v^a)^2 = \frac{1}{2}\int dt\left((\dot{x}^a)^2 + \frac{d}{dt}(2v^a x^a + v^2 t)\right). \tag{7.6}$$

Although now $S[f'] \neq S[f]$, it still turns solutions into solutions. Hence it would be reasonable to admit a total derivative term in the definition of an action symmetry.

One more example is the relativistic particle action

$$S = \frac{1}{2}\int dt\sqrt{c^2 - \dot{x}^a\dot{x}^a}, \qquad a = 1, 2, 3. \tag{7.7}$$

The Lorentz boost

$$t \to t' = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\left(t - \frac{v}{c^2}x^1\right), \quad x^1 \to x'^1 = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}(x^1 - vt),$$
$$x^2 \to x^2, \qquad x^3 \to x^3. \tag{7.8}$$

leaves the action invariant for any value of the numeric parameter v.

Exercise
Verify the invariance, $S[t', x'] = S[t, x]$, by direct computation using the rules (we will confirm them in the next section)

$$\frac{dx'^1}{dt'} = \frac{\frac{dx'^1}{dt}}{\frac{dt'}{dt}}, \qquad \frac{dx^\alpha}{dt'} = \frac{\frac{dx^\alpha}{dt}}{\frac{dt'}{dt}}, \quad \alpha = 1, 2. \tag{7.9}$$

This symmetry mixes space coordinates with time variables, that is, here we deal with coordinate transformation of the extended configuration space t, x^a.

As an example of non-linear symmetry, consider the action

$$S = \frac{1}{2}\int dt \left(\delta_{\alpha\beta} + \frac{x^\alpha x^\beta}{1-(x^\gamma)^2}\right)\dot{x}^\alpha \dot{x}^\beta, \quad \alpha, \beta, \gamma = 1, 2. \tag{7.10}$$

This describes a free particle moving on a two-dimensional semisphere of unit radius. Besides the two-dimensional rotations, $x'^\alpha = R_{\alpha\beta}x^\beta$, $R^T R = 1$, it is invariant under the transformation

$$x^1 \to x^1, \qquad x^2 \to x'^2 = x^2 \cos\varphi + \sqrt{1-(x^\gamma)^2}\sin\varphi, \tag{7.11}$$

for any value of the parameter φ. One more symmetry is obtained from (7.11) replacing $x^1 \leftrightarrow x^2$.

Exercise
Show the invariance. Hint: First notice that the Lagrangian can be written as $\frac{1}{2}((\dot{x}^\alpha)^2 + (\dot{x}^3)^2)$, where $x^3 \equiv \sqrt{1-(x^\alpha)^2}$. Second, show that (7.11) implies $x'^3 = -x^2 \sin\varphi + x^3 \cos\varphi$.

According to the examples discussed above, action symmetries can be induced by coordinate transformations of extended configuration space. Moreover we generally deal with a family of transformations parameterized by a set of continuous parameters. So the general case of the family reads

$$\begin{pmatrix} \tau \\ q^a \end{pmatrix} \leftrightarrow \begin{pmatrix} \tau' = \alpha(\tau, q^a, \omega^\alpha) \\ q'^a = \psi^a(\tau, q^a, \omega^\alpha) \end{pmatrix}. \tag{7.12}$$

This will be taken as the starting point of our discussion in the next section.

7.2 Coordinate Transformation, Induced Transformation of Functions and Symmetries of an Action

We present here two equivalent definitions for an action symmetry. We start from the definition that makes the concept clear and then obtain the other one usually used in practical calculations. It is worth noting that the notion of symmetry and the Noether theorem work both for mechanical and field theories that can be described by differential equations following from the variational principle for a functional. It can be a Lagrangian or Hamiltonian action functional, or some other. We start from the Euler–Lagrange equations. The Hamiltonian version of the Noether theorem is discussed in Sect. 7.11.

Consider a dynamical system described by equations of motion following from the action functional

$$S[q] = \int_{\tau_1}^{\tau_2} d\tau L(q^a, \dot{q}^a, \tau), \tag{7.13}$$

defined in the space of functions $q^a = f^a(\tau)$, $\tau \subset [\tau_1, \tau_2]$. In this section it will be convenient to use a different notation for coordinates of the configuration space: q^a, and for the curves: $q^a = f^a(\tau)$, that is for maps $f : \mathbb{R} = \{\tau\} \longrightarrow \mathbb{R}^n = \{q^a\}$.

It will be convenient to work in extended configuration space parameterized by the coordinates τ, q^a: $\mathbb{R}^{n+1} = \mathbb{R}\times\mathbb{R}^n = \{(\tau, q^a)\}$. Consider a family $G = \{g(\omega^\alpha)\}$ of *coordinate transformations* specified by given functions α, ψ^a, and parameterized by k parameters ω^α, $\alpha = 1, 2, \ldots, k$

$$g(\omega^\alpha) : (\tau, q^a) \longrightarrow (\tau', q'^a) = (\, \alpha(\tau, q^a, \omega^\alpha), \psi^a(\tau, q^a, \omega^\alpha)\,). \tag{7.14}$$

We adopt "an active point of view", that is the transformation g turns a point with coordinates (τ, q) into another point, with the coordinates ($\alpha(\tau, q^a, \omega^\alpha)$, $\psi^a(\tau, q^a, \omega^\alpha)$). That is (τ', q'^a) are labels of the transformed point in the *same* coordinate system. The transformation is presumed to be invertible

$$\frac{\partial(\alpha, \psi^a)}{\partial(\tau, q^b)} \neq 0. \tag{7.15}$$

Suppose also that the parametrization has been chosen in such a way that transformation with $\omega^\alpha = 0$ is the identity transformation

$$\alpha(\tau, q^a, 0) = \tau, \qquad \psi^a(\tau, q^a, 0) = q^a. \tag{7.16}$$

Where this cannot lead to confusion, we suppress the parameters ω^α (as well as the indices of the coordinates: $q^a \to q$, $\psi^a \to \psi$ and so on).

Comments (1) In general, neither a composition nor an inverse transformation are guaranteed to be members of a family. But the families arising in physical applications typically possess these additional properties, forming the so called *Lie group*.

Suppose a family of transformations obeys the following properties. (A) The product (that is, a consecutive application) of two transformations of the family is a member of the family as well: $g(\omega_2)g(\omega_1) = g(\omega_3(\omega_2, \omega_1))$. (B) The family contains a unit. It is an element e with the property $eg = ge = g$ for any g (for our case it is the identity transformation, $e = g(0)$, see (7.16)). (C) For any g an inverse transformation $\tilde{g}$ is a member of the family, that is $\tilde{g} = g(\omega)$ for some ω. A family equipped with a product obeying these properties is precisely a Lie group.

(2) All the families of transformations discussed in the previous section are examples of a Lie group. Specifically, for the Galileo boosts $g(v^a) : x^a \to x'^a = x^a + v^a t$, note that $x''^a = x'^a + v_2^a t = x^a + (v_2^a + v_1^a)t$. That is, the product gives a boost parameterized by sum, $g(v_2)g(v_1) = g(v_2 + v_1)$. The unit element is $e = g(0)$, and an inverse for $g(v^a)$ is given by $\tilde{g} = g(-v^a)$.

Exercises

(1) Show that the family (7.11) form a Lie group with the same composition law as for Galileo boosts: $g(\varphi_2)g(\varphi_1) = g(\varphi_2 + \varphi_1)$.

(2) Show that the transformations $g(R, a) : x \to x' = Rx + a, R \neq 0$ of one-dimensional space parameterized by x form a Lie group. Find the composition law, a unit and an inverse element.

(3) Coordinate τ' of the transformed point generally depends on q^a; see Eq. (7.14). Such transformations arise, in particular, in relativistic theories formulated in terms of physical variables; see example of a relativistic particle in Sect. 7.4. An example of a transformation with q' dependent on τ is the Galileo boost. We also point out that a typical form of transformations in classical (non relativistic) mechanics is either $\tau' = \alpha(\tau)$, $q' = q$; or $\tau' = \tau$, $q' = \psi(\tau, q)$. That is, either τ or q remains unaffected. In contrast, the form (7.14) turns out to be typical for symmetry transformations in field theories (with the corresponding substitutions $\tau \to x^\mu$, $q^a \to \varphi^a(x^\mu)$).

Remember that an action functional is an operation defined by functions instead of coordinates. So, to formulate consistently how S is affected by a coordinate transformation, we need to decide how the map g acts on a function $q^a = f^a(\tau)$. The idea is to identify the function with its graph (see Fig. 2.1 on page 92)

$$\Gamma_f = \{(\tau, f^a(\tau)), \ \tau \subset [\tau_1, \tau_2]\}. \tag{7.17}$$

The map g transforms it into another graph, $\Gamma_{f'}$, and the problem is to find the corresponding function. This leads to the following rule for transformation of a function:

$$\begin{aligned} * g : q^a = f^a(\tau) \ \to \ q^a = f'^a(\tau) \equiv \psi^a(\tilde{\alpha}(\tau), f(\tilde{\alpha}(\tau)), \\ \tau \subset [\alpha(\tau_1, f(\tau_1)), \alpha(\tau_2, f(\tau_2))]. \end{aligned} \tag{7.18}$$

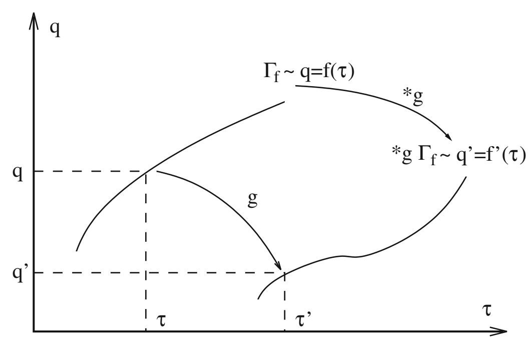

Fig. 7.1 Coordinate transformation g induces the map ${}^*g : f \to f'$ in the space of functions

Here $\tilde{\alpha}(\tau)$ is, for each given $f(\tau)$, an inverse function for $\alpha(\tau, f(\tau), \omega^\alpha)$ considered as a function of τ, that is:

$$\alpha(\tilde{\alpha}(\tau), f(\tilde{\alpha}(\tau)) = \tilde{\alpha}(\alpha(\tau, f(\tau)) = \tau. \tag{7.19}$$

To confirm the formula, represent the initial function in a parametric form $\tau = \sigma$, $q = f(\sigma)$, $\sigma \subset [\tau_1, \tau_2]$. Applying the map g, the point with these coordinates goes over to the point with coordinates

$$\tau = \alpha(\sigma, f(\sigma)), \qquad q = \psi(\sigma, f(\sigma)). \tag{7.20}$$

The equations determine the transformed function in a parametric form. Resolving the first equation with respect to σ, $\sigma = \tilde{\alpha}(\tau)$, and using this to eliminate the parameter from the second equation, we arrive at Eq. (7.18).

Note that the inverse function $\tilde{\alpha}$ depends on a particular f, so the obtained representation is rather formal.

Now, for a given function $f^a(\tau)$, let us construct the image $f'^a(\tau)$, and compute the *same* functional (7.13) on f'

$$S[f'] = S[*gf] = \int_{\alpha(\tau_1, f(\tau_1))}^{\alpha(\tau_2, f(\tau_2))} d\tau L(f'^a, \dot{f}'^a, \tau). \tag{7.21}$$

Definition 1 Coordinate transformations (7.14), (7.15) and (7.16) represent *symmetry of the action* (7.13) (*variational symmetry*), if for any $f^a(\tau)$ and ω^α there is a function $N(f, \dot{f}, \tau, \omega)$ such that:

$$\int_{\alpha(\tau_1, f(\tau_1))}^{\alpha(\tau_2, f(\tau_2))} d\tau L(f'^a, \dot{f}'^a, \tau) = \int_{\tau_1}^{\tau_2} d\tau \left[L(f^a, \dot{f}^a, \tau) + \frac{dN}{d\tau} \right]. \tag{7.22}$$

Let us stress that Eq. (7.22) represents equality of two *numbers*. In particular, in many practically interesting cases, one has $N = 0$, then for any f and the corresponding f', the number $S[f']$ must be equal to the number $S[f]$. The notion of symmetry turns out to be interesting, in particular, for the following reason: if f^a represents a solution to equations of motion, the same is true for f'^a; see the previous section.

Being written as a property of a functional under transformation acting on functions, we can generalize the invariance condition, admitting the terms with derivatives of the function

$$*g : f^a(\tau) \ \to \ f'^a(\tau) \equiv \psi^a(\tilde{\alpha}(\tau), f(\tilde{\alpha}(\tau), \dot{f}(\tilde{\alpha}(\tau), \ldots). \tag{7.23}$$

Generally it can not be treated as induced by a coordinate transformation.

As it is written, the invariance condition is presented in terms of the same functional computed on two different functions (initial and transformed). We now come back to the initial function on the l.h.s. of Eq. (7.22), thus obtaining the invariance condition in terms of the initial and some *transformed actions*, both computed on the same function. At the end, we obtain the invariance condition in a form which is convenient for applications, as an algebraic property of a Lagrangian function under the coordinate transformations (7.14).

To achieve this, we make a change of variables in the definite integral. The change will be chosen from the requirement that the integration limits on the left and right hand sides of the final expression coincide. Equation (7.19), suggests the change $\tau \to \alpha(\tau, f(\tau))$. Using the identity

$$\left.\frac{df'(x)}{dx}\right|_{x=\alpha(\tau,f(\tau))} = \left(\frac{d\alpha}{d\tau}\right)^{-1}\frac{df'(\alpha)}{d\tau}, \tag{7.24}$$

Equation (7.22) acquires the form

$$\int_{\tau_1}^{\tau_2} d\tau\dot{\alpha}L(\psi(\tau, f(\tau)), (\dot{\alpha})^{-1}\frac{d\psi(\tau, f(\tau))}{d\tau}, \alpha) = \int_{\tau_1}^{\tau_2} d\tau\left[L(f(\tau), \dot{f}, \tau) + \frac{dN}{d\tau}\right], \tag{7.25}$$

where it has been used the equality $f'(x)|_{x=\alpha(\tau,f(\tau))} = \psi(\tau, f(\tau))$, which follows from the representation (7.18) and from the identity (7.19). In Eq. (7.25) instead of the induced transformation of a function there appeared the coordinate transformation (7.14) itself. In contrast to Eq. (7.22), both sides of Eq. (7.25) are written for the same function $f(\tau)$. So, it can be said that under the transformation (7.14) the initial action transforms into another action (written on the l.h.s. of the last equation).

Definition 2 The action

$$S_g[q] \equiv \int_{\tau_1}^{\tau_2} d\tau\dot{\alpha}L(\,\psi(\tau, q), (\dot{\alpha})^{-1}\dot{\psi}(\tau, q), \alpha\,), \tag{7.26}$$

is called the *transformation of the action* (7.13) under the coordinate transformations

$$g : \tau \to \tau' = \alpha(\tau, q^a), \qquad q^a \to q'^a = \psi^a(\tau, q^a). \tag{7.27}$$

As we have just shown, (7.21) is equal to (7.26). That is, symbolically

$$S[*gf] = S_g[f]. \tag{7.28}$$

In terms of the transformed action, an invariance condition is formulated as follows:

Definition 3 The action (7.13) is invariant, if, disregarding the total derivative, the transformed action coincides with the initial one

$$S_g[q] = S[q] + \int d\tau\frac{dN}{d\tau}. \tag{7.29}$$

Since this equality must be satisfied for any integration interval, the integrals can be omitted. This gives the invariance condition as an algebraic property of a Lagrangian under the coordinate transformations

$$\dot{\alpha} L(\psi(\tau, q), (\dot{\alpha})^{-1} \dot{\psi}(\tau, q), \alpha) = L(q, \dot{q}, \tau) + \frac{dN}{d\tau}. \tag{7.30}$$

7.3 Examples of Invariant Actions, Galileo Group

Example 1. Consider rotations of the two-dimensional space (τ, q)

$$\theta : (\tau, q) \to (\tau', q') = (\tau \cos\theta - q \sin\theta, \tau \sin\theta + q \cos\theta). \tag{7.31}$$

Let us find the image of the linear function $q = f(\tau) = a\tau + b$. According to Eqs. (7.18) and (7.20), we need to eliminate σ from the equations $\tau = \sigma \cos\theta - (a\sigma + b)\sin\theta$, $q = \sigma \sin\theta + (a\sigma + b)\cos\theta$, which again gives the linear function $q = a'\tau + b'$ (one straight line is rotated into another straight line)

$$q = f'(\tau) = \frac{\sin\theta + a\cos\theta}{\cos\theta - a\sin\theta}\tau + \frac{b}{\cos\theta - a\sin\theta}. \tag{7.32}$$

Example 2. Consider translations of the evolution parameter

$$a : (\tau, q^a) \to (\tau', q'^a) = (\tau + a, q^a), \qquad a = \text{const}, \tag{7.33}$$

The image of the function $f(\tau)$ is obtained from the parametric equations $\tau = \sigma + a$, $q'^a = f^a(\sigma)$. We find

$$q^a = f'^a(\tau) = f^a(\tau - a). \tag{7.34}$$

Translations are the symmetry of any action which does not explicitly depend on τ: $S = \int d\tau L(q, \dot{q})$. The transformed functional is obtained according to Eq. (7.21), and coincides with the initial one after the change of variables $\tau \to \tau + a$

$$\int_{\tau_1+a}^{\tau_2+a} d\tau L(f(\tau - a), \frac{d}{d\tau} f(\tau - a)) = \int_{\tau_1}^{\tau_2} d\tau L(f(\tau), \dot{f}(\tau)) \tag{7.35}$$

Thus the condition (7.22) is satisfied with $N = 0$.

Intuitively, the physical interpretation of the time translations is that an experiment carried out during the time interval $[\tau_1, \tau_2]$, can be repeated at a different time: $[\tau_1 + a, \tau_2 + a]$. The invariance of the action implies that the same experiment carried out "today" and "tomorrow" gives identical results, since in both cases the same trajectory is an extremum of the functional (see Fig. 2.1 on page 92). Equations (7.33) and (7.35) can be thought of as the mathematical formulation of *homogeneity*

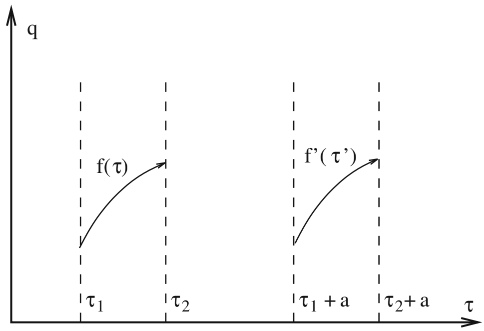

Fig. 7.2 Time translation: the same trajectory turns out to be an extremum of a functional at a different time

in time: the properties of a physical system at different times are the same. As will be seen below, symmetry under time translations implies the energy conservation law. That is, energy conservation is a consequence of homogeneity in time.

Example 3. Consider the *Galileo boosts*, which is a three-parameter family of transformations of $\mathbb{R} \times \mathbb{R}^3$

$$v : \tau \to \tau' = \tau, \qquad x^i \to x'^i = x^i + v^i \tau, \qquad v^i = \text{const.} \tag{7.36}$$

In three-dimensional Euclidean space, these equations can be thought of as related coordinates of two observers O and O', with the latter moving at velocity v^i in relation to O, and passing through the point (0, 0, 0) at $\tau = 0$. Since time is unchanged, the induced transformation of functions coincides with the x-transformation

$$*v : f^i(\tau) \to f'^i(\tau) = f^i(\tau) + v^i \tau. \tag{7.37}$$

The action of a free moving particle is invariant under the boosts. Indeed, Eq. (7.29) turns out to be satisfied

$$\int d\tau \frac{1}{2} m[(x^i + v^i \tau)^{\cdot}]^2 = \int d\tau \left(\frac{1}{2} m(\dot{x}^i)^2 + \frac{dN}{d\tau} \right), \tag{7.38}$$

with nontrivial $N(x, \tau, v) = x^i v^i + \frac{m}{2}(v^i)^2 \tau$. The same is true for a system of particles subject to a potential which depends only on relative distances among the particles. Equations (7.36) and (7.38) represent the mathematical formulation of the principle of Galilean relativity for the case of boosts: the properties of a given system as studied in laboratories O and O' are the same.

Example 4. Kepler's problem. Consider the action of a particle under a central field

$$S = \int d\tau \left(\frac{m}{2} (\dot{x}^i)^2 - U(r) \right), \quad \text{where} \quad r = (x^i)^2. \tag{7.39}$$

Besides the time translations, symmetries of this action are transformations generated by real orthogonal matrices

$$R : \tau \to \tau' = \tau, \quad x^i \to x'^i = R^{ij} x^j, \quad \text{where} \quad R^T = R^{-1}. \tag{7.40}$$

Its invariance can be immediately verified, in accordance with Eq. (7.29), by substituting x'^i instead of x^i into Eq. (7.39). Notice that the Galileo boosts are not symmetries of the action.

Example 5. A system of two particles labeled by Euclidean coordinates $x^i_{(1)}$, $x^i_{(2)}$, with a potential which depends on the relative distance between them, is described by the action

$$S = \int d\tau \left(\frac{1}{2} m_1 (\dot{x}^i_{(1)})^2 + \frac{1}{2} m_2 (\dot{x}^i_{(2)})^2 - U(r_{12}) \right), \tag{7.41}$$

where $(r_{12})^2 = \sum_{i=1}^3 (x^i_{(2)} - x^i_{(1)})^2$. Besides the time translations, rotations and the Galilean boosts, there is a symmetry under spacial translations with the parameters c^i

$$c : \tau \to \tau' = \tau, \quad x^i_{(a)} \to x'^i_{(a)} = x^i_{(a)} + c^i, \quad a = 1, 2. \tag{7.42}$$

Generalizing, let us write the action

$$S = \int d\tau \left(\frac{1}{2} \sum_{a=1}^{l} m_{(a)} (\dot{x}^i_{(a)})^2 - U(r_{ab}) \right), \tag{7.43}$$

where $(r_{ab})^2 = \sum_{i=1}^3 (x^i_{(b)} - x^i_{(a)})^2$. It describes a system of l particles, $x^1_{(a)}$, $x^2_{(a)}$, $x^3_{(a)}$ which are Euclidean coordinates of a particle with the number a, $a = 1, 2, \ldots, l$. They are under a potential $U(r_{ab})$, which is a function of the variables r_{ab}, $a, b = 1, 2, \ldots, l$. The action is invariant under the 10-parameter *Galileo group*

$$\begin{aligned} \tau &\to \tau' = \tau + a, \\ x^i_{(a)} &\to x'^i_{(a)} = R^{ij} x^j_{(a)} + v^i \tau + c^i. \end{aligned} \tag{7.44}$$

As we have discussed in Sect. 1.2, in classical mechanics it is postulated that the Galileo group relates different inertial frames. The invariance can be verified, in accordance with Eq. (7.29), by substituting τ', x'^i instead of τ, x^i into Eq. (7.43).

7.4 Poincaré Group, Relativistic Particle

As an example of the coordinate transformations of a general form (7.14) (when τ' depends on q^a), we discuss here a free-moving relativistic particle in terms of its physical coordinates.

Let us consider the action functional

$$S = -mc \int dx^0 \sqrt{1 - \left(\frac{dx^a}{dx^0}\right)^2}, \tag{7.45}$$

on the space of functions $x^a = f^a(x^0)$. According to Sect. 1.12.6, it describes a particle which moves along a straight line with constant velocity $(\frac{dx^a}{dt})^2 < c^2$. Let us confirm that the system obeys the principle of special relativity. We need to show that the action admits the Poincaré group as a symmetry group. Invariance under the translations is evident, so, let us discuss the Lorentz transformations

$$\begin{aligned} x^0 \to x'^0 &= \Lambda^0{}_0 x^0 + \Lambda^0{}_b x^b, \\ x^i \to x'^a &= \Lambda^a{}_0 x^0 + \Lambda^a{}_b x^b. \end{aligned} \tag{7.46}$$

Note that they represent an example of coordinate transformations, when transformed time x'^0 depends on spatial coordinates x^a. Starting from a function $x^a = f^a(x^0)$, the transformed function $x^a = f'^a(x^0)$ can be found in a parametric form

$$\begin{aligned} x^0 &= \Lambda^0{}_0 \sigma + \Lambda^0{}_b f^b(\sigma), \\ x^a &= \Lambda^a{}_0 \sigma + \Lambda^a{}_b f^b(\sigma). \end{aligned} \tag{7.47}$$

In the general case, the parameter σ cannot be eliminated from these equations by analytic methods, and we are not able to find closed expression for $f'^a(x^0)$. In other words, it can be said that the *Lorentz group acts on the physical dynamical variables* $f^a(x^0)$ *in a higher non-linear way, in contrast to its linear realization in coordinate space* (7.46). This represents a serious obstacle to the investigation of relativistic theories in terms of physical variables, since the relativistic invariance is not under control.[1] Fortunately, to check an invariance of the action, we do not need to know f'. According to Eq. (7.26), it is sufficient to replace x^μ in Eq. (7.45) by x'^μ given in Eq. (7.46), and to confirm the validity of the condition (7.29).

Exercise
Verify this invariance.

[1] For a free particle, solutions to equations of motion $x^a(x^0)$ are linear functions, and Eq. (7.47) can be resolved; see Example 1. Serious problems arise for particle and field theories with interaction.

Since the description based on the physical variables $f^a(t)$ is not very convenient, let us see what happens, when we try to avoid this problem. Consider the action

$$S = -mc \int d\tau \sqrt{\eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu}, \quad \text{where} \quad \dot{x}^\mu \equiv \frac{dx^\mu}{d\tau}, \tag{7.48}$$

in the space of functions $x^\mu = f^\mu(\tau)$. Now the evolution parameter is τ, while both x^0 and x^a are the-configuration space coordinates. Lorentz transformations are defined in extended space (τ, x^μ) according to $\tau \rightarrow \tau' = \tau$, $x^\mu \rightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu$, and represent symmetry of the action. As compared with the action (7.45), the advantages are:

(A) Invariance of the action is evident, since $\eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu$ is a scalar function with respect to the transformations.
(B) Since the evolution parameter τ is not affected by the transformations, *the transformation law for the function $x^\mu = f^\mu(\tau)$ coincides with the one for the coordinates x^μ: $f'^\mu(\tau) = \Lambda^\mu{}_\nu f^\nu(\tau)$.*

Of course, there is a price to pay. First, the formulation contains an additional variable (actually, it involves two evolution parameters, τ and x^0). Second, the theory is singular: $\det \frac{\partial^2 L}{\partial \dot{x}^\mu \partial \dot{x}^\nu} = 0$. The lesson is that formulation of a relativistic theory in a manifestly Lorentz invariant form (that is with linearly realized Lorentz transformations on dynamical variables) implies a singular action which involves the auxiliary variables.

7.5 Symmetries of Equations of Motion

As before, let g be the coordinate transformation (7.14) (not necessarily a symmetry of an action), and $*g$: $f(\tau) \rightarrow f'(\tau)$ represent the induced transformation (7.18). Consider the equations of motion $F_a(q^a, \dot{q}^a, \ddot{q}^a, \tau) = 0$ following from the action functional (7.13).

Definition g is a symmetry of the equations of motion, if it maps any solution into a solution

$$F_a(f, \ldots) = 0 \quad \Rightarrow \quad F_a(f', \ldots) = 0. \tag{7.49}$$

From a pragmatic point of view, the existence of the symmetry facilitates the search for a general solution to the equations of motion: starting from the known particular solution $q^a = f^a(\tau)$, one immediately obtains a family of solutions applying the transformation $*g$: $q^a = *g \cdot f^a(\tau) = f'^a(\tau, \omega^\alpha)$, which depends on k arbitrary constants ω^α. Sometimes, when the family is large enough, it is sufficient to find only one particular solution to generate the general solution.

As an illustration, consider a free particle $\ddot{x}^i = 0$. The six-parameter transformations $g(\vec{v}, \vec{a}) : \tau \to \tau' = \tau, x^i \to x'^i = x^i + v^i \tau + a^i$ form a symmetry group. In this case, the induced transformations coincide with the coordinate ones. Note that $x^i(\tau) = 0$ is a solution to the equations of motion, then $x'^i = 0 + v^i \tau + a^i$ turns out to be the general solution. Intuitively, a free-moving particle can be obtained from a particle at rest by the Galileo transformation.

As one more example, consider the system $\ddot{x}^i + x^i = 0$, $i = 1, 2$, which admits a symmetry generated by arbitrary non-degenerate matrices

$$a : \tau \to \tau' = \tau, \quad x^i \to x'^i = a^i{}_j x^j, \quad \text{where} \quad \det a \neq 0. \tag{7.50}$$

The general solution $x^1 = A\cos(t + \alpha)$, $x^2 = B\sin(t + \beta)$ can be generated from the particular solution $x^1 = \cos t$, $x^2 = \sin t$ by the application of a symmetry transformation of the form

$$a = \begin{pmatrix} A\cos\alpha & -A\sin\alpha \\ B\cos\beta & -B\sin\beta \end{pmatrix}. \tag{7.51}$$

There are non-trivial applications of this resource, see Sect. 1.12.7, where we computed the electromagnetic field of a moving charge. In a similar way can be obtained complete set of independent solutions to the Dirac equation (describing an electron in the relativistic field theory), see [36].

Exercise
For the second example, find (an invertible) symmetry transformation such that $(\cos t, 0) \to (0, \sin t)$.

Before discussing the relationship between symmetries of an action and symmetries of the corresponding equations of motion, let us consider the following task. Let f' be the image of f under a transformation $*g$. Supposing that f is an extremum of the functional (7.13), let us to find the functional that has f' as an extremum.

Let us suppose that the transformations (7.14) are invertible; see Eq. (7.15). Let us denote the inverse transformation as g^{-1}. Applied to the point τ, q it reads

$$g^{-1} : \tau \to \tau' = \tilde{\alpha}(\tau, q), \qquad q^a \to q'^a = \tilde{\psi}^a(\tau, q). \tag{7.52}$$

This implies transformation of the function f

$$* g^{-1} : f(\tau) \to f''(\tau). \tag{7.53}$$

Exercise
Suppose f is the solution to the equations of motion $F(q, \dot{q}, \ddot{q}, \tau) = 0$. Write equations of motion for f' (see Sect. 2.1.1).

Starting from the action (7.13), the inverse transformation can be equally used to construct a transformed action according to Eq. (7.26). We obtain

$$S_{g^{-1}}[q] = S[f''] = S[*g^{-1}f] = \int d\tau\dot{\tilde{\alpha}}L(\,\tilde{\psi}(\tau,q), (\dot{\tilde{\alpha}})^{-1}\dot{\tilde{\psi}}(\tau,q), \tilde{\alpha}\,). \quad (7.54)$$

Comment It should be mentioned that, if $g \subset G$, the inverse transformation is not generally an element of the family. Nevertheless, if g is a symmetry of the functional (7.13), the same is true for the inverse transformation (prove this by using Definition 1 of Sect. 7.2). If the family is a Lie group, an inverse transformation belongs to the group. Suppose that the transformation (7.14) is parameterized by ω^α, and $\tilde{\omega}^\alpha$ are parameters corresponding to the inverse element. Then, by construction, the functions $\tilde{\alpha}$, $\tilde{\psi}$ are the group functions: $\tilde{\alpha}(\omega) = \alpha(\tilde{\omega})$, $\tilde{\psi}(\omega) = \psi(\tilde{\omega})$. The action (7.54) then simply coincides with (7.26), where $\omega \to \tilde{\omega}$.

Computing $S_{g^{-1}}$ on f', we have $S_{g^{-1}}[f'] = S[*g^{-1} * gf] = S[f]$. So, if f' is the image of a function f under the transformation (7.14), then

$$S_{g^{-1}}[f'] = S[f]. \quad (7.55)$$

This resolves the task formulated above: if f is an extremum of S, then f' will be an extremum of $S_{g^{-1}}$. In other words, if f represents a solution to the equations of motion following from S, then f' obeys equations of motion obtained from $S_{g^{-1}}$.

This allows us to demonstrate the following remarkable fact: transformations leaving an action invariant, map solutions to equations of motion into solutions. It is worth noting that the inverse statement is not true.

Affirmation If the family G is a symmetry of the functional $S[q] = \int d\tau L(q^a, \dot{q}^a, \tau)$, then G is a symmetry of the corresponding equations of motion.

Proof According to the comment made above, together with g, the transformation g^{-1} represents a symmetry of the action. The invariance condition for g^{-1} reads $S[*g^{-1}f] = S[f] + \int d\tau\dot{N}$, or, equivalently

$$S_{g^{-1}}[q] = S[q] + \int d\tau\dot{N} \quad (7.56)$$

Let f be a solution to the equations of motion deriving from S. According to Eq. (7.55), f'^a is a solution to equations of motion deriving from $S_{g^{-1}}[q]$. According to (7.56), the equations simply coincide with those obtained from $S[q]$. Thus f and f' obey the same equation.

To confirm that the statement cannot be inverted, it is sufficient to return to the second example of the previous section. Equations $\ddot{x}^i + x^i = 0$ follow from the action $S = S\int d\tau((\dot{x}^i)^2 - (x^i)^2)$. The symmetry (7.50) of the equations is not a symmetry of the action (unless the matrix a is orthogonal).

7.6 Noether Theorem

We present here the Noether theorem in the form normally used by physicists.[2] Let G be a k-parameter family of coordinate transformations

$$\begin{aligned}
\tau \to \tau' = \alpha(\tau, q^a, \omega^\alpha) = \tau + G_\alpha(\tau, q^a)\omega^\alpha + O(\omega^2), \\
G_\alpha \equiv \left.\frac{\partial\alpha}{\partial\omega^\alpha}\right|_{\omega=0}, \\
q^a \to q'^a = \psi^a(\tau, q^a, \omega^\alpha) = q^a + R^a{}_\alpha(\tau, q^a)\omega^\alpha + O(\omega^2), \\
R^a{}_\alpha \equiv \left.\frac{\partial\psi^a}{\partial\omega^\alpha}\right|_{\omega=0}.
\end{aligned} \tag{7.57}$$

Here, with use of Eq. (7.16), the transition functions have been expanded, up to a linear order, in a power series at $\omega = 0$. So, *infinitesimal transformations* ($\omega \ll 1$) are characterized by the functions G and R, called *generators of the transformations*. We combine them into the quantity[3]

$$D^a{}_\alpha \equiv R^a{}_\alpha - \dot{q}^a G_\alpha, \tag{7.58}$$

and impose the following technical condition

$$\text{rank}\, D^a{}_\alpha = [\alpha] = k, \tag{7.59}$$

A family of transformations with this property is called a *family with k essential parameters*.

Noether Theorem. Let the action (7.13) be invariant under the family of transformations (7.57) with k essential parameters, that is

$$\int_{\tau_1}^{\tau_2} d\tau\dot{\alpha}L(\psi, (\dot{\alpha})^{-1}\dot{\psi}, \alpha) = \int_{\tau_1}^{\tau_2} d\tau \left(L(q, \dot{q}, \tau) + \frac{dN(q, \dot{q}, \tau)}{d\tau}\right). \tag{7.60}$$

Then there are k functions $Q_\alpha(q, \dot{q}, \tau)$ called *Noether charges*, namely

$$Q_\alpha = -\frac{\partial L}{\partial\dot{q}^a}(R^a{}_\alpha - \dot{q}^a G_\alpha) - LG_\alpha + N_\alpha, \quad N_\alpha \equiv \left.\frac{\partial N}{\partial\omega^\alpha}\right|_{\omega=0}, \tag{7.61}$$

which retain a constant value throughout any solution to equations of motion

$$\left.\frac{dQ_\alpha}{d\tau}\right|_{\frac{\delta S}{\delta q}=0} = 0. \tag{7.62}$$

---

[2] See [37] for discussion of the most general form of the Noether theorem.

[3] As we will see below, D determines an infinitesimal transformation of a function.

For the case of nonsingular theory the charges do not vanish identically. Moreover, they are functionally independent: rank $\frac{\partial Q}{\partial(q,\dot{q})} = k$.

Comment The Noether theorem gives the charges in terms of generators. It is possible to write an inverse formula for the generators through a given conserved charges, see Eq. (7.141) below.

Proof As has been discussed in Sect. 7.2, the integrals in Eq. (7.60) can be omitted. Further, the integrands can be expanded in a power series of ω. Since ω are arbitrary parameters, the identity (7.60) must be satisfied for each power order separately. The result we are interested in appears in the linear order. Let $A(\omega) = B(\omega)$ be a symbolic notation for the integrand of Eq. (7.60). Then the linear on the ω part is $\frac{\partial A}{\partial\omega^\alpha}\Big|_{\omega=0} = \frac{\partial B}{\partial\omega^\alpha}\Big|_{\omega=0}$. Let us write an explicit form of this expression. The right-hand side is

$$\frac{d}{d\tau}\frac{\partial N}{\partial\omega^\alpha}\bigg|_{\omega=0} \equiv \frac{d}{d\tau}N_\alpha, \tag{7.63}$$

with the known function N_α. With the use of Eqs. (7.57) and (7.16), the derivative of the left-hand side is

$$\frac{\partial(\text{l.h.s.})}{\partial\omega^\alpha}\bigg| = L\dot{G}_\alpha + \frac{\partial L}{\partial q^a}R^a{}_\alpha - \frac{\partial L}{\partial\dot{q}^a}\dot{G}_\alpha\dot{q}^a + \frac{\partial L}{\partial\dot{q}^a}\dot{R}^a{}_\alpha + \frac{\partial L}{\partial\tau}G_\alpha \tag{7.64}$$

Here $L \equiv L(q, \dot{q}, \tau)$. Extracting a total derivative with respect to τ from the first and fourth terms, we write

$$\left(LG_\alpha + \frac{\partial L}{\partial\dot{q}^a}R^a{}_\alpha\right)^{\cdot} - \frac{\partial L}{\partial q^a}G_\alpha\dot{q}^a - \frac{\partial L}{\partial\dot{q}^a}(G_\alpha\dot{q}^a)^{\cdot} + R^a{}_\alpha\frac{\delta S}{\delta q^a}. \tag{7.65}$$

Further, extracting a total derivative from the third term in (7.65) we obtain

$$\left(LG_\alpha + \frac{\partial L}{\partial\dot{q}^a}(R^a{}_\alpha - \dot{q}^aG_\alpha)\right)^{\cdot} + (R^a{}_\alpha - \dot{q}^aG_\alpha)\frac{\delta S}{\delta q^a}. \tag{7.66}$$

Combining the expression (7.63) and this last one, the linear with respect to ω part of Eq. (7.60) acquires the form

$$(R^a{}_\alpha - \dot{q}^aG_\alpha)\frac{\delta S}{\delta q^a} = \frac{dQ_\alpha}{d\tau}, \quad \text{forany} \quad q^a(\tau), \tag{7.67}$$

with Q given by Eq. (7.61). Let us stress once again that this equality is an identity, that is, it is true for any function $q^a(\tau)$. So, invariance of an action implies that some combinations of the equations of motion form total derivatives of the charges Q_α. The equations (7.67) are called *Noether identities*. We discuss below how the Noether identities can be used to simplify the equations of motion.

The Noether theorem follows immediately from Eq. (7.67): when the equations of motion holds, $\frac{\delta S}{\delta q^a} = 0$, one has $\frac{dQ_\alpha}{d\tau} = 0$. The charges Q^α do not vanish identically; besides that, they are linearly independent. Indeed, suppose, for example, $Q_1 = 0$ for any $q(\tau)$. Then the identity (7.67) acquires the form (see (2.16)) $D_1^a(M_{ab}\ddot{q}^b - K_a) = 0$. This implies that the matrix M has the null-vector D_1, which contradicts the nonsingular character of the theory.[4] Similarly, linear dependence of Q_α would contradict the condition (7.59). Functional independence of the charges will be demonstrated in Sect. 7.9.

Exercises

1. Confirm that $N|_{\omega=0} = 0$.
2. Work out the quadratic term of the power expansion of Eq. (7.60). What is the information contained in it?

7.7 Infinitesimal Symmetries

The reader possibly observed that only the linear part of the power expansion was used in the proof of the Noether theorem (we will return to the discussion of that point in the next section). This justifies the notions of infinitesimal transformation and infinitesimal symmetry discussed here.

The linear on ω^α coordinate transformation

$$\begin{aligned} \tau \to \tau' &= \tau + G_\alpha(\tau, q^a)\omega^\alpha \equiv \tau + \delta\tau \\ q^a \to q'^a &= q^a + R^a{}_\alpha(\tau, q^a)\omega^\alpha \equiv q^a + \delta q^a, \end{aligned} \tag{7.68}$$

is called *infinitesimal transformation*. The functions G_α, $R^a{}_\alpha$ are called generators. With any coordinate transformation can be associated an infinitesimal transformation. It is obtained keeping the first two terms of the power expansion, two see Eq. (7.57).

Similarly to the general case discussed in Sect. 7.2, the infinitesimal transformation (7.68) induces transformation of functions, $*g : q^a = f^a(\tau) \to q^a = f'^a(\tau)$, where $f'^a(\tau)$ is given in parametric form by

$$\begin{aligned} \tau &= \sigma + G_\alpha(\sigma, f^a(\sigma))\omega^\alpha \\ q^a &= f^a(\sigma) + R^a{}_\alpha(\sigma, f^a(\sigma))\omega^\alpha, \end{aligned} \tag{7.69}$$

According to Eq. (7.18) this implies

[4] Note that in singular theory it can happen that $Q \equiv 0$, which implies identities among the equations of motion. This is closely related with the presence of local symmetries, see Chap. 8.

$$f'^a(\tau) = f^a(\tilde{\alpha}(\tau)) + R^a{}_\alpha(\tilde{\alpha}(\tau), f(\tilde{\alpha}(\tau))\omega^\alpha, \tag{7.70}$$

where $\tilde{\alpha}(\tau)$ is the inverse function for the following function of τ: $\tau + G_\alpha(\tau, f(\tau))\omega^\alpha$. In the linear approximation it reads $\tilde{\alpha} = \tau - G_\alpha(\tau, f(\tau))\omega^\alpha + O^2(\omega)$. This implies

$$\left.\frac{\partial \tilde{\alpha}}{\partial \omega^\alpha}\right|_{\omega=0} = -G_\alpha. \tag{7.71}$$

In the linear approximation, there is a simple formula for transformed function in terms of the initial one. Using Eqs. (7.70) and (7.71) this reads

$$\begin{aligned} f'^a(\tau) &= f'^a(\tau)\big|_{\omega=0} + \left.\frac{\partial f'^a(\tau)}{\partial \omega^\alpha}\right|_{\omega=0} \omega^\alpha + O^2(\omega) \\ &= f^a(\tau) + \left(R^a{}_\alpha(\tau, f(\tau)) + \dot{f}^a(\tau) \left.\frac{\partial \tilde{\alpha}}{\partial \omega^\alpha}\right|_{\omega=0} \right) \omega^\alpha + O^2(\omega) \\ &= f^a(\tau) + \left(R^a{}_\alpha(\tau, f(\tau)) - \dot{f}^a(\tau) G_\alpha(\tau, f(\tau))\right) \omega^\alpha + O^2(\omega). \end{aligned} \tag{7.72}$$

we come back to our notation $f(\tau) = q(\tau)$, then Eq. (7.72) implies

$$q^a(\tau) \to q'^a(\tau) = q^a(\tau) + \tilde{\delta} q^a, \quad \tilde{\delta} q^a \equiv \delta q^a - \dot{q}^a \delta\tau \equiv D^a \alpha \omega^\alpha, \tag{7.73}$$

with $\delta\tau, \delta q^a$ specified in (7.68), and D is given by (7.58). The symbol $\tilde{\delta}$ is used for variation of a function. In the books on quantum field theory it is called an *infinitesimal transformation of form of a field (function).*

An infinitesimal transformation (7.68) is an *infinitesimal symmetry* of an action, if

$$\left.\int_{\tau_1 + G_\alpha(\tau_1, f(\tau_1))\omega^\alpha}^{\tau_2 + G_\alpha(\tau_2, f(\tau_2))\omega^\alpha} d\tau L(f'^a, \dot{f}'^a, \tau)\right|_{O(\omega)} = \int_{\tau_1}^{\tau_2} d\tau \left[L(f^a, \dot{f}^a, \tau) + \frac{dN}{d\tau} \right]. \tag{7.74}$$

Here f' is given by Eq. (7.70), and the notation $O(\omega)$ means that we keep only the linear term of the expansion in power series of ω around $\omega = 0$. As compared with the symmetry condition (7.22), we now require it to be satisfied only in linear order with respect to ω. So each symmetry (7.14) implies an infinitesimal symmetry, which appears as linear in the ω_α part of the symmetry transformation. Note that an infinitesimal symmetry *is not* generally a symmetry.

Let us compute the left-hand side of Eq. (7.74). First, take into account integration limits that give an extra total derivative term

$$\int_{\tau_1}^{\tau_2} d\tau \left[L(f'^a, \dot{f}'^a, \tau)\big|_{O(\omega)} + \frac{d}{d\tau}(LG_\alpha \omega^\alpha) \right]$$
$$= \int_{\tau_1}^{\tau_2} d\tau \left[L(f^a, \dot{f}^a, \tau) + \frac{dN}{d\tau} \right]. \quad (7.75)$$

Second, substitute (7.73) and omit the integration, thus obtaining

$$L(q^a + \tilde{\delta} q^a, (q^a + \tilde{\delta} q^a)^{\cdot}, \tau)\Big|_{O(\tilde{\delta} q)} - L(q, \dot{q}, \tau) = \frac{d}{d\tau}(N - LG_\alpha \omega^\alpha). \quad (7.76)$$

Note that the left-hand side is just the usual variation of a Lagrangian due to the variation of coordinates $\tilde{\delta} q^a$. Hence the invariance condition (7.74) is equivalent to the statement that the variation can be presented as a total derivative of some function. Computing the variation according to the known formula (1.105), we reproduce the Noether identities (7.67)

$$\frac{\delta S}{\delta q^a} \tilde{\delta} q^a = \frac{d}{d\tau} \left(\frac{\partial L}{\partial \dot{q}^a} \tilde{\delta} q^a - L \delta \tau + N \right). \quad (7.77)$$

Remember that the invariance condition (7.22) is equivalent to Eq. (7.30). The infinitesimal invariance condition (7.74) can also be written in a similar form; it is sufficient to replace $\alpha \to \tau + \delta\tau$, $\psi \to q + \delta q$ in Eq. (7.30) and keep only the linear part. It reads

$$(1 + (\delta\tau)^{\cdot}) L(q + \delta q, (q + \delta q)^{\cdot}, \tau + \delta\tau)|_{O(\delta\tau, \delta q)} - L(q, \dot{q}, \tau) = \frac{dN}{d\tau}. \quad (7.78)$$

Computing the linear part, we arrive at the Noether identities once again (do the computation!).

As in Sect. 7.2, we can generalize the notion of infinitesimal symmetry allowing the generator R in the transformation law of a function (Eq. (7.70)) to depend on time derivatives of f.

Example An infinitesimal transformation with $\delta\tau = 0$, $\delta q^a = B^{ab} \frac{\delta S}{\delta q^b}$, where B^{ab} is an antisymmetric matrix, represents an infinitesimal symmetry of any action. In fact, in this case, the variation of the Lagrangian reads (omitting a total derivative) $\delta L = \frac{\delta S}{\delta q^b} B^{ab} \frac{\delta S}{\delta q^b} = 0$. This is called a *trivial infinitesimal symmetry*. Being present in any action, the trivial symmetry does not lead to physical consequences. In particular, the corresponding charge vanishes on equations of motion.

In short, we have obtained two formal recipes allowing us to check whether an infinitesimal transformation (7.68) represents an infinitesimal symmetry of an action functional. We can compute the usual variation of the action under transformations of the form (7.73) and see whether it can be presented as a total derivative of some function. Equivalently, we can see whether the left-hand side of Eq. (7.78) forms a total derivative.

We have also demonstrated that the infinitesimal symmetry condition (7.74) is equivalent to the Noether identities (7.77) and thus leads to the conserved charges.

7.8 Discussion of the Noether Theorem

Let us return to the discussion of the Noether theorem. A brief inspection of Sect. 7.6 shows that invariance of an action under the symmetry transformations (7.57) is not necessary in the proof of the Noether theorem. Since only the linear on ω_α part of Eq. (7.60) has been used in the proof, the conservation law is already guaranteed by the infinitesimal symmetry. This was shown once again in the previous section: the infinitesimal symmetry condition (7.74) is equivalent to the Noether identities (7.77), which imply conservation of the charges, Eq. (7.62).

As we have discussed, symmetries of an action reflect fundamental properties assumed for our space-time. Since an infinitesimal symmetry is not generally a symmetry, the question arises whether such properties as homogeneity, isotropy, and so on are actually related to the existence of conservation laws. Here we fill this gap by showing that *any infinitesimal symmetry of an action implies its invariance under certain symmetry transformation*.

Remember that, due to the identity $e^{a(x)\partial_x}G(x) = G\left(e^{a(x)\partial_x}x\right)$, the function $f(\epsilon, x) = e^{\epsilon a(x)\partial_x}x$ is a formal solution to the equation (see Sect. 2.3)

$$\frac{\partial f}{\partial \epsilon} = a(f). \tag{7.79}$$

Besides, it obeys the initial condition $f(0, x) = x$.

Now, starting from the infinitesimal symmetry (7.68), construct the functions

$$\begin{aligned} \alpha(\tau, q^\alpha, \epsilon\omega^\alpha) &= e^{G_\alpha \epsilon\omega^\alpha \partial_\tau + R^a{}_\alpha \epsilon\omega^\alpha \partial_a}\tau, \\ \psi^a(\tau, q^\alpha, \epsilon\omega^\alpha) &= e^{G_\alpha \epsilon\omega^\alpha \partial_\tau + R^a{}_\alpha \epsilon\omega^\alpha \partial_a}q^a. \end{aligned} \tag{7.80}$$

These obey the multi-variable generalization of the Eq. (7.79)

$$\begin{aligned} \frac{\partial \alpha}{\partial \epsilon} &= G_\alpha(\alpha, \psi^a)\omega^\alpha, \\ \frac{\partial \psi^a}{\partial \epsilon} &= R^a{}_\alpha(\alpha, \psi^a)\omega^\alpha, \end{aligned} \tag{7.81}$$

as well as the initial conditions $\alpha(\epsilon=0)=\tau$, $\psi^a(\epsilon=0)=q^a$.

We show that the transformation

$$g(\omega): \quad \tau \to \tau' = \alpha(\tau, q^\alpha, \omega^\alpha), \quad q^a \to q'^a = \psi^a(\tau, q^\alpha, \omega^\alpha), \tag{7.82}$$

represents a symmetry of the action (7.13).

To see this, we construct the function $S(\epsilon)$ transforming the action (7.26) by means of (7.80)

$$S(\epsilon) = S_{g(\epsilon\omega)}[q] = \int_{\tau_1}^{\tau_2} d\tau \dot{\alpha} L(\psi, (\dot{\alpha})^{-1}\dot{\psi}, \alpha). \tag{7.83}$$

Note that $S(0)=S[q]$, while $S(1)=S_{g(\omega)}[q]$ is transformation of the action (7.26) under (7.82). Using Eq. (7.81) we compute

$$\frac{\partial S(\epsilon)}{\partial \epsilon} = \int d\tau \left(L|(G_\alpha|)^{\dot{}} + \dot{\alpha}\frac{\partial L}{\partial q^a} R^a{}_\alpha| + \right.$$
$$\left. \dot{\alpha}\frac{\partial L}{\partial \dot{q}^a}| \left(-(\dot{\alpha})^{-2}(G_\alpha|)^{\dot{}}\dot{\psi}^a + (\dot{\alpha})^{-1}(R^a{}_\alpha|)^{\dot{}} \right) + \dot{\alpha}\frac{\partial L}{\partial \tau}G_\alpha| \right) \omega^\alpha. \tag{7.84}$$

Here the notation $A(\tau, q)|$ implies substitution of the functions (7.80) instead of τ, q. Integrating by parts (the computation is similar to the one leading to Eq. (7.67)), we obtain

$$\frac{\partial S(\epsilon)}{\partial \epsilon} = \int d\tau \left(-\frac{dQ_\alpha|}{d\tau} + \dot{\alpha}(R^a{}_\alpha - \dot{q}^a G_\alpha)\frac{\delta S}{\delta q^a}| + (N_\alpha|)^{\dot{}} \right) \omega^\alpha. \tag{7.85}$$

Since the transformation (7.68) is an infinitesimal symmetry, the first two terms cancel each other (see Eqs. (7.24) and (7.67)), and we have

$$\frac{\partial S(\epsilon)}{\partial \epsilon} = \int d\tau \frac{d}{d\tau} N_\alpha | \omega^\alpha. \tag{7.86}$$

Computing the integral of this expression with respect to ϵ on the interval [0, 1], we arrive at the invariance condition (7.60).

7.9 Use of Noether Charges for Reduction of the Order of Equations of Motion

As we have seen, the invariance of an action functional implies a special structure of the corresponding equations of motion. This is given by Eq. (7.67). Since some combinations of the equations of motion are total derivatives, they can immediately be integrated, which simplifies (sometimes solves) the problem to find a general solution. Let us discuss this point in further detail. We demonstrate that knowledge of k Noether charges allows us to replace the initial system of n second-order Euler-Lagrange equations by an equivalent system, which has only $n-k$ second-order equations.

According to the Noether theorem, the equations $\dot{Q}_\alpha = 0$ are consequences of equations of motion. So, they can be added to the system, which gives an equivalent system

$$\frac{\delta S}{\delta q^a} = 0, \qquad \dot{Q}_\alpha = 0. \tag{7.87}$$

Now some k among the Euler–Lagrange equations are consequences of others, and thus can be omitted. Indeed, Eq. (7.67) states that there are identities present among equations of the system (7.87). Suppose that a minor rank of the matrix D is placed in the upper k lines (if not, this can be achieved by reordering the initial variables q^a in an appropriate way). Thus we write: $D^a_\alpha = (D^\beta_\alpha, D^i_\alpha)$, $\det D^\beta_\alpha|_{\frac{\delta S}{\delta q}=0} \neq 0$. Then the identity (7.67) can be written in the form ($\tilde{D}$ is the inverse matrix for D^γ_α)

$$\frac{\delta S}{\delta q^\alpha} = \tilde{D}^\gamma_\alpha (\dot{Q}_\gamma - D^i_\gamma \frac{\delta S}{\delta q^i}), \tag{7.88}$$

that is, the equations $\frac{\delta S}{\delta q^\alpha} = 0$ are consequences of other equations of the system (7.87). Then the initial system is equivalent to

$$\begin{aligned} &\frac{\delta S}{\delta q^i} = 0, && i = 1, 2, \ldots, n-k, \\ &Q_\alpha(q, \dot{q}) = c_\alpha = const, && \alpha = 1, 2, \ldots, k. \end{aligned} \tag{7.89}$$

This contains $n-k$ second-order equations and k first-order equations, that is, the order has been decreased by k units.

Q_α are functionally independent, otherwise some of them could be omitted from the system (7.89). Then we should have a system with a number of equations less than the number of variables. But this would contradict the theorem of the existence of a unique solution for normal system of equations.

7.10 Examples

In Sect. 7.3, it was mentioned that the Lagrangian action for a system of particles subject to distance-dependent forces is invariant under a ten-parameter Galileo group, which includes translations, rotations and boosts. Here we write the corresponding Noether charges. To put this in concrete terms, we consider two particles with Euclidean coordinates $x^i_{(1)}, x^i_{(2)}$, $i = 1, 2, 3$

$$S = \int d\tau \left(\frac{1}{2} m_1 (\dot{x}^i_{(1)})^2 + \frac{1}{2} m_2 (\dot{x}^i_{(2)})^2 - U(r_{12}) \right). \tag{7.90}$$

In this case, the expression for the Noether charge (7.61) is

$$Q_\alpha = -m_1 \dot{x}^i_{(a)} \left[R^i_{(a)\alpha} - \dot{x}^i_{(a)} G_\alpha \right] - L G_\alpha + N_\alpha. \tag{7.91}$$

Example 1. For the time translations $\tau' = \tau + a$, $x'^i_{(a)} = x^i_{(a)}$, we have $G = 1$, $R^i_{(a)} = 0$, $N = 0$, and the Noether charge is the *total energy* of the system

$$E = \frac{1}{2} m_1 (\dot{x}^i_{(1)})^2 + \frac{1}{2} m_2 (\dot{x}^i_{(2)})^2 + U(r_{12}). \tag{7.92}$$

Intuitively, homogeneity in time implies conservation of the total energy of a closed system.

Example 2. For the space translations $\tau' = \tau$, $x'^i_{(a)} = x^i_{(a)} - c^i$, we have $G = 0$, $R^i_{(a)j} = \delta^i{}_j$, $N_i = 0$, which leads to conservation of *total momentum*

$$P^i = m_1 \dot{x}^i_{(1)} + m_2 \dot{x}^i_{(2)} = p^i_{(1)} + p^i_{(2)}. \tag{7.93}$$

Thus conservation of the total momentum is a consequence of homogeneity of space. In the present case, the total momentum turns out to be the sum of the conjugated momenta of the particles. Note also, that the individual momenta are not conserved during the evolution, as soon as $\dot{p}^i_{(a)} = \frac{\partial U}{\partial x^i_{(a)}} \neq 0$.

Example 3. For the Galileo boosts $\tau' = \tau$, $x'^i_{(a)} = x^i_{(a)} + v^i \tau$, we have $G = 0$, $R^i_{(a)j} = \tau \delta^i{}_j$, $N = (m_1 x^i_{(1)} + m_2 x^i_{(2)}) v^i + \frac{1}{2}(m_1 + m_2)(v^i)^2 \tau$, $N_i = m_1 x^i_{(1)} + m_2 x^i_{(2)}$, which gives the conserved charge

$$-(m_1 \dot{x}^i_{(1)} + m_2 \dot{x}^i_{(2)}) \tau + m_1 x^i_{(1)} + m_2 x^i_{(2)} = D^i = \text{const.} \tag{7.94}$$

Let us write this in the form

$$m_1 x^i_{(1)} + m_2 x^i_{(2)} = D^i + P^i \tau. \tag{7.95}$$

We conclude that during the evolution of two particles, the point[5] $X^i = \frac{1}{m_1 + m_2}(m_1 x^i_{(1)} + m_2 x^i_{(2)})$ moves along a straight line with a velocity proportional to the total momentum. In other words, from the Noether theorem we have discovered the concept of *center of mass of a system*. Since the dynamics of X^i is already known, the convenient coordinates for description of the particles are X^i and, for example, $Y^i = \frac{1}{m_1 + m_2}(m_1 x^i_{(1)} - m_2 x^i_{(2)})$.

[5] It is reasonable to divide by $m_1 + m_2$, then X has the dimension of a length.

Example 4. Consider the rotations (see Sect. 1.2)

$$\tau' = \tau, \quad x'^i_{(a)} = \left(e^{\omega}\right)^{ij} x^j_{(a)} = x^i_{(a)} + \omega^{ij} x^j_{(a)} + O(\omega^2), \tag{7.96}$$

where $\omega^{ij} = -\omega^{ji}$ are three independent parameters. We have $G = 0$, $N_\alpha = 0$. To find the generators R, we need to represent Eq. (7.96) in the form $x'^i_{(a)} = x^i + R^i_{12}\omega^{12} + R^i_{13}\omega^{13} + R^i_{23}\omega^{23}$, and then to find an explicit form of these nine quantities R^i_{jk}. To avoid the problem,[6] we look for a sum of the charges, and then separate them

$$Q_\alpha \omega^\alpha = -\frac{\partial L}{\partial \dot{q}^a} R^a{}_\alpha \omega^\alpha = -\sum_{i,j=1}^{3} \left(\frac{\partial L}{\partial \dot{x}^i_{(1)}} \omega^{ij} x^j_{(1)} + \frac{\partial L}{\partial \dot{x}^i_{(2)}} \omega^{ij} x^j_{(2)} \right) =$$
$$\sum_{a=1}^{2} \Big[\left(x^1_{(a)} p^2_{(a)} - x^2_{(a)} p^1_{(a)} \right) \omega^{12} + \left(x^2_{(a)} p^3_{(a)} - x^3_{(a)} p^2_{(a)} \right) \omega^{23} +$$
$$\left(x^1_{(a)} p^3_{(a)} - x^3_{(a)} p^1_{(a)} \right) \omega^{13} \Big]. \tag{7.97}$$

Thus, the isotropy of a space implies three charges which turn out to be components of the *total angular momentum* of the system

$$J^i = \sum_{a=1}^{2} \epsilon^{ijk} x^j_{(a)} p^k_{(a)}. \tag{7.98}$$

Summing up, we have seen that conservation of energy, momentum and angular momentum can be thought as a consequence of the homogeneity of time and space and isotropy of space.

Exercises

1. Check by direct computations that the infinitesimal rotations$x'^i_{(a)} = x^i_{(a)} + \omega^{ij} x^j_{(a)}$ represent symmetry of the action (7.90).
2. Check by direct computations a preservation of the charges obtained inthis section.
3. Check whether the angular momentum of each particle separately is acon-served quantity.
4. Warning exercise. Try to find the charges (7.97) directly, repeating the computation carried out in the proof of the Noether theorem (that is, by extracting the terms of the transformed action that are linear on ω^{ij}).

[6] Of course, the problem can easily be solved. We write $\omega^{ij} x^j \equiv \frac{1}{2}(\delta^{ik} x^j - \delta^{ij} x^k)\omega^{kj}$. Then the quantities $R^i_{kj} = \frac{1}{2}(\delta^{ik} x^j - \delta^{ij} x^k)$ with $k < j$ represent the generators.

5. Find the Noether charges of the relativistic particle in the formulations (7.45) and (7.48).
6. Verify that the action $S = \int d\tau \frac{1}{2} g_{ab} \dot{y}^a \dot{y}^b$, $a = 1, 2$, where $g_{ab} \equiv \delta_{ab} + (l^2 - y^2)^{-1} y^a y^b$, $l =$ const, has infinitesimal symmetry with the parameters ϵ^{ab}, c^a

$$y'^a = y^a + \epsilon^{ab} y^b + (l^2 - y^2)^{\frac{1}{2}} c^a, \tag{7.99}$$

where $\epsilon^{12} = -\epsilon^{21}$, $\epsilon^{aa} = 0$. Derive the corresponding Noether charges. Verify by direct computations that they are actually conserved.

7.11 Symmetries of Hamiltonian Action

It has already been mentioned that the notion of symmetry and the Noether theorem machinery can be applied to any system of differential equations arising from a variational problem. So it is true for Hamiltonian equations as they follow from the Hamiltonian action functional

$$S_H = \int_{\tau_1}^{\tau_2} d\tau (p_a \dot{q}^a - H(q^a, p_a, \tau)). \tag{7.100}$$

The basic notions of the previous sections remain true for this case, with the corresponding replacements: $q^a \to z^i = (q^a, p_b)$, $L \to p\dot{q} - H(q, p)$. Below we present this reformulation. It is interesting for the following reasons.

(a) The formalism can be applied to Hamiltonian systems that do not admit a Lagrangian formulation. We have already discussed the Schrödinger equation as an example of this kind; see Sect. 2.10.
(b) Due to certain specific properties of a phase space, we can go further in the analysis of Hamiltonian symmetries as compared with Lagrangian ones. Moreover, this allows us to obtain additional information on Lagrangian symmetries, giving an expression for Lagrangian symmetry in terms of the Noether charge; see Eq. (7.141) below.
(c) The tool developed below can be applied to much more complicated case of *local* symmetries, see Chap. 8.

7.11.1 Infinitesimal Symmetries Given by Canonical Transformations

Here we obtain the necessary and sufficient condition for an infinitesimal canonical transformation to be an infinitesimal symmetry. In the next section we analyze the structure of infinitesimal symmetry of a general form. It will be shown that a general infinitesimal symmetry is equivalent to an infinitesimal symmetry given by canonical transformation. Hence, looking for infinitesimal symmetries of S_H we can restrict our scope and look for them in the class of canonical transformations.

In the extended phase space $(\tau, z^i) = (\tau, q^a, p_b)$ consider an infinitesimal canonical transformation

$$\delta\tau = 0, \qquad \delta z^i = -\{z^i, Q\}, \tag{7.101}$$

with the generator $Q = Q_\alpha(z,\tau)\omega^\alpha$. Notice that instead of Q one can equally take $Q' = Q + h_\alpha(\tau)\omega^\alpha$ where h_α stands for an arbitrary z-independent function. The generators Q and Q' are equivalent since they produce the same transformation.

Assertion A canonical transformation (7.101) is an infinitesimal symmetry of the Hamiltonian action if and only if its generator obeys the equation

$$\frac{\partial Q}{\partial\tau} + \{Q, H\} = 0, \quad \text{for any} \quad z(\tau). \tag{7.102}$$

Besides, the generator represents a conserved charge, that is $\frac{dQ}{d\tau} = 0$ on equations of motion.

Proof As $\delta\tau = 0$, the induced transformation of a function has the same form as δz^i. According to Sect. 7.7, δz^i will be a symmetry of S_H, if its variation forms a total derivative. The variation reads (see Eq. (2.127))

$$\begin{aligned} \delta S_H &= \int d\tau \left[(\dot p_a - \{p_a, H\})\delta q^a - (\dot q^a - \{q^a, H\})\delta p_a + (p_a\delta q^a)^{\cdot}\right] \\ &= \int d\tau \left[-\frac{\partial Q}{\partial\tau} - \{Q, H\} + \frac{d}{d\tau}(Q - p_a\{q^a, Q\})\right]. \end{aligned} \tag{7.103}$$

If Q satisfies Eq. (7.102), the variation is a total derivative, so Eq. (7.101) represents a symmetry. Conversely, suppose that $\delta S_H = \int d\tau \frac{dN}{d\tau}$, where $N = N_\alpha\omega^\alpha$. Then Eq. (7.103) implies the Noether identity

$$(\dot p_a - \{p_a, H\})\delta q^a - (\dot q^a - \{q^a, H\})\delta p_a = \frac{d}{d\tau}(-p_a\delta q^a + N). \tag{7.104}$$

Computing the derivative with respect to τ which appears on the right-hand side of Eq. (7.104), the latter reads

$$\begin{aligned} &\dot q^a \left(-\delta p_a + p_b\frac{\partial\delta q^b}{\partial q^a} - \frac{\partial N}{\partial q^a}\right) + \dot p_a \left(2\delta q^a + p_b\frac{\partial\delta q^b}{\partial p_a} - \frac{\partial N}{\partial p_a}\right) \\ &\qquad + \frac{\partial}{\partial\tau}(p_a\delta q^a - N) + \{Q, H\} = 0. \end{aligned} \tag{7.105}$$

Since this is true for any functions $q(\tau)$, $p(\tau)$, the terms in front of $\dot{q}$, $\dot{p}$ vanish separately. So Eq. (7.105) decomposes into three identities. Taking into account the expression (7.101) for δz, they can be written in the form

$$\{q^a, p_b\delta q^b - N - Q\} = 0, \qquad \{p_a, p_b\delta q^b - N - Q\} = 0, \tag{7.106}$$

$$\frac{\partial}{\partial\tau}(p_a\delta q^a - N) + \{Q, H\} = 0. \tag{7.107}$$

Equation (7.106) implies $p_b\delta q^b - N - Q = h(\tau)$. If $h \neq 0$, we replace the initial Q with an equivalent generator $Q' = Q + h$. Then equation (7.107) states that Q' satisfies the Eq. (7.102).

To see that Q represents a conserved charge, notice that Eq. (7.102) can be identically rewritten as

$$\frac{dQ}{d\tau} = \frac{\partial Q}{\partial q^a}(\dot{q}^a - \{q^a, H\}) + \frac{\partial Q}{\partial p_a}(\dot{p}_a - \{p_a, H\}), \tag{7.108}$$

so $\frac{dQ}{d\tau} = 0$ on equations of motion.

7.11.2 Structure of Infinitesimal Symmetry of a General Form

The general form of an infinitesimal transformation with k essential parameters ω^α is

$$\begin{aligned} \tau \to \tau' &= \tau + G_\alpha(\tau, z^i)\omega^\alpha \equiv \tau + \delta\tau, \\ z^i \to z'^i &= z^i + R^i{}_\alpha(\tau, z^i)\omega^\alpha \equiv z^i + \delta z^i. \end{aligned} \tag{7.109}$$

The generator R contains two blocks: $R^i{}_\alpha = (R^a{}_\alpha, T_{b\alpha})$, where $R^a{}_\alpha$ corresponds to q^a and $T_{b\alpha}$ corresponds to p_b.

The coordinate transformation induces the transformation of a function, $z^i = f^i(\tau) \to z^i = f'^i(\tau)$, where $f'^i(\tau)$ is given in parametric form by

$$\begin{aligned} \tau &= \sigma + G_\alpha(\sigma, f^i(\sigma))\omega^\alpha, \\ f^i &= f^i(\sigma) + R^i{}_\alpha(\sigma, f^i(\sigma))\omega^\alpha, \end{aligned} \tag{7.110}$$

Denoting by $\tilde{\alpha}(\tau)$ the inverse function to $\tau + G_\alpha(\tau, f^i(\tau))\omega^\alpha$ we obtain

$$\begin{aligned} f'^i(\tau) &= f^i(\tilde{\alpha}(\tau)) + R^i{}_\alpha(\tilde{\alpha}(\tau), f(\tilde{\alpha}(\tau))\omega^\alpha \\ &= f^i(\tau) + \Big(R^i{}_\alpha(\tau, f(\tau)) - \dot{f}^i(\tau)G_\alpha(\tau, f(\tau))\Big)\,\omega^\alpha + O^2(\omega), \end{aligned} \tag{7.111}$$

or, in condensed notation, the transformed function in linear approximation is

$$z^i(\tau) \to z'^i(\tau) = z^i(\tau) + \tilde{\delta}z^i, \qquad \tilde{\delta}z^i \equiv \delta z^i - \dot{z}^i\delta\tau, \tag{7.112}$$

with $\delta\tau, \delta z^i$ specified in Eq. (7.109).

An infinitesimal transformation (7.109) is an *infinitesimal symmetry* of S_H, if

$$\int_{\tau_1+G_\alpha\omega^\alpha}^{\tau_2+G_\alpha\omega^\alpha} d\tau(f'_a\dot{f}'^a - H(f'^i,\tau))\Big|_{O(\omega)} = \int_{\tau_1}^{\tau_2} d\tau\left(f_a\dot{f}^a - H(f^i,\tau) + \frac{dN}{d\tau}\right). \tag{7.113}$$

Here f' is given by Eq. (7.111). Computing the left-hand side of (7.113) (the computation is similar to those on page 221), we conclude that the invariance condition is equivalent to the following Noether identity for S_H

$$\frac{d}{d\tau}(-p_a\delta q^a + H\delta\tau + N) = (\dot{q}^a - \{q^a, H\})\tilde{\delta}p_a - (\dot{p}_a - \{p_a, H\})\tilde{\delta}q^a. \tag{7.114}$$

This implies the following two properties of the Hamiltonian Noether charge

$$Q_\alpha(q,p,\tau) \equiv -p_a\delta q^a + H\delta\tau + N_\alpha. \tag{7.115}$$

First, it is preserved on solutions to equations of motion. Second, it obeys the equation (see Sect. 2.5)

$$\frac{\partial Q_\alpha}{\partial\tau} + \{Q_\alpha, H\} = 0, \quad \text{for any} \quad z(\tau). \tag{7.116}$$

In turn, this guarantees (see the previous section) that the following canonical transformation

$$\tilde{\tilde{\delta}}\tau = 0, \qquad \tilde{\tilde{\delta}}z^i = -\{z^i, Q_\alpha\}\omega_\alpha, \tag{7.117}$$

represents an infinitesimal symmetry of S_H. *The general (7.109) and canonical (7.117) symmetry transformations are equivalent since they produce the same conserved charge* Q_α.

It would be interesting to find an explicit form of $\tilde{\tilde{\delta}}z^i$ in terms of $\delta\tau, \delta z^i$. It can be deduced from detailed analysis of the Noether identity.[7] To see this, compute a derivative with respect to τ which appears on l.h.s. of Eq. (7.114). Then it reads

[7] One cannot compute the bracket $\{z^i, Q\}$ directly, since Q contains the unspecified function N.

$$
\begin{aligned}
&\dot{q}^a\left(-\delta p_a - p_b\frac{\partial\delta q^b}{\partial q^a} + H\frac{\partial\delta\tau}{\partial q^a} + \frac{\partial N}{\partial q^a}\right)\\
&\quad +\dot{p}_a\left(-p_b\frac{\partial\delta q^b}{\partial p_a} + H\frac{\partial\delta\tau}{\partial p_a} + \frac{\partial N}{\partial p_a}\right)\\
&\quad -p_b\frac{\partial\delta q^b}{\partial\tau} + \frac{\partial H\delta\tau}{\partial\tau} + \frac{\partial N}{\partial\tau} + \frac{\partial H}{\partial p_b}\delta p_b + \frac{\partial H}{\partial q^b}\delta q^b = 0. \qquad (7.118)
\end{aligned}
$$

Since this is true for any functions $q(\tau)$, $p(\tau)$, the terms in front of $\dot{q}$, $\dot{p}$ vanish separately. So Eq. (7.118) decomposes into three identities. Taking into account the expression (7.115) for the Noether charge, they can be written in the form

$$
\begin{aligned}
\delta q^a - \frac{\partial H}{\partial p_a}\delta\tau &= -\frac{\partial Q_\alpha}{\partial p_a}\omega^\alpha \equiv -\{q^a, Q_\alpha\}\omega^\alpha,\\
\delta p_a + \frac{\partial H}{\partial q^a}\delta\tau &= \frac{\partial Q_\alpha}{\partial q^a}\omega^\alpha \equiv -\{p_a, Q_\alpha\}\omega^\alpha, \qquad (7.119)\\
\frac{\partial Q_\alpha}{\partial\tau}\omega^\alpha &= -\frac{\partial H}{\partial p_a}\delta p_a - \frac{\partial H}{\partial q^a}\delta q^a. \qquad (7.120)
\end{aligned}
$$

Equation (7.119) gives the desired Poisson bracket $\{z^i, Q_\alpha\}$ in terms of $\delta\tau, \delta z^i$. Comparing (7.119) with (7.117) we find $\tilde{\tilde{\delta}}$ through δ

$$
\tilde{\tilde{\delta}}\tau = 0, \qquad \tilde{\tilde{\delta}}z^i = \delta z^i - \{z^i, H\}\delta\tau. \qquad (7.121)
$$

It is also instructive to compare transformations of a function induced by $\delta\tau$, δz^i and by $\tilde{\tilde{\delta}}z^i$. Comparing (7.112) with (7.121) we write

$$
\tilde{\delta}z^i = \tilde{\tilde{\delta}}z^i + \delta_0 z^i, \qquad (7.122)
$$

where $\delta_0 z^i = -(\dot{z}^i - \{z^i, H\})\delta\tau$. $\delta_0 z^i$ turns out to be a trivial symmetry of S_H. To check that it obeys the invariance condition, it is sufficient to confirm that the variation of S_H under the transformation can be presented as a total derivative. We obtain (see (2.127)) $\delta S_H = (\dot{q}^a - \{q^a, H\})\delta_0 p_a - (\dot{p}_a - \{p_a, H\})\delta_0 q^a + (p_a\delta q^a)^\cdot \equiv (p_a\delta_0 q^a)^\cdot$. In this case $N = p_a\delta_o q^a$, $\delta_0\tau = 0$, so the Noether charge (7.115) produced by the symmetry vanishes identically. Hence $\tilde{\delta}z$ and $\tilde{\tilde{\delta}}z$ are equivalent, leading to the same Noether identity, conserved charge and so on.

The remaining Eq. (7.120) does not contain any new information. Actually, its right-hand side reads

$$
\begin{aligned}
&-\frac{\partial H}{\partial p_a}\delta p_a - \frac{\partial H}{\partial q^a}\delta q^a \equiv -\frac{\partial H}{\partial p_a}\tilde{\tilde{\delta}}p_a - \frac{\partial H}{\partial q^a}\tilde{\tilde{\delta}}q^a\\
&\quad = -\frac{\partial H}{\partial p_a}\frac{\partial Q_\alpha}{\partial q^a}\omega^\alpha + \frac{\partial H}{\partial q^a}\frac{\partial Q_\alpha}{\partial p_a}\omega^\alpha = -\{Q_\alpha, H\}\omega^\alpha, \qquad (7.123)
\end{aligned}
$$

so (7.120) is simply the Eq. (7.116).

The results obtained can be summarized as follows. Let $\delta\tau, \delta z^i$ be an infinitesimal symmetry of S_H, $\tilde{\delta} z^i$ stands for the induced transformation of a function and $Q_\alpha(z^i, \tau)$ is the corresponding Noether charge, $\frac{dQ_\alpha}{d\tau} = 0$ on-shell. Being a conserved phase-space quantity, it also obeys the equation

$$\frac{\partial Q_\alpha}{\partial \tau} + \{Q_\alpha, H\} = 0, \quad \text{for any } z^i(\tau). \tag{7.124}$$

With the symmetry one associates the transformation

$$\tilde{\tilde{\delta}}\tau = 0, \qquad \tilde{\tilde{\delta}} z^i = \delta z^i - \{z^i, H\}\delta\tau, \tag{7.125}$$

which does not affect the time variable. Then

(a) $\tilde{\tilde{\delta}} z^i$ is a canonical transformation with the generator being the charge Q,

$$\tilde{\tilde{\delta}} z^i = \{z^i, Q_\alpha\}\omega_\alpha. \tag{7.126}$$

(b) Equation (7.124) implies that $\tilde{\tilde{\delta}} z^i$ is an infinitesimal symmetry of S_H. The corresponding Noether charges are Q_α.

(c) The transformations $\delta\tau$, δz^i and $\tilde{\tilde{\delta}} z^i$ are equivalent, as soon as they generate the same Noether charges. This can also be seen from comparison of the induced transformations of a function that coincide except for the trivial symmetry $\delta_0 z^i = -(\dot{z}^i - \{z^i, H\})\delta\tau$

$$\tilde{\delta} z^i = \tilde{\tilde{\delta}} z^i + \delta_0 z^i. \tag{7.127}$$

Finite symmetries of a general form and canonical transformations. We finish this section with two comments on the relationship between finite canonical and symmetry transformations of S_H.

1. Let $z'^i(z^k, \tau)$ stand for a univalent canonical transformation. Since time is not affected, the induced transformation of the function $z^i(\tau)$ has the same form. The formula for the corresponding transformation of S_H was obtained in Sect. 4.5

$$p_a\dot{q}^a - H(z) = p'_a\dot{q}'^a - \tilde{H}(z', \tau) + \frac{dF(z', \tau)}{d\tau}. \tag{7.128}$$

So, the canonical transformation is a symmetry of S_H if and only if it does not modify the Hamiltonian $\tilde{H}(z') = H(z)$.

2. We have seen above that an infinitesimal symmetry of S_H of a general form can be replaced by an infinitesimal symmetry that represents a canonical transformation. And what about a finite symmetry transformation? Let us take the

exponential of an infinitesimal canonical transformation (7.117) associated with the general symmetry (7.109)

$$z'^i = e^{-\omega^\alpha \{z^k, Q_\alpha\}\partial_k} z^i. \tag{7.129}$$

According to Sect. 7.8 it is a symmetry transformation of S_H. At the same time, according to Sect. 4.2, it represents a canonical transformation.

7.11.3 Hamiltonian Versus Lagrangian Global Symmetry

Lagrangian and Hamiltonian symmetries have been discussed in the previous sections in an independent manner. Remember that Lagrangian and Hamiltonian formulations are related by a change of variables that in practice reduces to the substitution $\dot{q}^a \to v^a(q, p, \tau)$; see Sect. 2.1.2. So each property of a Lagrangian formulation has its Hamiltonian counterpart.

The aim of this section is to show the relationship between the infinitesimal Lagrangian and Hamiltonian symmetries. Among other things, we find a remarkable expression for Lagrangian symmetry in terms of the Hamiltonian Noether charge, $\delta q^a = -\omega^\alpha \{q^a, Q_{H\alpha}\}|_{p \to \frac{\partial L}{\partial \dot{q}}}$.

We start with a couple of auxiliary formulas. Let A_H stand for the Hamiltonian counterpart of a Lagrangian quantity $A(q, \dot{q}, \tau)$, that is, $A_H \equiv A(q, \dot{q}, \tau)|_{\dot{q} \to v(q,p,\tau)}$. Then

$$\left.\frac{\partial A(q, \dot{q}, \tau)}{\partial \dot{q}^a}\right|_{\dot{q} \to v(q,p,\tau)} = M_{ab} \frac{\partial A_H}{\partial p_b}, \tag{7.130}$$

$$\left.\frac{\partial A(q, \dot{q}, \tau)}{\partial q^a}\right|_{\dot{q} \to v(q,p,\tau)} = \frac{\partial A_H}{\partial q^a} - \frac{\partial A_H}{\partial p_c} M_{cb} \frac{\partial v^b}{\partial q^a}, \tag{7.131}$$

$$\left.\frac{\partial A(q, \dot{q}, \tau)}{\partial q^a}\dot{q}^a\right|_{\dot{q} \to v(q,p,\tau)} = \{A_H, H\} - \frac{\partial A_H}{\partial p_c} M_{cb} \{v^b, H\}. \tag{7.132}$$

The last equality is the previous one multiplied by $v^a = \{q^a, H\}$, see (2.31).

Let

$$\delta\tau = G_\alpha(q, \dot{q}, \tau)\omega^\alpha, \qquad \delta q^a = R^a{}_\alpha(q, \dot{q}, \tau)\omega^\alpha, \tag{7.133}$$

be an infinitesimal Lagrangian symmetry, $Q_\alpha(q, \dot{q}, \tau)$ is the corresponding Noether charge, and $Q_{H\alpha}(q, p, \tau) \equiv Q_\alpha(q, v(q, p), \tau)$ stands for its Hamiltonian counterpart. Let us introduce the infinitesimal canonical transformation generated by Q_H

$$\delta_H z^i = -\{z^i, Q_{H\alpha}\}\omega^\alpha. \tag{7.134}$$

We show that it is a symmetry of S_H.

Computing the time derivative in the expression for the Lagrangian Noether identity

$$\frac{\delta S}{\delta q^a} D^a{}_\alpha = \frac{d}{d\tau}(-\frac{\partial L}{\partial \dot{q}^a} D^a{}_\alpha - L G_\alpha + N_\alpha), \tag{7.135}$$

this reads

$$\frac{\partial L}{\partial q^a} D^a{}_\alpha = -\frac{\partial L}{\partial \dot{q}^a}\left(\frac{\partial D^a{}_\alpha}{\partial q^b}\dot{q}^b + \frac{\partial D^a{}_\alpha}{\partial \dot{q}^b}\ddot{q}^b + \frac{\partial D^a{}_\alpha}{\partial \tau}\right) +$$
$$\frac{\partial(-L G_\alpha + N_\alpha)}{\partial q^b}\dot{q}^b + \frac{\partial(-L G_\alpha + N_\alpha)}{\partial \dot{q}^b}\ddot{q}^b + \frac{\partial(-L G_\alpha + N_\alpha)}{\partial \tau}. \tag{7.136}$$

Since this is true for arbitrary functions $q^a(\tau)$, the term in front of $\ddot{q}^a$ vanishes separately. So Eq. (7.136) implies two identities

$$-\frac{\partial L}{\partial \dot{q}^a}\frac{\partial D^a{}_\alpha}{\partial \dot{q}^b} + \frac{\partial(-L G_\alpha + N_\alpha)}{\partial \dot{q}^b} = 0, \tag{7.137}$$

$$\frac{\partial L}{\partial q^a} D^a{}_\alpha + \frac{\partial L}{\partial \dot{q}^a}\frac{\partial D^a{}_\alpha}{\partial q^b}\dot{q}^b + \frac{\partial(-L G_\alpha + N_\alpha)}{\partial q^b}\dot{q}^b =$$
$$-\frac{\partial L}{\partial \dot{q}^a}\frac{\partial D^a{}_\alpha}{\partial \tau} + \frac{\partial(-L G_\alpha + N_\alpha)}{\partial \tau}. \tag{7.138}$$

We cast Eq. (7.137) into the Hamiltonian form substituting $v(q, p, \tau)$ instead of $\dot{q}$. Using Eq. (7.130) and the fact that M is invertible we have

$$-p_a \frac{\partial D^a_{H\alpha}}{\partial p_b} + \frac{\partial(-L_H G_{H\alpha} + N_{H\alpha})}{\partial p_b} = 0, \tag{7.139}$$

or, equivalently,

$$\{q^a, Q_{H\alpha}\} + D^a_{H\alpha} = 0. \tag{7.140}$$

Remind that D determines the transformation of a function induced by the Lagrangian symmetry, see (7.73). So, comparing (7.140) with (7.134), we conclude that D is a Lagrangian counterpart of the canonical transformation generated by the Hamiltonian Noether charge, $\delta_H q^a = -\{z^i, Q_{H\alpha}\}\omega^\alpha = D^a_{H\alpha}\omega^\alpha = \delta q^a|_{\dot{q}\to v}$. Carrying out the inverse transformation, we obtain a formula giving an *expression for Lagrangian symmetry through the Noether charge*

$$\delta q^a = -\omega^\alpha \{q^a, Q_{H\alpha}\}\big|_{p\to\frac{\partial L}{\partial \dot{q}}}. \tag{7.141}$$

The Hamiltonian form of Eq. (7.138) is obtained using the identities (2.30) and (7.132). The result is

$$
\begin{aligned}
&\{p_a, H\}D^a{}_\alpha + p_a\{D^a{}_\alpha, H\} + \{LG_\alpha - N_\alpha, H\} + \\
&\left(-p_a\frac{\partial D^a{}_\alpha}{\partial p_b} + \frac{\partial(-LG_\alpha + N_\alpha)}{\partial p_b}\right) M_{bc}\{v^c, H\} \\
&\qquad = -p_a \left.\frac{\partial D^a{}_\alpha}{\partial \tau}\right|_v + \left.\frac{\partial(-LG_\alpha + N_\alpha)}{\partial \tau}\right|_v .
\end{aligned} \tag{7.142}
$$

The last bracket on l.h.s. of this expression vanishes due to the identity (7.139). Further, using the identity

$$
\left.\frac{\partial A(q, \dot{q}, \tau)}{\partial \tau}\right|_{\dot{q}\to v(q,p,\tau)} = \frac{\partial A}{\partial \tau} - \frac{\partial A}{\partial p_c} M_{cb}\frac{\partial v^c}{\partial \tau}, \tag{7.143}
$$

that follows from (7.130), we substitute v into r.h.s. of Eq. (7.142). Again the terms containing derivatives with respect to p vanish due to the identity (7.139). Then Eq. (7.142) acquires the form $\frac{\partial Q_\alpha}{\partial \tau} + \{Q_\alpha, H\} = 0$.

In the result, the Lagrangian Noether identity (7.67), being rewritten in phase-space variables, acquires the form

$$
\{q^a, Q_{H\alpha}\} = -D^a_{H\alpha} \equiv -(R^a_{H\alpha} - \{q^a, H\}G_{H\alpha}), \tag{7.144}
$$

$$
\frac{\partial Q_{H\alpha}}{\partial \tau} + \{Q_{H\alpha}, H\} = 0. \tag{7.145}
$$

As we know, Eq. (7.145) implies that an *infinitesimal canonical transformation (7.134) associated with the initial Lagrangian symmetry (7.133) is an infinitesimal symmetry of* S_H.

Chapter 8
Hamiltonian Formalism for Singular Theories

Modern particle and field theories often involve auxiliary variables which have no direct physical meaning. We have seen examples of this kind at the end of first chapter: Lagrangian multipliers for holonomic constraints, forceless Hertz mechanics, electrodynamics and the relativistic particle.

Their auxiliary character is supplied either by local symmetries presented in the Lagrangian action, or by the algebraic character of equations of motion for these variables. So, in Lagrangian formalism we deal with a singular theory. Equations of motion can have rather a complicated structure, including in general differential equations of the second and the first order as well as algebraic equations. Besides, identities among the equations can be present (see Eq. (1.267) for the case of electrodynamics). As a consequence, there is an ambiguity in solutions for *any* given initial conditions. So, the physical content of a theory with local symmetry is not a simple question.

We point out that the appearance of auxiliary variables is mostly due to our desire to incorporate the manifest Poincaré invariance (and locality) as the leading principle of formalism. As we have discussed in Sect. 1.12.5, the auxiliary variables allow us to "decompose" a nonlinear global symmetry carried out on physical dynamical variables into a linear global symmetry plus a local symmetry.

Hamiltonian formalism is well adapted for the investigation of a singular theory. The phase-space description allows us to automatically separate the dynamical part of equations of motion from the algebraic part as well as to analyze the ambiguity present in solutions of equations of motion. Besides, the transition to Hamiltonian formulation is a necessary step in the process of canonical quantization of classical theory. Systematic analysis of an arbitrary singular theory was started in the pioneer works of Dirac [8] and Bergmann [38], and is posed at present on a solid mathematical ground [10, 11].

Notation In this chapter it will be convenient to change the notation as follows. Generalized coordinates of the configuration space are denoted by q^A, where A ranges from 1 to $[A]$, that is $[A]$ stands for the number of variables q^A. For the phase-space variables we write $z \equiv (q^A, p_B)$.

A. Deriglazov, *Classical Mechanics*, DOI 10.1007/978-3-642-14037-2_8,

8.1 Hamiltonization of a Singular Theory: The Recipe

Here we discuss the working recipe for Hamiltonization of a singular theory. We first illustrate the recipe on two simple examples that reveal the improvements that are necessary when we apply a Hamiltonization procedure of Sect. 2.1.3 to a singular theory.

8.1.1 Two Basic Examples

Example 1. Consider the Lagrangian action

$$S = \frac{1}{2}\int d\tau \left[[(xy)^{\cdot}]^2 + x^2 + y^2 \right], \tag{8.1}$$

written for the configuration variables $x(\tau)$, $y(\tau)$. Its Hessian is degenerate

$$\det M = \det \frac{\partial^2 L}{\partial \dot{q}^A \partial \dot{q}^B} = \det \begin{pmatrix} y^2 & xy \\ xy & x^2 \end{pmatrix} = 0, \tag{8.2}$$

so we deal with a singular theory. Lagrangian equations can be presented in the form

$$(x^2)^{\cdot\cdot} = 1, \qquad x^2 - y^2 = 0, \tag{8.3}$$

and can be integrated out

$$x = \pm\sqrt{\frac{1}{2}\tau^2 + c\tau + d}, \qquad y = \pm x. \tag{8.4}$$

To find the Hamiltonian formulation, the first step is to solve the defining equations for momenta

$$p = \frac{\partial L}{\partial \dot{x}} = y^2\dot{x} + xy\dot{y}, \qquad \pi = \frac{\partial L}{\partial \dot{y}} = x^2\dot{y} + xy\dot{x}. \tag{8.5}$$

The first equation can be resolved with respect to $\dot{x}$

$$\dot{x} = \frac{p}{y^2} - \frac{x\dot{y}}{y}. \tag{8.6}$$

Its substitution into the second equation gives

$$G \equiv xp - y\pi = 0. \tag{8.7}$$

This is the first characteristic property of a singular theory: only some of the equations determining momenta can be resolved with respect to velocities. The remaining equations do not contain velocities at all, representing algebraic equations relating coordinates and momenta. They are called *primary constraints of the Dirac procedure*.

As the second step, let us try to construct a Hamiltonian according to the standard prescription: we write $H_0 = p\dot{x} + \pi\dot{y} - L$ and use Eqs. (8.6) and (8.7) to exclude all the velocities from this expression. This gives the expression $H_0 = \frac{p^2}{2y^2} - \frac{1}{2}x^2 - \frac{1}{2}y^2$. Using the recipe $\dot{z} = \{z, H_0\}$, we obtain the equations

$$\dot{x} = \frac{p}{y^2}, \quad \dot{p} = x, \quad \dot{y} = 0, \quad \dot{\pi} = \frac{p^2}{y^3} + y. \tag{8.8}$$

Multiply the first equation by x, compute a derivative of the resulting expression and use the other equations of the system, obtaining the consequence $(x^2)^{\cdot\cdot} = 2 + 2\dot{x}^2$. This is different from Eq. (8.3).

So, the recipe of Sect. 2.1.3 does not lead to the right Hamiltonian and must be modified. According the Dirac procedure, the modification is rather nontrivial and is as follows. We introduce an extended phase space with the coordinates x, y, p, π, v, where $v(\tau)$ is an additional variable. It is sometimes called the *Lagrangian multiplier*. On the reason that will become clear in the next section, we will call it a *velocity*. The right Hamiltonian is given by

$$H = \frac{p^2}{2y^2} - \frac{1}{2}x^2 - \frac{1}{2}y^2 + v(xp - y\pi), \tag{8.9}$$

that is, we add to H_0 the primary constraint multiplied by v. H is called the *complete Hamiltonian*.

This is the second characteristic property of a singular theory: the rule for constructing the Hamiltonian of the theory must be modified according to Eq. (8.9).

Of course, now we need to do a little work to show that this H leads to the right dynamics. The Poisson bracket on the extended space is defined in the standard way, $\{z^i, z^j\} = \omega^{ij}$, that is it does not involve v: $\{v, z^i\} = 0$. Then (8.9) implies the Hamiltonian equations

$$\dot{x} = \frac{p}{y^2} + vx, \quad \dot{p} = x - vp, \quad \dot{y} = -vy, \quad \dot{\pi} = \frac{p^2}{y^3} + y + v\pi. \tag{8.10}$$

Together with the primary constraint (8.7), there are five equations for five variables.

We can try to use the algebraic Eq. (8.7) to obtain more *algebraic* consequences of the system (8.10), (8.7), if any. This works as follows. Compute the derivative of G, $\dot{G} = \dot{x}p + x\dot{p} - \dot{y}\pi - y\dot{\pi} = 0$. Use Eqs. (8.10) to exclude all the derivatives appearing in this expression. This gives

$$T \equiv x^2 - y^2 = 0. \tag{8.11}$$

So, the Eqs. (8.10) and (8.7) imply one more algebraic equation; occasionally it is the second equation from (8.3). This is called the *secondary constraint of the Dirac procedure* or the *second-stage constraint*.

In turn, the derivative of T implies

$$\dot{T} = 2\left(\frac{xp}{y^2} + vx^2 + vy^2\right) = 0, \quad \text{then} \quad v = -\frac{p}{2x^3}. \tag{8.12}$$

Instead of a new constraint, we have obtained an algebraic expression for v in terms of the phase-space variables. The derivative of this equation can not lead to new algebraic equations since the equations at our disposal do not allow us to exclude $\dot{v}$.

Exercise
Show that the derivative of a constraint can be computed according to the formula

$$\dot{G} = \{G, H\}, \tag{8.13}$$

which is usually used in practice. Reproduce Eqs. (8.11) and (8.12) using this formula.

Not all equations of the complete system (8.10), (8.7), (8.11) and (8.12) are independent. Keeping the independent equations only, they read

$$\dot{x} = \frac{p}{2x^2}, \qquad \dot{p} = x + \frac{p^2}{2x^3}; \tag{8.14}$$

$$y^2 = x^2, \qquad \pi = \pm p. \tag{8.15}$$

Note that the auxiliary variable v has disappeared from the final equations. Equations for the pair x, p do not involve y, π, and form a normal system. Equations for y, π are algebraical.

Now we are ready to confirm that the Hamiltonian equations are equivalent to the Lagrangian ones. Multiply the first equation from (8.14) by x, and compute the derivative of the resulting expression. Using the Eqs. (8.14), it reads $(x^2)^{\cdot\cdot} = 1$. Together with the first equation from (8.15), this reproduces the Lagrangian dynamics (8.3).

Example 2. Consider the manifestly Poincaré-invariant action

$$S = \int d\tau \left[\frac{1}{2e}\dot{x}_\mu \dot{x}^\mu + \frac{1}{2}em^2 \right], \tag{8.16}$$

written for the configuration-space variables $e(\tau)$, $x^\mu(\tau)$, $\mu = 0, 1, 2, 3$. Besides the global Poincaré transformations, it is also invariant under the local transformations of reparametrization

$$\tau = \tau(\tau'), \qquad x^\mu(\tau) = x'^\mu(\tau'), \qquad e(\tau) = \frac{\partial \tau'}{\partial \tau} e'(\tau). \tag{8.17}$$

The Lagrangian equations read

$$\frac{d}{d\tau}\left(\frac{\dot{x}^\mu}{e}\right) = 0, \qquad \dot{x}^2 - e^2 m^2 = 0. \tag{8.18}$$

The second equation implies $e = \pm\frac{1}{m}\sqrt{\dot{x}^2}$; therefore the first equation reads $\frac{d}{d\tau}\left(\frac{\dot{x}^\mu}{\sqrt{\dot{x}^2}}\right) = 0$. There are just the equations of motion for a free relativistic particle (1.274). So, we have found one more Lagrangian action for the description of a relativistic particle. The Eqs. (8.18) can be immediately integrated out (it is one of the advantages of the action (8.16) as compared with the square-root action (1.273))

$$x^\mu = x_0^\mu + p^\mu \int e d\tau, \tag{8.19}$$

where $p^\mu =$ const such that $p^2 = m^2$, and $e(\tau)$ remains an arbitrary function. Since the action (8.16) does not involve $\dot{e}$, we have a singular theory. Let us construct its Hamiltonian formulation. The equations determining momenta are

$$p_\mu = \frac{1}{e}\eta_{\mu\nu}\dot{x}^\nu, \quad \text{then} \quad \dot{x}^\mu = e\eta^{\mu\nu}p_\nu; \tag{8.20}$$

$$p_e = 0, \tag{8.21}$$

where $p_e = \frac{\partial L}{\partial \dot{e}}$ is the conjugate momentum for e. Equation (8.21) represents the primary constraint. Then the total Hamiltonian reads

$$H = \frac{e}{2}(p^2 - m^2) + v p_e. \tag{8.22}$$

Using Eq. (8.13) we find the secondary constraint

$$\dot{p}_e = \frac{1}{2}(p^2 - m^2) = 0, \quad \text{or} \quad T \equiv p^2 - m^2 = 0. \tag{8.23}$$

Its derivative vanishes, so there is no new constraint nor an equation for determining the Lagrangian multiplier v. In the result, the dynamics is governed by the Hamiltonian equations following from (8.22)

$$\dot{x}^\mu = ep^\mu, \qquad \dot{p}_\mu = 0, \quad \dot{e} = v, \qquad \dot{p}_e = 0. \tag{8.24}$$

These are accompanied by the constraints (8.21) and (8.23).

Note that these equations do not determine the variable $v(\tau)$. Indeed, for an arbitrary $v(\tau)$ the functions

$$\begin{aligned} e &= \int d\tau v, & p_e &= 0, \\ x^\mu &= x_0^\mu + \int e d\tau, \quad & p_\mu &= \text{const} \quad \text{such that} \quad p^2 = 0, \end{aligned} \tag{8.25}$$

resolve the Eqs. (8.24), (8.21) and (8.23).

This is one more characteristic property of a singular theory: the Lagrangian multipliers that have not been determined in the course of the Dirac procedure are not determined by the complete system of equations of motion either. They enter into solutions as arbitrary functions. The advantage of the Hamiltonian formalism is that the arbitrariness is detected and described automatically in the course of the procedure.

As has been noted above, we are dealing with the relativistic particle. The arbitrariness presented in solutions, as well as the procedure for its removal, were discussed in the Lagrangian framework in Sect. 1.12.4. This works in Hamiltonian formulation as well. Suppose that $x^\mu = x^\mu(\tau)$, $p^\mu = p^\mu(\tau)$ represent parametric equations for the observable quantities $x^a(t)$, $p_a(t)$. As a consequence of (8.24), they obey the equations of motion

$$\frac{dx^a}{dt} = -\frac{cp_a}{\sqrt{m^2 + (p^b)^2}}, \qquad \frac{dp_a}{dt} = 0, \tag{8.26}$$

which reproduce our previous result (1.276). Note that neither e nor v enter into these equations.

Note that the two examples of a singular theory have essentially different structures of equations of motion. When all the velocities have been determined in the course of the Dirac procedure, the singular theory is called *non-degenerate* [10]. Otherwise, it is called a *singular degenerate theory*. It will be seen below that the difference is encoded in the algebraic properties of constraints with respect to the Poisson bracket.

8.1.2 Dirac Procedure

Now we are ready to discuss the Dirac recipe for an arbitrary singular theory

$$S = \int d\tau L(q^A, \dot{q}^A), \qquad \det \frac{\partial^2 L}{\partial \dot{q}^A \partial \dot{q}^B} = 0. \tag{8.27}$$

We impose the following rank condition on the Hessian matrix M:

$$\text{rank}\, M_{AB} \equiv \text{rank} \frac{\partial^2 L}{\partial \dot{q}^A \partial \dot{q}^B} = [i] < [A]. \tag{8.28}$$

The first stage of the Dirac procedure consists of Hamiltonization of the theory.

(1) Introduce conjugate momenta for the variables q^A according to the equations

$$p_A = \frac{\partial L(q, \dot{q})}{\partial \dot{q}^A}. \tag{8.29}$$

They are considered as algebraic equations for determining the velocities $\dot{q}^A$. According to the well-known theorem about implicit function, Eq. (8.28) guarantees that some $[i]$ velocities among $\dot{q}^A$, say $\dot{q}^i$, $i = 1, 2, \ldots, [i]$, can be found from these equations. Let us denote the solution as[1]

$$\dot{q}^i = v^i(q^A, p_j, \dot{q}^\alpha). \tag{8.30}$$

These expressions can be substituted into the remaining $[\alpha]$ equations for the momenta (8.29). By construction, the resulting expressions do not depend[2] on $\dot{q}^\alpha$ and represent the primary constraints $\Phi_\alpha(q, p)$ of the theory. They read

$$\Phi_\alpha \equiv p_\alpha - f_\alpha(q^A, p_j) = 0, \qquad \alpha = 1, 2, \ldots, [\alpha] = [A] - [i], \tag{8.31}$$

where

$$f_\alpha(q^A, p_j) \equiv \left. \frac{\partial L}{\partial \dot{q}^\alpha} \right|_{\dot{q}^i = v^i(q^A, p_j, \dot{q}^\alpha)}. \tag{8.32}$$

The Eqs. (8.30) and (8.31) are thus equivalent to the system (8.29).

(2) Introduce an *extended phase space* parameterized by the coordinates q^A, p_A, v^α, and define a *complete Hamiltonian* H according to the rule

$$\begin{aligned} H(q^A, p_A, v^\alpha) &= p_A \dot{q}^A - L(q^A, \dot{q}^A) + v^\alpha \Phi_\alpha(q^A, p_B) \\ &\equiv H_0(q^A, p_j) + v^\alpha \Phi_\alpha(q^A, p_B), \end{aligned} \tag{8.33}$$

---

[1] Latin indices from the middle of the alphabet, i, j, k, are reserved for those coordinates whose velocities can be found from (8.29). Greek indices from the beginning of the alphabet, α, β, γ, are used to denote the remaining coordinates.

[2] Indeed, if they depended on one of $\dot{q}^\alpha$, it would be possible to find it in terms of q, p, contradicting the rank condition (8.28).

where we use Eqs. (8.30) and (8.31) to exclude all the velocities $\dot{q}^A$, that is

$$H_0(q^A, p_j) = \left[p_A\dot{q}^A - L(q^A, \dot{q}^A)\right]\Big|_{(8.30),\,(8.31)}. \tag{8.34}$$

H_0 is called the *Hamiltonian* of the theory.

Exercise
Show that H_0 does not depend on v^α: $\frac{\partial H_0}{\partial v^\alpha} = 0$, or on p_α: $\frac{\partial H_0}{\partial p_\alpha} = 0$.

(3) Write the equations of motion

$$\dot{q}^A = \{q^A, H\}, \qquad \dot{p}_A = \{p_A, H\}, \tag{8.35}$$

$$\Phi_\alpha(q^A, p_B) = 0, \tag{8.36}$$

where $\{q^A, p_B\} = \delta^A{}_B$ stands for the Poisson bracket. For later use, we write the Hamiltonian equations in detail

$$\dot{q}^i = \frac{\partial H_0}{\partial p_i} - \frac{\partial f_\beta}{\partial p_i}v^\beta,$$

$$\dot{p}_i = -\frac{\partial H_0}{\partial q^i} + \frac{\partial f_\beta}{\partial q^i}v^\beta; \tag{8.37}$$

$$\dot{q}^\alpha = v^\alpha, \tag{8.38}$$

$$\dot{p}_\alpha = -\frac{\partial H_0}{\partial q^\alpha} + \frac{\partial f_\beta}{\partial q^\alpha}v^\beta. \tag{8.39}$$

The system is equivalent to the Lagrangian equations for the action (8.27). This equivalence will be proved in the next section.

The second, third, …stages of the Dirac procedure consist in revealing all the algebraic consequences of the system (8.35) and (8.36).

According to Eqs. (8.36), all the solutions are confined to lying on a surface of the extended phase space defined by the algebraic equations $\Phi_\alpha = 0$. It may happen that the system (8.35) and (8.36) contains in reality more then $[\alpha]$ algebraic equations. Indeed, computing the derivative of the primary constraints with respect to time and using (8.35), we obtain the algebraic consequences of the system

$$\{\Phi_\alpha, H\} \equiv \{\Phi_\alpha, \Phi_\beta\}v^\beta + \{\Phi_\alpha, H_0\} = 0. \tag{8.40}$$

Henceforth they are called *second-stage equations* of the Dirac procedure. They can be added to Eqs. (8.35) and (8.36), which gives an equivalent system.

After that the dynamical equations $\dot{p}_\alpha = \{p_\alpha, H\}$ turn into consequences of other equations of the system, as the following computation shows:

$$\dot{\Phi}_\alpha = \frac{\partial \Phi_\alpha}{\partial z^A}\left(\dot{z}^A - \{z^A, H\}\right) + \frac{\partial \Phi_\alpha}{\partial z^A}\{z^A, H\}$$
$$= \delta^{\alpha\beta}\left(\dot{p}_\beta - \{p_\beta, H\}\right) + \{\Phi_\alpha, H\} = \dot{p}_\alpha - \{p_\alpha, H\}. \tag{8.41}$$

So they can be omitted from consideration. In many cases we will keep them, with the aim to preserve the symmetric form of the system (8.35).

Let us analyze the structure of the second-stage system. It is considered as a system of linear equations for determining the velocities v^α. According to known theorems of linear algebra, if

$$\text{rank}\{\Phi_\alpha, \Phi_\beta\} = [\alpha'] \leq [\alpha], \tag{8.42}$$

then $[\alpha']$ equations can be used to represent some $v^{\alpha'}$ through other variables. The velocities $v^{\alpha'}$ thus determined can be substituted into the remaining $[\alpha''] \equiv [\alpha] - [\alpha']$ equations; the resulting expressions do not contain v^α at all. After doing this, the second-stage system acquires the form

$$v^{\alpha'} = v^{\alpha'}(q^A, p_j, v^{\alpha''}), \qquad \Phi_{\alpha''}(q^A, p_j) = 0. \tag{8.43}$$

Functionally-independent equations among $\Phi_{\alpha''} = 0$, if any, represent the *second-stage Dirac constraints*. Thus all the solutions of the system (8.35) and (8.36) are confined to lying on the surface defined by $\Phi_\alpha = 0$ and by the Eqs. (8.43).

The velocities that have been determined can be substituted into the expression for the complete Hamiltonian.[3]

The procedure described above can be now repeated for the second-stage constraints, which can produce non-trivial *third-stage* algebraic equations,

$$\{\Phi_{\alpha''}, H\} = 0. \tag{8.44}$$

This may determine some of the velocities and may produce new constraints. As above, adding them to the system (8.35), (8.36) and (8.40), some of the dynamical equations can be omitted from consideration. If the system (8.44) implies new constraints, we start the fourth stage of the Dirac procedure, and so on.

Since the number of functionally-independent constraints can not be more than dimension $2[A]$ of the phase space, the procedure necessarily stops at some stage, say N.

[3] Note that we can substitute the velocities $v^{\underline{\alpha}}(z, v^{\bar{\beta}})$ into the complete Hamiltonian before computing the Poisson brackets, which does not alter the resulting equations of motion. This follows from the fact that the velocities enter into H multiplied by the primary constraints. So, on the constraint surface, $\{z, v^{\alpha'}(z)\Phi_\alpha\} = \{z, \Phi_\alpha\}v^{\alpha'}(z) + \{z, v^{\alpha'}(z)\}\Phi_\alpha = \{z, \Phi_\alpha\}v^{\alpha'}$.

The complete set of higher-stage constraints is denoted by $\Phi_a(q^A, p_j) = 0$ (Latin indices from the beginning of the alphabet, a, b, c, are used to denote the higher-stage constraints). Then the complete constraint system is

$$\Phi_I \equiv (\Phi_\alpha, \Phi_a) = 0, \qquad a = 1, 2, \ldots [a]. \tag{8.45}$$

Then all the higher-stage algebraic equations are given by the system

$$\{\Phi_I, H\} = 0. \tag{8.46}$$

All the solutions to Eq. (8.35) and (8.36) are confined to lying on the surface defined by the equations $\Phi_\alpha = 0$ and (8.46); see Fig. 8.1 on page 246. By construction, after substitution of the velocities v^α determined in the course of the Dirac procedure, the Eq. (8.46) hold on the complete constraint surface $\Phi_I = 0$.

In short, after completing the Dirac procedure, the theory can be described by Hamiltonian equations (8.35) which are accompanied by the constraints (8.45). Besides, some of the velocities (or all of them) have been determined in the process. Equivalently, the complete system of equations is given by formulas (8.35), (8.36) and (8.46). We also repeat (see page 245) that $[I]$ dynamical equations are consequences of other equations of the complete system (where $[I]$ is the number of all constraints).

Exercise
Construct the Hamiltonian formulation for the action (2.162) and find the Eqs. (2.165).

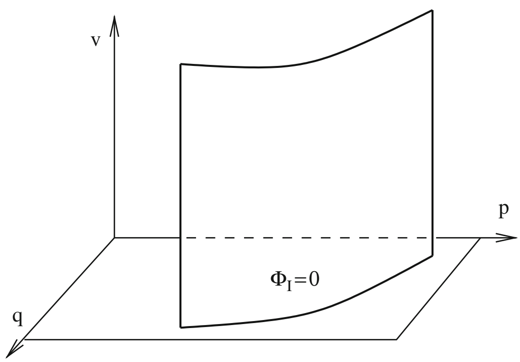

Fig. 8.1 All the trajectories of a singular theory lie on the surface $\Phi_\alpha = 0$, $\{\Phi_I, H\} = 0$. The surface belongs to the cylindrical surface $\Phi_I = 0$ of the extended phase space

8.2 Justification of the Hamiltonization Recipe

Remember that the Lagrangian equations for the action (8.27)

$$\frac{d}{d\tau}\left(\frac{\partial L(q,\dot{q})}{\partial \dot{q}^A}\right) - \frac{\partial L(q,\dot{q})}{\partial q^A} = 0. \tag{8.47}$$

can be identically rewritten in the form

$$M_{AB}\ddot{q}^B = K_B, \tag{8.48}$$

where $M_{AB}(q,\dot{q})$ stands for the Hessian matrix, and $K_B(q,\dot{q})$ has been specified in Eq. (1.115).

To construct the corresponding Hamiltonian formulation, we basically follow the same ideology as in Sect. 2.1.2. First we rewrite the Lagrangian equations in a first-order form, and then we look for the change of variables that supplies the Hamiltonian form of these equations.

8.2.1 Configuration-Velocity Space

We introduce $2[A]$-dimensional *configuration-velocity space* parameterized by *independent* coordinates q^A, v^B (sometimes the coordinates v^B are called *(generalized) velocities*). Let us define the evolution of these variables according to the equations $M_{AB}\ddot{q}^B = K_B$, $v^A = \dot{q}^A$. As before, time dependence of the coordinates $q^A(\tau)$ is determined by the Lagrangian equations (8.48), while $v^A(\tau)$ accompanies $\dot{q}^A(\tau)$: $v^A(\tau)$ is determined from the known $q^A(\tau)$ taking its derivative. Evidently, this system is equivalent to (8.48). Substitution of the second equation into the first one gives the desired first-order system

$$\dot{q}^A = v^A, \qquad \bar{M}_{AB}\dot{v}^B = \bar{K}_A, \tag{8.49}$$

where $\bar{M}$, $\bar{K}$ are obtained from M, K by the replacement $\dot{q} \to v$, for example

$$\bar{M}_{AB} \equiv M_{AB}(q,\dot{q})|_{\dot{q}^A \to v^A} = \frac{\partial L(q,v)}{\partial v^A \partial v^B}. \tag{8.50}$$

For a quantity defined on the configuration space we indicate its arguments manifestly: $A(q,\dot{q})$. Where arguments of a quantity are not indicated, we adopt the following conventions. Symbols with a bar are used to denote functions on the configuration-velocity space: $\bar{A} \equiv A(q,v)$, while symbols without a bar denote functions on the space q^A, p_j, v^α (see below):

$$A = A(q^A, p_j, v^\alpha) \equiv A(q^A, v^B)|_{v^j \to v^j(q^A, p_j, v^\alpha)}. \tag{8.51}$$

To analyze the system (8.49), we suppose that the rank minor of the matrix M is placed in the upper left corner.[4] Then, according to the rank condition (8.28), our variables can be decomposed into two groups, $q^A = (q^i, q^\alpha)$, $i = 1, 2, \ldots, [i]$, $\alpha = 1, 2, \ldots, [\alpha] = [A] - [i]$. The Hessian matrix reads

$$\bar{M}_{AB} = \begin{pmatrix} \bar{M}_{ij} & \bar{M}_{i\beta} \\ \bar{M}_{\alpha j} & \bar{M}_{\alpha\beta} \end{pmatrix}, \quad \det \bar{M}_{ij} \neq 0. \tag{8.52}$$

Thus, the Latin indices from the middle of the alphabet, i, j, k, correspond to the coordinates related with the invertible block of matrix M. Greek indices from the beginning of the alphabet, α, β, γ, are used to denote the remaining variables.

The inverse matrix for $\bar{M}_{ij}$ is denoted $\tilde{\bar{M}}^{ij}$, we have $\bar{M}_{ij}\tilde{\bar{M}}^{jk} = \delta_i{}^k$. The equations of motion (8.49) read

$$\bar{M}_{ij}\dot{v}^j + \bar{M}_{i\beta}v^\beta = K_i, \tag{8.53}$$

$$\bar{M}_{\alpha j}\dot{v}^j + \bar{M}_{\alpha\beta}v^\beta = K_\alpha. \tag{8.54}$$

According to the rank condition, the Eqs. (8.53) can be resolved with respect to $\dot{v}^i$, $\dot{v}^i = \tilde{\bar{M}}^{ij}(K_j - \bar{M}_{j\alpha}\dot{v}^\alpha)$, and then substituted into the Eqs. (8.54) with the result being $[\bar{M}_{\alpha\beta} - \bar{M}_{\alpha i}\tilde{\bar{M}}^{ij}\bar{M}_{j\beta}]\dot{v}^\beta = \bar{K}_\alpha - \bar{M}_{\alpha i}\tilde{\bar{M}}^{ij}\bar{K}_j$. It must be $\operatorname{rank}[\bar{M}_{\alpha\beta} - \bar{M}_{\alpha i}\tilde{\bar{M}}^{ij}\bar{M}_{j\beta}] = 0$ (otherwise, we would be able to resolve the equations in relation to some of $\dot{v}^\alpha$, contradicting (8.28)). Hence the matrix M obeys the identity

$$\bar{M}_{\alpha\beta} - \bar{M}_{\alpha i}\tilde{\bar{M}}^{ij}\bar{M}_{j\beta} = 0. \tag{8.55}$$

Exercise
Show that the quantities

$$\bar{C}_\alpha{}^A \equiv \left(-\bar{M}_{\alpha j}\tilde{\bar{M}}^{ji},\ \delta_\alpha{}^\beta\right), \tag{8.56}$$

form a basis on the space of null vectors of the matrix $\bar{M}_{AB}$

$$\bar{C}_\alpha{}^A\bar{M}_{AB} = 0. \tag{8.57}$$

In the result, the equations of motion (8.49) are presented in the equivalent form

$$\dot{q}^A = v^A, \tag{8.58}$$

[4] If not, the initial variables q^A can be re-numbered to achieve this.

$$\dot{v}^i = \tilde{\bar{M}}^{ij}(\bar{K}_j - \bar{M}_{j\alpha}\dot{v}^\alpha), \tag{8.59}$$

$$\bar{K}_\alpha - \bar{M}_{\alpha i}\tilde{\bar{M}}^{ij}\bar{K}_j = 0. \tag{8.60}$$

The Eqs. (8.60) are algebraic ones. As will be seen below, they coincide with second-stage equations of the Dirac procedure, $\{\Phi_\alpha, H\} = 0$. For the first example of the previous section, this is $x^2 - y^2 = 0$. For the second example it is $v^2 - e^2m^2 = 0$.

8.2.2 Hamiltonization

The next step is to show whether the Eqs. (8.58) and (8.59) form a Hamiltonian system in properly chosen coordinates of the configuration-velocity space. Our aim now will be to demonstrate the following

Assertion Given the singular theory (8.27) and (8.28), consider the following change of variables

$$\begin{pmatrix} q^A \\ v^\alpha \\ v^i \end{pmatrix} \leftrightarrow \begin{pmatrix} q^A \\ v^\alpha \\ p_i \end{pmatrix}, \quad \text{where} \quad p_i = \frac{\partial L(q^A, v^i, v^\alpha)}{\partial v^i}. \tag{8.61}$$

According to the rank condition (8.52), it is invertible. Let us denote the inverse transformation as[5]

$$v^i = v^i(q^A, p_j, v^\alpha). \tag{8.62}$$

We also introduce the reduced Poisson bracket, that is the Poisson bracket computed with respect to (q^i, p_j)-variables only. For the functions $A(q^A, p_i, v^\alpha)$, $B(q^A, p_i, v^\alpha)$ this reads

$$\{A, B\}_R = \frac{\partial A}{\partial q^i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q^i}. \tag{8.63}$$

Then, in the variables q^A, p_i, v^α, the equations of motion (8.58), (8.59) and (8.60) acquire the form

$$\begin{aligned} \dot{q}^i &= \{q^i, H_R\}_R = \frac{\partial H_0}{\partial p_i} - \frac{\partial f_\beta}{\partial p_i}v^\beta, \\ \dot{p}_i &= -\{p_i, H_R\}_R = -\frac{\partial H_0}{\partial q^i} + \frac{\partial f_\beta}{\partial q^i}v^\beta \end{aligned} \tag{8.64}$$

[5] The velocities v^α which "survive" after the variable change are just the Lagrangian multipliers of the Dirac recipe, see below.

$$\dot{q}^\alpha = v^\alpha, \tag{8.65}$$

$$\triangle_{\alpha\beta}(q^A, p_j)v^\beta + H_\alpha(q^A, p_j) = 0. \tag{8.66}$$

The basic quantity of the Hamiltonian formulation turns out to be

$$H_R \equiv H_0(q^A, p_j) - f_\alpha(q^A, p_j)v^\alpha, \tag{8.67}$$

where

$$H_0(q^A, p_j) = \left(\frac{\partial \bar{L}}{\partial v^A}v^A - \bar{L}\right)\bigg|_{v^i} \equiv \left(p_i v^i - \bar{L}(q, v) + v^\alpha \frac{\partial \bar{L}}{\partial v^\alpha}\right)\bigg|_{v^i}, \tag{8.68}$$

is just the Hamiltonian of the Dirac recipe. It is also denoted

$$f_\alpha(q^A, p_j) \equiv \frac{\partial \bar{L}}{\partial v^\alpha}\bigg|_{v^i}, \tag{8.69}$$

$$\begin{aligned} \triangle_{\alpha\beta} &= \{f_\alpha, f_\beta\}_R - \left(\frac{\partial f_\alpha}{\partial q^\beta} - \frac{\partial f_\beta}{\partial q^\alpha}\right), \\ H_\alpha &= -\{f_\alpha, H_0\}_R - \frac{\partial H_0}{\partial q^\alpha}. \end{aligned} \tag{8.70}$$

According to the Assertion, equations for the (q^i, p_j)-sector are presented in the Hamiltonian form (8.64). As compared with the Dirac system (8.36), (8.37), (8.38) and (8.39), there is no conjugated momenta for the variables q^α. Hence, we have not yet arrived at the Dirac recipe.

Proof of the Assertion To prove the assertion, we will need to know properties of the transition function $v^i(q^A, p_j, v^\alpha)$, which is given in an implicit form by Eq. (8.61), as well as a structure of the Lagrangian as a function of q^A, p_j, v^α, $L(q^A, p_j, v^\alpha) \equiv \bar{L}(q^A, v^i, v^\alpha)|_{v^i}$.

Derivatives of the identity $p_i \equiv \frac{\partial \bar{L}(q^A, v^i, v^\alpha)}{\partial v^i}|_{v^i(q^A, p_j, v^\alpha)}$ give the relationships

$$\frac{\partial v^i}{\partial p_j} = \tilde{M}^{ij}, \qquad \frac{\partial v^i}{\partial v^\alpha} = -\tilde{M}^{ij}M_{j\alpha}, \qquad \frac{\partial v^i}{\partial q^A} = -\tilde{M}^{ij}\frac{\partial^2 \bar{L}}{\partial v^j \partial q^A}\bigg|_{v^i}. \tag{8.71}$$

Then the identity (8.55) acquires the form $\frac{\partial}{\partial v^\alpha}\left(\frac{\partial \bar{L}}{\partial v^\beta}|_{v^i}\right) = 0$, so the quantity

$$f_\alpha(q^A, p_j) \equiv \frac{\partial \bar{L}}{\partial v^\alpha}\bigg|_{v^i}, \tag{8.72}$$

does not depend on v^α. Its derivative reads

$$\frac{\partial f_\alpha(q^A, p_j)}{\partial p_i} = \tilde{M}^{ij} M_{j\alpha}, \tag{8.73}$$

so the quantity $\tilde{M}^{ij} M_{j\alpha}$ does not depend on v^α either. Remember that a quantity without a bar denotes a function defined on the space q^A, p_j, v^α; see our notation (8.51).

Using the known formula for the derivative of a composed function, Eq. (8.72) can be identically rewritten in terms of L

$$f_\alpha(q^A, p_j) = -\frac{\partial}{\partial v^\alpha}\Big(p_i v^i - L\Big), \tag{8.74}$$

which implies that the quantity

$$H_R(q^A, p_j, v^\alpha) \equiv p_i v^i - L, \tag{8.75}$$

is at most linear on v^α. Integrating out the Eq. (8.74) we obtain

$$H_R(q^A, p_j, v^\alpha) = H_0(q^A, p_j) - v^\alpha f_\alpha(q^A, p_j), \tag{8.76}$$

where H_0 stands for an integration constant. It can be found in terms of the initial Lagrangian by comparison of Eqs. (8.75) and (8.76)

$$H_0(q^A, p_j) = \left(\frac{\partial \bar{L}}{\partial v^A} v^A - \bar{L}\right)\bigg|_{v^i}. \tag{8.77}$$

From Eq. (8.75) we obtain useful relationships

$$\frac{\partial H_R}{\partial p_i} = v^i, \qquad \frac{\partial H_R}{\partial q^A} = -\frac{\partial \bar{L}}{\partial q^A}\bigg|_{v^i}, \qquad \frac{\partial H_R}{\partial v^\alpha} = -f_\alpha(q^A, p_j). \tag{8.78}$$

In particular, the first equation together with (8.76) implies that v^i is at most linear on v^α and has the representation

$$v^i(q^A, p_j, v^\alpha) = \frac{\partial H_0}{\partial p_i} - \frac{\partial f_\alpha}{\partial p_i} v^\alpha. \tag{8.79}$$

This implies (see Eq. (8.71)) that the matrix $\tilde{M}^{ij}$ is at most linear on v^α, and has the representation $\tilde{M}^{ij} = \frac{\partial^2 H_R}{\partial p_i \partial p_j}$. Finally, Eqs. (8.75), (8.76) and (8.79) imply that L is at most linear on v^α and has the representation

$$L(q^A, p_j, v^\alpha) = p_i \frac{\partial H_R}{\partial p_i} - H_R. \tag{8.80}$$

Let us summarize the results. Given singular theory, both the transition function (8.62) and the Lagrangian $L(q^A, p_i, v^\alpha)$ are linear functions of the velocities v^α. The Lagrangian has been presented in terms of the function H_R, which is a linear function of v^α as well. Besides, the Hamiltonian H_0 and the function f_α do not depend on v^α.

Using these formulas, the reader can check that Eqs. (8.64), (8.65), (8.66), (8.67), (8.68), (8.69) and (8.70) can be obtained by direct substitution of the transition function into the first-order Eqs. (8.58), (8.59) and (8.60). Therefore the two formulations are equivalent.

As in the case of the Dirac procedure, we could now start to look for, from Eq. (8.66), all the algebraic consequences of the system (8.64), (8.65) and (8.66). Instead, we now establish the equivalence of the formulation (8.64), (8.65) and (8.66) with the more symmetric Dirac formulation (8.36), (8.37), (8.38) and (8.39).

8.2.3 Comparison with the Dirac Recipe

In the previous section the first-order equations of motion (8.58), (8.59) and (8.60) were identically rewritten in special coordinates of $2[A]$ - dimensional configuration-velocity space. Let us compare the resulting system (8.64), (8.65) and (8.66) with the Dirac system (8.36), (8.37), (8.38) and (8.39). The latter is formulated in $2[A]+[\alpha]$-dimensional space, where the additional variables are the conjugate momenta p_α for q^α. So, to see the equivalence of the two formulations, we need to extend our space adding the auxiliary variables

$$(q^A, p_i, v^\alpha) \longrightarrow (q^A, p_i, p_\alpha, v^\alpha). \tag{8.81}$$

We also complete the reduced bracket (8.63) up to the Poisson bracket, defining $\{q^\alpha, p_\beta\} = \delta^\alpha{}_\beta$. Let us define the dynamics on this space as follows. By definition, the initial variables (q^A, p_j, v^α) obey Eqs. (8.64), (8.65) and (8.66). As the "equations of motion" for p_α we write the primary Dirac constraints

$$\Phi_\alpha \equiv p_\alpha - f_\alpha(q^A, p_j) = 0. \tag{8.82}$$

By construction, the system (8.64), (8.65) and (8.66), (8.82) is equivalent to (8.64), (8.65) and (8.66). On the other hand, it is equivalent to the Dirac system.

To see this, we notice the relationship between the complete Hamiltonian (8.33) of the Dirac recipe and our basic quantity H_R: $H = H_R + p_\alpha v^\alpha$, and make the following observations.

(a) We can replace $H_R \to H$ in the Eqs. (8.64) without altering their form. They then coincide with (8.37).
(b) Equation (8.65) coincides with (8.38).

(c) Using the complete Poisson bracket, Eq. (8.66) can be written in the form $\{\Phi_\alpha, H\} = 0$, that is, it represents the second-stage system of the Dirac procedure.

(d) The equation

$$\dot{p}_\alpha = -\frac{\partial H}{\partial q^\alpha} \tag{8.83}$$

is a consequence of Eqs. (8.64), (8.65) and (8.66) and (8.82), as the following computation shows:

$$\begin{aligned} 0 = \dot{\Phi}_\alpha = \dot{p}_\alpha - \frac{\partial f_\alpha}{\partial q^A}\left[\dot{q}^A - \{q^A, H\}\right] - \frac{\partial f_\alpha}{\partial p_i}\left[\dot{p}_i - \{p_i, H\}\right] \\ -\frac{\partial f_\alpha}{\partial q^A}\{q^A, H\} - \frac{\partial f_\alpha}{\partial p_i}\{p_i, H\} \\ = \dot{p}_\alpha - \{f_\alpha, H\} = \dot{p}_\alpha - \{p_\alpha, H\} + \{\Phi_\alpha, H\} = \dot{p}_\alpha - \frac{\partial H}{\partial q^\alpha}. \end{aligned} \tag{8.84}$$

Thus *we have reproduced the Dirac system (8.36), (8.37), (8.38) and (8.39), together with its consequence, which is the second-stage system* $\{\Phi_\alpha, H\} = 0$.

In the result, we have demonstrated that the Dirac procedure produces a Hamiltonian formulation which is equivalent to the initial Lagrangian formulation. The procedure of Hamiltonization can be schematically resumed as follows

$$(q^A, \dot{q}^A) \to (q^A, v^i, v^\alpha) \leftrightarrow (q^A, p_i, v^\alpha) \to (q^A, p_i, p_\alpha, v^\alpha). \tag{8.85}$$

Some relevant comments are in order.

1. The mysterious Lagrangian multipliers of the Dirac recipe represent just the velocities that remain "untouched" by the change of variables (8.61).
2. In an arbitrary singular theory, the Hamiltonian $H_0(q^A, p_i)$ does not depends on p_α.
3. Our discussion reveals that the only role played by the momenta p_α is to represent equations of motion in a completely symmetric form, with the Poisson bracket defined in relation to all variables q^A, p_A. The momenta p_α are, in fact, the auxiliary variables of a singular theory.
4. Primary and secondary[6] constraints have very different origins. The secondary constraints represent, in fact, part of the initial equations of motion, rewritten in the Hamiltonian form. The primary constraints, together with the momenta p_α, have been added by hand, with the aim of obtaining a more symmetric formalism.

[6] Remember that the secondary constraints consist of the second-stage, third-stage, ... constraints.

Exercise
Obtain Eqs. (8.64), (8.65), (8.66), (8.67), (8.68), (8.69) and (8.70) by direct substitution of the transition function into the first-order Eqs. (8.58), (8.59) and (8.60).

8.3 Classification of Constraints

Let the matrix $K_I{}^J(q^A, p_B)$ be invertible on the constraint surface

$$\det K_I{}^J \Big|_{\Phi_I=0} \neq 0. \tag{8.86}$$

Then the functions $K_I{}^J \Phi_J$ are called constraints that are *equivalent* to the initial constraints Φ_I. Note that the equations $\Phi_I = 0$ and $K_I{}^J \Phi_J = 0$ determine the same surface in the phase space.

As will be seen below, the structure of Hamiltonian equations (8.35), (8.36) and (8.46) essentially depends on properties of the constraints Φ_I with respect to the Poisson bracket.

Definition 1 The constraint Φ_{I_1} is called a *first class constraint* if its bracket with *any* constraint is proportional to the constraints

$$\{\Phi_{I_1}, \Phi_J\} \sim \Phi_L, \qquad J = 1, 2, \ldots, [J]. \tag{8.87}$$

Any subset of constraints, say Φ_{I_2}, $I_2 = 1, 2, \ldots, [I_2]$, for which the matrix of Poisson brackets is invertible on the constraint surface, is called a set of *second-class constraints*. That is they obey

$$\{\Phi_{I_2}, \Phi_{J_2}\} = \triangle_{I_2 J_2}, \qquad \text{where} \qquad \det \triangle_{I_2 J_2}\big|_{\Phi_I=0} \neq 0. \tag{8.88}$$

Now we demonstrate the central result on the structure of a constraint system. Let us denote the Poisson bracket of all constraints by $\triangle$

$$\{\Phi_I, \Phi_J\} = \triangle_{IJ}(q^A, p_B). \tag{8.89}$$

Suppose that its rank on the constraint surface is equal to a number $[I_2]$, $\operatorname{rank} \triangle_{IJ}|_{\Phi_I=0} = [I_2] < [I]$. We demonstrate that there is an equivalent system of constraints formed by $[I_2]$ second-class constraints, and by $[I_1] = [I] - [I_2]$ first-class constraints.

To see this, we note that according to the rank condition, there are $[I_1] = I - [I_2]$ independent null-vectors $\vec{K}_{I_1}$ of the matrix $\triangle$ on the surface $\Phi_I = 0$. Their components are denoted by $K_{I_1}{}^J(q^A, p_B)$. Then, for the combinations $\tilde{\Phi}_{I_1} \equiv K_{I_1}{}^J \Phi_J$, we find

$$\{\tilde{\Phi}_{I_1}, \Phi_L\} = \{K_{I_1}{}^J, \Phi_L\}\Phi_J + K_{I_1}{}^J \triangle_{JL} \sim \Phi_I, \tag{8.90}$$

since $K\triangle$ vanishes on surface $\Phi_I = 0$. Hence Φ_{I_1} represent the first-class constraints.

Exercise
Show that the presence of n first-class constraints among Φ_I implies rank $\triangle_{IJ}|_{\Phi_I=0} \le [I] - n$.

We can choose the vectors $\vec{K}_{I_2}(q^A, p_j)$ to complete $\vec{K}_{I_1}$ up to a basis of $[I]$-dimensional vector space. By construction, the matrix

$$K_I{}^J \equiv \begin{pmatrix} K_{I_1}{}^J \\ K_{I_2}{}^J \end{pmatrix}, \tag{8.91}$$

is invertible. Let us denote $\tilde{\Phi}_I \equiv (\tilde{\Phi}_{I_1}, \tilde{\Phi}_{I_2})$, where $\tilde{\Phi}_{I_1} \equiv K_{I_1}{}^J \Phi_J$, $\tilde{\Phi}_{I_2} \equiv K_{I_2}{}^J \Phi_J$. The system $\tilde{\Phi}_I$ is equivalent to the initial system of constraints Φ_I. The constraints $\tilde{\Phi}_{I_2}$ form the second-class subset (prove this!).

Therefore, properties of the set $\tilde{\Phi}_I$ can be summarized as follows

$$\begin{gathered}\{\tilde{\Phi}_I, \tilde{\Phi}_J\} = \triangle_{IJ}(q^A, p_B), \\ \{\tilde{\Phi}_{I_1}, \Phi_J\} = c_{I_1 J}{}^K(q^A, p_B)\Phi_K, \quad \{\tilde{\Phi}_{I_2}, \tilde{\Phi}_{J_2}\} = \triangle_{I_2 J_2}(q^A, p_B),\end{gathered} \tag{8.92}$$

where

$$\operatorname{rank} \triangle_{IJ}|_{\Phi_I=0} = [I_2], \qquad \det \triangle_{I_2 J_2}\big|_{\Phi_I=0} \neq 0. \tag{8.93}$$

For later use we mention one more property

$$\{\tilde{\Phi}_{I_1}, H_0\} = b_{I_1}{}^J(q^A, p_B)\Phi_J, \tag{8.94}$$

which follows from the conservation of the first-class constraints in time, Eq. (8.46).

8.4 Comment on the Physical Interpretation of a Singular Theory

After completing the Dirac procedure, we deal with equations of motion in the form

$$\dot{z} = \{z, H\}, \qquad \Phi_\alpha = 0, \qquad \{\Phi_I, H\} = 0, \tag{8.95}$$

where $z = (q^A, p_B)$ and $\Phi_I = (\Phi_\alpha, \Phi_a)$. We have also seen that some of the dynamical equations of the system are a consequence of other equations. To continue the analysis, we now separate independent equations.

Since the constraints $\Phi_I = 0$ are functionally independent, we can resolve them with respect to $[I]$ phase-space variables. That is, they can be presented in the form

$$\Omega_I \equiv z^I - f^I(z^{\bar{a}}) = 0, \tag{8.96}$$

where $[\bar{a}] = [2A] - [I]$. Just the equations for z^I are the consequences, as follows from the computation

$$\begin{aligned}
\dot{\Phi}_I &= \frac{\partial \Phi_I}{\partial z}\left[\dot{z} - \{z, H\}\right] + \frac{\partial \Phi_I}{\partial z}\{z, H\} \\
&= \frac{\partial \Phi_I}{\partial z^J}\left[\dot{z}^J - \{z^J, H\}\right] + \frac{\partial \Phi_I}{\partial z^{\bar{a}}}\left[\dot{z}^{\bar{a}} - \{z^{\bar{a}}, H\}\right] + \{\Phi_I, H\} \\
&= \frac{\partial \Phi_I}{\partial z^J}\left[\dot{z}^J - \{z^J, H\}\right].
\end{aligned} \tag{8.97}$$

Besides, the higher-stage equations allow us to present some of the velocities, say $v^{\underline{\alpha}}$, through other variables, $v^{\underline{\beta}} = d^{\underline{\beta}}{}_{\bar{\alpha}}(z)v^{\bar{\alpha}}$. Using all this, the Eqs. (8.95) are equivalent to

$$\dot{z}^{\bar{a}} = h^{\bar{a}}(z^{\bar{a}}, v^{\bar{\alpha}}), \tag{8.98}$$

$$z^I = f^I(z^{\bar{a}}), \qquad v^{\underline{\beta}} = d^{\underline{\beta}}{}_{\bar{\alpha}}(z^A)v^{\bar{\alpha}}, \tag{8.99}$$

where[7] $h^{\bar{a}} = \{z^{\bar{a}}, H(z^{\bar{a}}, z^I, v^{\underline{\alpha}}(z, v^{\bar{\alpha}}, v^{\bar{\alpha}}))\}\big|_{z^I = f^I(z^{\bar{a}})}$.

Now it is clear that the velocities $v^{\bar{\alpha}}$ are not determined by these equations. Similarly to the case of the relativistic particle, they "parameterize" the ambiguity presented in solutions of the problem. Given arbitrary functions $v^{\bar{\alpha}}(\tau)$, the Eq. (8.98) represent a normal system and can be solved under a given initial conditions, for example $z^{\bar{a}}(0) = z_0^{\bar{a}}$. Then we determine z^I, $v^{\underline{\beta}}$ according to (8.99). Solution of the Cauchy problem is not unique, since taking some other functions $v'^{\bar{\alpha}}(\tau)$ such that $v'^{\bar{\alpha}}(0) = v^{\bar{\alpha}}(0)$, we obtain a different solution which obeys the same initial conditions.

Moreover, a solution is not unique for any other type of initial conditions! To illustrate this, let us discuss the two-dimensional case

$$\dot{z} = h(z, v). \tag{8.100}$$

To put this in concrete terms, impose the initial conditions $z(0) = c$, $\dot{z}(0) = c_1$, $v(0) = b$. To find a solution with these conditions, solve the equation with respect to v, $v = \tilde{h}(z, \dot{z})$. Note that the initial conditions are consistent with the problem, if $b = \tilde{h}(c, c_1)$. It is clear that the following functions: $z = c\tau + c_1\tau + g(\tau)$,

[7] Recall that we can substitute the velocities $v^{\underline{\alpha}}(z, v^{\bar{\beta}})$ into the complete Hamiltonian before computing the Poisson brackets, which does not alter the resulting equations of motion.

$v = \tilde{h}(z, \dot{z})$, where $g(\tau)$ is an arbitrary function obeying $g(0) = \dot{g}(0) = 0$, represent a solution to the problem (8.100).

Therefore, for a singular theory, in which some of velocities have not been determined in the course of the Dirac procedure, it is impossible to formulate the Cauchy problem. For any given initial conditions, a solution of the problem is not unique.[8]

Singular theory in special coordinates. Further progress in analysis of equations of motion can be achieved in special coordinates of the phase space. They are related with the initial coordinates by a time-independent canonical transformation. In this paragraph we give a brief review of the formalism developed in [10].

As we have seen in Sect. 4.4.3, there is a canonical transformation such that the constraints Ω specified in Eq. (8.96) turn into coordinates of the new system. So, all solutions to equations of motion lie on the hyper-plane $\Omega = 0$. The set Ω can be composed of pairs of conjugated variables (hence they form a second-class subset), and by conjugated momenta for some coordinates Q, which do not belong to the set Ω (then the momenta represent all the first-class constraints) .

Accordingly, the special coordinates can be divided into three groups as follows: (ω, Ω, Q), where ω is a set of pairs of canonically conjugated variables. In these coordinates, equations of motion (8.95) acquire the form (this remarkable result has been proved in [10])

$$\dot{\omega} = \{\omega, H_{ph}(\omega)\}, \quad \Omega = 0, \quad \dot{\bar{Q}}^{\bar{\alpha}} = v^{\bar{\alpha}}, \quad \dot{\tilde{Q}} = A(\omega, \bar{Q}, \tilde{Q}), \tag{8.101}$$

where H_{ph}, A are known functions.

Exercise
Confirm that the number of first-class constraints among Φ_α is equal to the number of undetermined velocities $v^{\bar{\alpha}}$.

Now the ambiguity in solutions becomes transparent: varying the functions $v^{\bar{\alpha}}$, we alter only the variables $Q = (\bar{Q}, \tilde{Q})$. The variables ω have unambiguous dynamics, since the equations for them are separated from others. Equations for ω are Hamiltonian with the Hamiltonian being H_{ph}.

A classical-mechanics system has unambiguous behavior. So, if we wish to describe it by a singular theory, we need some convention that will remove the discrepancy between causal evolution of a physical system and the ambiguity presented in solutions of Eqs. (8.101). The form of these equations suggests the natural possibility of doing this: the physical dynamics can be associated with the ω-sector of a singular theory. In further detail, we adopt the following convention:

---

[8] Remember that *all* the variables $x^\mu(\tau)$, used for the description of the relativistic particle, have no direct physical meaning. The same can happen in the general case. Generally, neither z^A nor any part of them represent physically measurable positions and velocities.

1. Two solutions $Z = (\omega, \Omega, Q, v)$ and $Z' = (\omega', \Omega', Q', v')$ to the system (8.101) are said to be equivalent, $Z' \sim Z$, if $\omega' = \omega$. According to the equivalence relationship, the space of solutions decomposes into non-intersecting classes $\tilde{Z}$ of mutually equivalent solutions.
2. The physically measurable quantity A of the theory is a function on the space of classes, that is $Z' \sim Z$ implies, at equal-time points, $A(Z') = A(Z)$.

Note that the variables ω (as well as the functions $f(\omega)$) represent examples of physical quantities.

According to this interpretation, the Cauchy problem is formulated for the ω-variables, $\omega(0) = \omega_0$. One consequence is that it is not sufficient to fix $v^{\bar{\alpha}}$ in Eqs. (8.101) to remove the ambiguity, since, given initial conditions for ω, there remains the ambiguity related to the choice of initial conditions for the variables Q.

Another consequence of the interpretation adopted is that the same physical system can be described by theories with different numbers of variables. Since a physical quantity acquires the same value for all the representatives of a given class, it is not necessary to know all representatives for the description of physical quantities.

Consider the theory with the variables ω and the equations of motion $\dot{\omega} = \{\omega, H_{ph}\}$. It is obtained from the initial theory disregarding all the ambiguous variables. It has the same physical content as the theory (8.101). Note also that it is a non-singular theory.

Other examples of equivalent formulations can be obtained either by partial or by complete fixation of the ambiguity presented in Eqs. (8.101).

For instance, we can partially fix the dynamics (8.101), adding the equations $\bar{Q}^{\bar{\alpha}} = 0$. This implies the fixation of $v^{\bar{\alpha}}$ as well, $v^{\bar{\alpha}} = 0$.

Adding the equations $\bar{Q}^{\bar{\alpha}} = 0$, and *replacing* the last set of equations from (8.101) by $\tilde{Q} = Y(\omega)$, where $Y(\omega)$ are given functions, we completely remove the ambiguity. Note that the *number of equations that must be added to completely remove the ambiguity is equal to the number of first-class constraints of the theory*.

The transition from the initial formulation to the other one, with the variables of the ambiguous sector fixed in one or another way, is called the *fixation of gauge*. The resulting formulation is called a *gauge* of the initial theory. By construction, the initial theory and its gauge have the same physical content.

In practical computations, it is desirable to pass from the initial theory to its gauge, working in terms of initial variables. We can do this according to the following scheme. Let $F_{I_1}(z^A)$ be functions such that the functions $\boldsymbol{\Phi} = (\Phi_I, F_{I_1})$ form the second-class set on the surface $\boldsymbol{\Phi} = 0$

$$\det\{\boldsymbol{\Phi}, \boldsymbol{\Phi}\}|_{\boldsymbol{\Phi}=0} \neq 0. \tag{8.102}$$

Then the theory

$$\dot{z}^A = \{z^A, H\}, \qquad \Phi_I = 0, \qquad F_{I_1} = 0, \tag{8.103}$$

is equivalent to (8.95). The equations $F_{I_1}(z^A) = 0$ are called *gauge conditions* imposed on the theory (8.95). Thus, to fix a gauge it is sufficient to add a gauge condition to each first-class constraint of the initial formulation.

8.5 Theory with Second-Class Constraints: Dirac Bracket

Consider a singular theory that, after completing the Dirac procedure, leads to the equations of motion

$$\dot{z} = \{z, H\}, \quad \Phi_\alpha = 0, \quad \{\Phi_I, H_0\} + \{\Phi_I, \Phi_\alpha\}v^\alpha = 0, \tag{8.104}$$

where all the constrains Φ_I form the second-class set

$$\{\Phi_I, \Phi_J\} = \triangle_{IJ}, \quad \det \triangle_{IJ}|_{\Phi_K=0} \neq 0. \tag{8.105}$$

The inverse matrix for $\triangle$ is denoted by $\tilde{\triangle}^{IJ}$, $\tilde{\triangle}^{IJ}\triangle_{JK} = \delta^I{}_K$.

Applying $\tilde{\triangle}^{IJ}$ to the last equation from (8.104), we obtain

$$v^\alpha = -\tilde{\triangle}^{\alpha I}\{\Phi_\alpha, H_0\}, \tag{8.106}$$

$$\tilde{\triangle}^{aI}\{\Phi_\alpha, H_0\} \sim \Phi_I. \tag{8.107}$$

Substituting these into the first equation from (8.104), we obtain $\dot{z} = \{z, H_0\} - \{z, \Phi_\alpha\}\tilde{\triangle}^{\alpha J}\{\Phi_J, H_0\}$. Using Eq. (8.107), this can be written in a more symmetric form

$$\dot{z} = \{z, H_0\} - \{z, \Phi_I\}\tilde{\triangle}^{IJ}\{\Phi_J, H_0\}. \tag{8.108}$$

Repeating the analysis carried out in Sect. 8.4 (see Eqs. (8.96), (8.97), (8.98) and (8.99)), we conclude that the dynamics of the theory is governed by

$$\dot{z}^{\bar{a}} = h^{\bar{a}}(z^{\bar{a}}), \tag{8.109}$$

$$z^I = f^I(z^{\bar{a}}), \qquad v^\alpha = -\tilde{\triangle}^{\alpha I}\{\Phi_\alpha, H_0\}. \tag{8.110}$$

Thus, in a singular theory with second-class constraints, all the velocities are determined algebraically through the phase-space variables. The dynamics is unambiguous. The variables $z^{\bar{a}}$ obey the first-order equations, so there is a natural formulation of the Cauchy problem, $z^{\bar{a}}(0) = z_0^{\bar{a}}$. The variables z^I, v^α are determined algebraically through $z^{\bar{a}}$, so they have no independent temporal evolution.

Dirac bracket. The equations of motion (8.108) can be written in a compact form

$$\dot{z} = \{z, H_0\}_D, \tag{8.111}$$

if we introduce the *Dirac bracket* constructed with help of the set of second-class constraints Φ_I

$$\{A, B\}_D = \{A, B\} - \{A, \Phi_I\}\tilde{\triangle}^{IJ}\{\Phi_J, B\}. \tag{8.112}$$

This possesses all the properties of the Poisson bracket, see Eqs. (2.40), (2.41), (2.42) and (2.43). Besides, the remarkable property is that the Dirac bracket of the constraint Φ_I with any function vanishes

$$\{\Phi_I, A\}_D = 0, \qquad I = 1, 2, \dots, [I], \tag{8.113}$$

in particular

$$\{\Phi_I, \Phi_J\}_D = 0. \tag{8.114}$$

This implies, in particular, that second-class constraints can be taken into account inside the Dirac bracket (that is, before computing the bracket).

This can be used to write the function $h^{\bar{a}}$ in Eq. (8.109) in terms of the Dirac bracket. Indeed, let us write equations for $z^{\bar{a}}$ which are contained in (8.111), $\dot{z}^{\bar{a}} = \{z^{\bar{a}}, H_0\}_D$. Using the constraints (8.110), z^I can be substituted into H_0 before computing the Dirac bracket. In the result, the dynamical variables $z^{\bar{a}}$ obey the equations

$$\dot{z}^{\bar{a}} = \{z^{\bar{a}}, H_0(z^{\bar{a}})\}_D, \quad \text{where} \quad H_0(z^{\bar{a}}) = H_0(z^{\bar{a}}, z^I(z^{\bar{a}})). \tag{8.115}$$

Exercise
Show that the rank of the matrix of fundamental brackets $\{z^A, z^B\}_D$ on the surface $\Phi_I = 0$ is equal to $[2A] - [I]$. (Hint: compute the fundamental brackets for the special coordinates (ω, Ω, Q) specified in Sect. 8.4.

Covariant canonical quantization by means of Dirac bracket. To quantize a given nonsingular theory we need, among other things, to associate with the classical quantities $A(q, p)$ certain operators $\hat{A}$, acting on a Hilbert space whose elements describe states of the theory. According to the canonical quantization paradigm, operators of position and momentum must be chosen in such a way that their commutators resemble the fundamental Poisson brackets of classical variables[9]

---

[9] The operators $\hat{q}$, $\hat{p}$ are taken as hermitian, which guarantees that their eigenvalues are real numbers. Since the commutator of Hermitian operators is an anti-Hermitian operator, the factor i appears on the r.h.s. of Eq. (8.116).

$$\hat{q}^A \hat{p}_B - \hat{p}_B \hat{q}^A \equiv [\hat{q}^A, \hat{p}_B] = i\hbar \delta^A{}_B,$$
$$[\hat{q}^A, \hat{q}^B] = [\hat{p}_A, \hat{p}_B] = 0. \tag{8.116}$$

For a theory with second-class constraints, the recipe would not be consistent. Indeed, since in classical theory $\Phi_I = 0$, one expects that the corresponding operators vanish on physical state vectors, $\hat{\Phi}_I \Psi_{ph} = 0$. Quantizing the theory by means of the Poisson bracket, we obtain $(\hat{\Phi}_I \hat{\Phi}_J - \hat{\Phi}_J \hat{\Phi}_I)\Psi_{ph} = \triangle_{IJ} \Psi_{ph}$. The left-hand side of this expression vanishes, but the right-hand side does not.

The problem is resolved by postulating commutators that resemble the Dirac bracket instead of the Poisson one

$$[\hat{q}^A, \hat{p}_B] = i\hbar \left. \{q^A, p_B\}_D \right|_{z \to \hat{z}} = i\hbar \triangle_{AB}. \tag{8.117}$$

Owing to Eq. (8.114), this is consistent with the condition $\hat{\Phi}_I \Psi_{ph} = 0$.

Formal quantum realization of the Dirac bracket. Quantum realization of the Poisson bracket can be achieved in a well-known way

$$\hat{q}^A = q^A, \qquad \hat{p}_A = -i\hbar \frac{\partial}{\partial q^A}. \tag{8.118}$$

This implies $[\hat{q}^A, \hat{p}_B] = i\hbar \delta^A{}_B$, as should be the case. To find the quantum realization for the Dirac bracket, we associate the function A_d with a phase-space function $A(q, p)$

$$A_d(q, p) = A(q, p) - \{A, \Phi_I\} \tilde{\triangle}^{IJ} \Phi_J. \tag{8.119}$$

On the constraint surface $\Phi_I = 0$ they coincide: $A_d = A$. Note that the Poisson bracket of d -functions on the constraint surface coincides with the Dirac bracket of the initial functions

$$\{A_d, B_d\}_P = \{A, B\}_D. \tag{8.120}$$

Since the quantum realization of the Poisson bracket is known, we can now realize the Dirac bracket, associating the following operators[10] with the classical quantities $A(q, p)$

$$A(q, p) \ \to \ \hat{A}_d = A_d(q, p)|_{q \to \hat{q}, p \to \hat{p}} \tag{8.121}$$

with $\hat{q}$, $\hat{p}$ specified in Eq. (8.118). Commutators of these operators resemble the Dirac bracket.

---

[10] We do not discuss the problem of ordering of operators which must be solved in each concrete case.

In particular, operators corresponding to the phase-space variables are

$$q^A \to \hat{q}^A_d = \hat{q}^A - [\hat{q}^A, \Phi_I]\tilde{\triangle}^{IJ}\Phi_J,$$
$$p_A \to \hat{p}_{(d)A} = \hat{p}_A - [\hat{p}_A, \Phi_I]\tilde{\triangle}^{IJ}\Phi_J, \tag{8.122}$$

Note also that in this realization the constraints become null operators: $\Phi_I \to \hat{\Phi}_{(d)I} = 0$.

8.6 Examples of Theories with Second-Class Constraints

8.6.1 Mechanics with Kinematic Constraints

Consider a mechanical system with configuration variables q^a, which is confined to move on the non-degenerate surface

$$G_i(q^a) = 0, \qquad \text{rank}\,\frac{\partial G_i}{\partial q^a} = [i] < [a] \tag{8.123}$$

Suppose that in the absence of the constraints the theory is described by a nonsingular Lagrangian $L_0(q^a, \dot{q}^a)$. Then, as we have seen in Sect. 1.10, the theory can be described by an action with the Lagrangian multipliers $\lambda^i(\tau)$

$$S = \int d\tau \left[L_0(q, \dot{q}) + \lambda^i G_i(q)\right]. \tag{8.124}$$

Let us construct the Hamiltonian formulation. Since L_0 is non singular, equations for the momenta p_a, $p_a = \frac{\partial L_0}{\partial \dot{q}^a}$, can be resolved with respect to $\dot{q}^a$. Let $\dot{q}^a = f^a(q, p)$ be a solution:

$$\left.\frac{\partial L_0}{\partial \dot{q}^a}\right|_{\dot{q}=f(q,p)} \equiv p_a, \qquad \det\frac{\partial f^a}{\partial p_b} \neq 0. \tag{8.125}$$

Conjugated momenta for λ^i represent $[i]$ primary constraints of the theory, $p_{\lambda i} = 0$. Then we obtain the complete Hamiltonian

$$H = H_0 - \lambda^i G_i(q) + v^i_\lambda p_{\lambda i}, \qquad H_0 \equiv p_a f^a - L_0(q, f). \tag{8.126}$$

Conservation in time of the primary constraints: $\dot{p}_{\lambda i} = \{p_{\lambda i}, H\} = 0$ implies secondary constraints $G_i(q) = 0$. In turn, conservation of G_i, $\dot{G}_i = \{G_i, H\} = 0$, gives the third-stage constraints. Using Eq. (8.125), they read $F_i \equiv G_{ia}(q)f^a(q, p) = 0$, where $G_{ia} \equiv \frac{\partial G_i}{\partial q^a}$. The Poisson brackets of the constraints are $\{G_i, F_j\} = G_{ia}\frac{\partial f^c}{\partial p_a}G_{jc} \equiv \triangle_{ij}$. Since $\det\frac{\partial f^a}{\partial p_b} \neq 0$ and $\text{rank}\,G_{ia} = [i]$, the known theorem of linear algebra guarantees $\det\triangle_{ij} \neq 0$. The inverse matrix for $\triangle$ is denoted as $\tilde{\triangle}^{ij}$.

Further, the condition $\dot{F}_i = 0$ implies fourth-stage constraints $\lambda^i - \tilde{\triangle}^{ij}\{F_j, H_0\} = 0$. Finally, conservation in time of these constraints determines all the velocities: $v^i_\lambda = \{\tilde{\triangle}^{ij}\{F_j, H_0\}, H_0 - \lambda^k G_k\}$. Thus, the Hamiltonian formulation of a Lagrangian theory with $[i]$ kinematic constraints implies $4[i]$ second-class Hamiltonian constraints

$$p_{\lambda i} = 0, \quad G_i = 0, \quad f^a G_{ia} = 0, \quad \lambda^i - \tilde{\triangle}^{ij}\{F_j, H_0\} = 0. \tag{8.127}$$

Let us specify these results for the free particle constrained to move on a 2 -sphere of radius c. The action is

$$S = \int d^3x \left[\frac{1}{2}m(\dot{\mathbf{x}}^2 + \lambda(\mathbf{x}^2 - c^2)\right]. \tag{8.128}$$

This implies the following chain of four second-class constraints

$$p_\lambda = 0, \quad \mathbf{x}^2 - c^2 = 0, \quad \mathbf{xp} = 0, \quad \mathbf{p}^2 + 2mc^2\lambda = 0, \tag{8.129}$$

as well as $v_\lambda = 0$. Using the complete Hamiltonian

$$H = \frac{1}{2m}\mathbf{p}^2 - \lambda(\mathbf{x}^2 - c^2), \tag{8.130}$$

we obtain the equations $m\dot{x}^i = p^i$, $\dot{p}^i = 2\lambda x^i$. Using the last constraint from (8.129), this leads to closed equations for the (x, p) -sector. They read

$$m\dot{x}_i = p^i, \qquad \dot{p}_i = -\frac{\mathbf{p}^2}{mc^2}x_i. \tag{8.131}$$

These imply the following second-order equations for x:

$$c^2\ddot{x}_i = -\dot{\mathbf{x}}^2 x_i. \tag{8.132}$$

The Dirac brackets for the (x, p) -sector are

$$\{x_i, x_j\}_D = 0, \qquad \{x_i, p_j\}_D = \delta_{ij} - \frac{1}{\mathbf{x}^2}x_i x_j,$$
$$\{\pi_i, \pi_j\}_D = -\frac{1}{\mathbf{x}^2}(x_i\pi_j - x_j\pi_i). \tag{8.133}$$

Exercises

1. Obtain the equations (8.130), (8.131), (8.132) and (8.133).
2. Confirm that the Hamiltonian equations (8.131) can be written in the form $\dot{z} = \{z, H_0\}_D$.

3. Write the Lagrangian equations following from the action (8.128). Obtain Eq. (8.132) from the resulting system.
4. We know that the quantities $L_i = \epsilon_{ijk} x_j p_k$ obey the angular-momentum algebra with respect to the Poisson bracket. Confirm that the same is true for the Dirac brackets (8.133): $\{L_i, L_j\}_D = \epsilon_{ijk} L_k$. We use this observation in Sect. 8.7.2 for construction of a semiclassical model for the spin one-half particle.
5. Construct the Hamiltonian formulation for the field theory (called sigma-model on a sphere)

$$S = \int d^4x \left[\frac{1}{2} (\partial_\mu \phi^a)^2 + \lambda((\phi^a)^2 - 1) \right]. \tag{8.134}$$

8.6.2 Singular Lagrangian Action Underlying the Schrödinger Equation

In Sect. 2.10.1 we discussed the remarkable similarity existing between mathematical structures of electrodynamics and quantum mechanics. As electric and magnetic fields can be obtained from the vector potential A_a, the real and imaginary parts of the wave function can be obtained from the scalar potential ϕ, see Eq. (2.135). The real field ϕ obeys the equation

$$\hbar^2 \ddot{\phi} + (\triangle - V)^2 \phi = 0, \tag{8.135}$$

which follows from the Lagrangian action

$$S[\phi] = \int dt d^3x \left[\frac{\hbar}{2} \dot{\phi} \dot{\phi} - \frac{1}{2\hbar} [(\triangle - V)\phi]^2 \right]. \tag{8.136}$$

Here we obtain a further relationship between the Shrödinger and the scalar potential equations, following the work [39]. We show that there is a Lagrangian theory subject to second-class constraints underlying both the Schrödinger equation and the classical field theory (8.135). This possibility is based on the fact that in a theory with second-class constraints, we can take different subsets as the independent variables when we look for a solution of the constraints. For the model presented below, there are two natural possibilities to choose the independent variables. By one option, they obey the Hamiltonian equations which correspond to the theory (8.136). By the other option we reach the Schrödinger system (2.130) and (2.131).

Consider the following Lagrangian theory:

$$S[\phi, \varphi] = \int dt d^3x \left[\frac{\hbar}{2} \dot{\phi} \dot{\phi} + \frac{1}{2\hbar} \varphi^2 + \frac{1}{\hbar} \varphi (\triangle - V) \phi \right], \tag{8.137}$$

written for two real fields $\phi(t, x^i)$, $\varphi(t, x^i)$ on the given external background $V(x^i)$. This implies the Lagrangian equations

$$\hbar^2\ddot{\phi} - (\triangle - V)\varphi = 0, \qquad \varphi = -(\triangle - V)\phi. \tag{8.138}$$

As a consequence, both ϕ and φ obey the second order Eq. (8.135). After the shift $\tilde{\varphi} \equiv \varphi + (\triangle - V)\phi$, the action acquires the form $S[\phi, \varphi] = S[\phi] + \frac{1}{2\hbar}\int\tilde{\varphi}^2$, where $S[\phi]$ is the action (8.136). Hence in this parametrization the fields ϕ and $\tilde{\varphi}$ decouple, and the only dynamical variable is ϕ. Once again, its evolution is governed by Eq. (8.135). Although it is natural, this is not the only possible parametrization of the dynamical sector. To find another relevant parametrization, we would like to construct a Hamiltonian formulation of the theory. We introduce the conjugate momenta p, π for the fields ϕ, φ and define their evolution according to the standard rule

$$p = \frac{\partial L}{\partial\dot{\phi}} = \hbar\dot{\phi}, \qquad \pi = \frac{\partial L}{\partial\dot{\varphi}} = 0. \tag{8.139}$$

The second equation represents a primary constraint of the theory. Then the complete Hamiltonian is

$$H = \int d^3x\left[\frac{1}{2\hbar}(p^2 - \varphi^2) - \frac{1}{\hbar}\varphi(\triangle - V)\phi + v\pi\right], \tag{8.140}$$

Preservation in time of the primary constraint, $\dot{\pi} = \{\pi, H\} = 0$, implies the secondary one $\varphi + (\triangle - V)\phi = 0$. In turn, its preservation in time determines the velocity $v = -\frac{1}{\hbar}(\triangle - V)p$. Hence the Dirac procedure stops at this stage. Evolution of the phase-space variables is governed by the Hamiltonian equations

$$\begin{aligned} \dot{\phi} &= \frac{1}{\hbar}p, \qquad \dot{p} = \frac{1}{\hbar}(\triangle - V)\varphi, \\ \dot{\varphi} = v &= -\frac{1}{\hbar}(\triangle - V)p, \qquad \dot{\pi} = 0, \end{aligned} \tag{8.141}$$

and by the constraints

$$\pi = 0, \qquad \varphi + (\triangle - V)\phi = 0. \tag{8.142}$$

The system implies that both ϕ and φ obey Eq. (8.135). Computing the Poisson bracket of the constraints, we obtain an on-shell non-vanishing result $\{\varphi + (\triangle - V)\phi, \pi\} = \delta^3(x - y)$. So the constraints form a second-class system.

Let us construct the Dirac bracket corresponding to the constraints

$$\begin{aligned} \{A, B\}_D =& \{A, B\} - \{A, \pi\}\{\varphi + (\triangle - V)\phi, B\} \\ &+ \{A, \varphi + (\triangle - V)\phi\}\{\pi, B\}. \end{aligned} \tag{8.143}$$

This implies $\{\pi, A\}_D = 0$, $\{\phi, \varphi\}_D = 0$, as well as

$$\{\phi, p\}_D = \delta^3(x - y), \qquad \{\phi, \phi\}_D = \{p, p\}_D = 0; \tag{8.144}$$

$$\begin{aligned} &\{\varphi, p\}_D = -(\triangle - V)\delta^3(x - y), \\ &\{\varphi, \varphi\}_D = \{p, p\}_D = 0. \end{aligned} \tag{8.145}$$

Note that for the pair ϕ, p the Dirac brackets coincide with the Poisson ones. For the pair φ, p the Dirac brackets coincide exactly with the non-canonical ones, (2.145).

According to the constraints (8.142), either φ, p or ϕ, p can be taken to parameterize the dynamical sector of the theory.

Parameterizing it by the pair φ, p, the Eqs. (8.141) reduce to the Schrödinger system (2.130) and (2.131), while the Hamiltonian (8.140) acquires the form (2.144). Note that p is the conjugate momenta for ϕ but not for φ. Using this Hamiltonian and the Dirac bracket (8.145), Eqs. (2.130) and (2.131) can also be obtained according to the rule (8.115).

Parameterizing the dynamical sector by the pair ϕ, p, the Eqs. (8.141) reduce to the system $\hbar\dot{\phi} = p$, $\hbar\dot{p} = -\left(\frac{\hbar^2}{2m}\triangle - V\right)^2\phi$, while the Hamiltonian (8.140) acquires the form

$$H(\phi, p) = \int d^3x \frac{1}{2\hbar}\left[p^2 + [(\triangle - V)\phi]^2\right]. \tag{8.146}$$

This is precisely the Hamiltonian formulation of the theory (8.135), (8.136).

Hence the classical field theory (8.135) and the Schrödinger equation can be identified with two possible parameterizations of the dynamical sector of the singular Lagrangian theory (8.137).

8.7 Examples of Theories with First-Class Constraints

8.7.1 Electrodynamics

Remember that a free electromagnetic field can be described by the action

$$\begin{aligned} S &= \int d^4x \left[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\right] \\ &= \int d^4x \left[\frac{1}{2}(\partial_0 A_b - \partial_b A_0)^2 - \frac{1}{4}F_{ab}^2\right]. \end{aligned} \tag{8.147}$$

written for the four-dimensional vector potential A_μ. We have denoted $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$. In the second line we have separated the terms containing temporal derivatives.

Equations determining the conjugate momenta are

$$p^0 = \frac{\partial L}{\partial \dot{A}_0} = 0, \tag{8.148}$$

$$p^a = \frac{\partial L}{\partial \dot{A}_a} = \partial_0 A_a - \partial_a A_0, \quad \text{then} \quad \partial_0 A_a = -p_a + \partial_a A_0. \tag{8.149}$$

So there is one primary constraint (8.148). Computing the Hamiltonian $H_0 = p^a \partial_0 A_a - L$ and adding the primary constraint multiplied by the velocity v_0, the complete Hamiltonian reads

$$H = \int d^3x \left[\frac{1}{2} p_a^2 - p_a \partial_a A_0 + \frac{1}{4} F_{ab}^2 + v_0 p_0 \right]. \tag{8.150}$$

Conservation in time of the primary constraint produces the second-stage constraint[11]

$$\begin{aligned} \dot{p}^0(x) &= \{p^0(x), H\} = \{p^0(x), -\int d^3x p_a \partial_a A_0\} \\ &= \int d^3y \{p^0(x), \partial_a p_a(y) A_0(y)\} \\ &= -\int d^3x \partial_a p_a(y) \delta^3(x-y) = -\partial_a p_a(x) = 0. \end{aligned} \tag{8.151}$$

Carrying out a similar computation, the reader can verify that it preserves in time, $\{\partial_a p_a, H\} = 0$, so the Dirac procedure stops at the second stage. In the result, the evolution is governed by the Hamiltonian equations

$$\dot{A}_0 = v_0, \qquad \dot{p}^0 = 0, \tag{8.152}$$

$$\dot{A}_a = -p_a + \partial_a A_0, \qquad \dot{p}_b = -\partial_a F_{ab}. \tag{8.153}$$

These are accompanied by two first-class constraints

$$p^0 = 0, \qquad \partial_a p_a = 0. \tag{8.154}$$

Exercise
Show that the Lagrangian equations $\partial_\mu F^{\mu\nu} = 0$ follow from the system (8.153) and (8.154).

[11] Poisson bracket in field theory is defined by $\{A(x), B(y)\} = \int d^3z \left[\frac{\delta A(x)}{\delta \phi^A(z)} \frac{\delta B(y)}{\delta p_A(z)} - \frac{\delta A(x)}{\delta p_A(z)} \frac{\delta B(y)}{\delta \phi^A(z)} \right]$. A and B are taken at the same instance of time. The working formula for computing the variational derivative is $\frac{\delta A(\phi(x), \partial_b \phi(x))}{\delta \phi^A(z)} = \left. \frac{\partial A}{\partial \phi^A} \right|_{\phi \to \phi(x)} \delta^3(x-z) + \left. \frac{\partial A}{\partial \partial_b \phi^A} \right|_{\phi \to \phi(x)} \frac{\partial}{\partial x b} \delta^3(x-z)$.

According to Sect. 8.4, the unique representative in a class of equivalent trajectories can be obtained imposing two gauge conditions. They can be taken as

$$A_0 = 0, \qquad \partial_a A_a = 0. \tag{8.155}$$

In this gauge, the Eqs. (8.153) and (8.155) imply the wave equation for the three-dimensional vector potential.

8.7.2 Semiclassical Model for Description of Non Relativistic Spin

The data of some experiments with elementary particles and atoms (Stern–Gerlach experiment, fine structure of hydrogen atom, Zeeman effect) shows that the Schrödinger equation for a one-component wave function is not adequate to describe the behavior of these systems in the presence of a magnetic field. This implies a radical modification of the formalism, see, for example, the book [40] for a detailed discussion. Roughly speaking, besides the position and the momentum, the state of an electron is specified by some additional numbers, which are eigenvalues of the properly defined operators. The mathematical theory of these operators is similar to the formalism of angular momentum. So, intuitively, an elementary particle can carry an intrinsic angular momentum called spin.

To describe a particle with spin $s = \frac{1}{2}$ (electron, proton, neutron), in quantum mechanics we introduce the two-component wave function Ψ_α, $\alpha = 1, 2$. The spin operators $\hat{J}_i$ act on Ψ_α as 2×2 -matrices, and are defined by

$$\hat{J}_i = \frac{\hbar}{2}\sigma_i, \tag{8.156}$$

where σ_i stands for the *Pauli matrices*, which we choose traceless and Hermitean

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8.157}$$

We also recall their basic algebraical properties

$$\sigma_i\sigma_j = i\epsilon_{ijk}\sigma_k + \mathbf{1}\delta_{ij}, \tag{8.158}$$

$$\sigma_i\sigma_j + \sigma_j\sigma_i = 2 \times \mathbf{1}\delta_{ij}, \tag{8.159}$$

$$\sigma_i\sigma_j - \sigma_j\sigma_i \equiv [\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k, \tag{8.160}$$

$$(\sigma_i)^2 = \mathbf{1}, \quad \text{for any fixed} \quad i,$$

$$\sum_i \sigma_i\sigma_i = 3 \times \mathbf{1}, \quad [\sigma_k, \sum_i \sigma_i\sigma_i] = 0. \tag{8.161}$$

Note that the commutators (8.160) of σ -matrices are the same as for the angular-momentum operators. The spin operators, being proportional to the Pauli matrices, have similar properties, in particular

$$[\hat{J}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{J}_k, \tag{8.162}$$

$$\hat{\mathbf{J}}^2 = \hbar^2 s(s+1) \times \mathbf{1} = \frac{3\hbar^2}{4}\mathbf{1}. \tag{8.163}$$

Evolution of the spin one-half particle in a given electromagnetic field is described by the *Pauli equation*

$$i\hbar\frac{\partial\Psi}{\partial t} = \left(\frac{1}{2m}(\hat{\mathbf{p}} - \frac{e}{c}\mathbf{A})^2 - eA_0 - \frac{e}{mc}\mathbf{B}\hat{\mathbf{J}}\right)\Psi. \tag{8.164}$$

To formulate the problem that we wish to discuss in this section, remember that the quantum mechanics of a spinless particle can be obtained applying the canonical quantization procedure to a classical-mechanics system with the Lagrangian $L = \frac{1}{2}mx^2 - U(x)$. Roughly speaking, it works as follows. We construct a Hamiltonian formulation for the system, then associate with the phase-space variables the operators with commutators resembling the Poisson brackets, and write on this base the Schrödinger equation $i\hbar\dot{\Psi} = \hat{H}\Psi$.

It is natural to ask whether this ideology can be realized for the Pauli equation as well. Since the quantum-mechanical description of a spin implies the use of three extra operators $\hat{J}_i$, the problem can be formulated as follows. We look for a classical-mechanics system which, besides the position variables x_i, contains additional degrees of freedom, appropriate for the description of a spin: in the Hamiltonian formulation the spin space should be parameterized by three variables that must obey the following classical algebra: $\{J_i, J_j\} = \epsilon_{ijk}J_k$. Then canonical quantization of these variables will give the spin operators (8.162). The problem is that, in Hamiltonian formalism, the initial phase-space algebra is always Poisson algebra. An algebra, different from the Poisson one, can be obtained in the framework of a constrained system.[12] So, we expect that the desired algebra can be produced by an appropriately chosen singular Lagrangian for the description of the spin sector.

In short, our aim will be to construct a dynamical system that, *in the end*, admits three degrees of freedom J_i as the spin-space basic variables, which obey the angular-momentum algebra and the condition (8.163). In this section we follow the work [41].

Construction of the spin space. Consider the canonical pairs $\omega_i, \pi_j, i, j = 1, 2, 3$ with the Poisson bracket algebra being[13]

[12] There is an elegant formalism developed by Berezin and Marinov [42] based on using anticommuting variables for semi-classical description of spin. We present here a formulation based on commuting variables, without appealing to the rather formal methods of Grassmann mechanics.

[13] It is well known that these six variables can be used to construct the quantities, $L_i = \epsilon_{ijk}\omega_j\pi_k$, which obey the angular-momentum algebra. The problem is that, according to (8.163), we need a space with two degrees of freedom instead of six.

$$\{\omega_i, \pi_j\} = \delta_{ij}. \tag{8.165}$$

To arrive at the desired spin space, we restrict the initial system to lying on a $d = 2$ surface of the six dimensional phase space. This will be done in two steps. First we constrain the coordinates to lying on a $d = 4$ surface specified by

$$\omega^2 = a^2, \qquad \omega\pi = 0, \qquad a = \text{const}. \tag{8.166}$$

The constraints form a second-class system, $\{\omega^2 - a^2, \omega\pi\} = 2\omega^2$. So we take them into account by transition from the Poisson to Dirac bracket, $\{,\}_{D1}$, which reads

$$\{A, B\}_{D1} = \{A, B\} - \{A, \omega\pi\}\frac{1}{2\omega^2}\{\omega^2, B\} \\ +\{A, \omega^2\}\frac{1}{2\omega^2}\{\omega\pi, B\}. \tag{8.167}$$

For the phase-space coordinates this implies the algebra

$$\{\omega_i, \omega_j\}_{D1} = 0, \qquad \{\omega_i, \pi_j\}_{D1} = \delta_{ij} - \frac{1}{\omega^2}\omega_i\omega_j, \tag{8.168}$$

$$\{\pi_i, \pi_j\}_{D1} = -\frac{1}{\omega^2}(\omega_i\pi_j - \omega_j\pi_i). \tag{8.169}$$

The constraints are consistent with the Dirac algebra, that is $\{A, \omega^2 - a^2\}_{D1} = 0$, $\{A, \omega\pi\}_{D1} = 0$ for any phase-space function $A(\omega, \pi)$. So we can resolve the constraints, keeping four independent coordinates and the corresponding algebra.

We introduce the coordinates that are convenient for canonical quantization of the system. Consider the quantities

$$J_i \equiv \epsilon_{ijk}\omega_j\pi_k, \qquad \tilde{\pi}_1 = \pi_1, \qquad \tilde{\pi}_2 = \pi_2, \qquad s = \omega_i\pi_i. \tag{8.170}$$

The quantities J_i obey $SO(3)$ angular-momentum algebra with respect to both the Poisson and the Dirac brackets. Equations (8.170) can be inverted

$$\pi_1 = \tilde{\pi}_1, \qquad \pi_2 = \tilde{\pi}_2, \qquad \pi_3 = -\frac{J_1\tilde{\pi}_1 + J_2\tilde{\pi}_2}{J_3}, \tag{8.171}$$

$$\omega_i = \frac{1}{\pi^2}(\epsilon_{ijk}\pi_j J_k + s\pi_i), \tag{8.172}$$

where π_i in the last equality are given by (8.171). Hence $J_i, \tilde{\pi}_1, \tilde{\pi}_2, s$ can be used as coordinates of the six-dimensional space instead of ω_i, π_i. Equations of the surface (8.166) in these coordinates acquire the form $J^2 = a^2(\tilde{\pi}_1^2 + \tilde{\pi}_2^2 + \pi_3^2)$, $s = 0$.

Let us compute the Dirac bracket (8.167) for the new coordinates. To write it in a compact form, the Eq. (8.171) prompts us to *denote* $\tilde{\pi}_3(J_i, \tilde{\pi}_\alpha) \equiv -\frac{J_1\tilde{\pi}_1 + J_2\tilde{\pi}_2}{J_3}$ (or, equivalently, $J\tilde{\pi} = 0$). The Dirac brackets of $J_i, \tilde{\pi}_i$ are (the identity $\epsilon_{ijk}\epsilon_{iab} = \delta_{ja}\delta_{kb} - \delta_{jb}\delta_{ka}$ is used)

$$\{J_i, J_j\}_{D1} = \epsilon_{ijk} J_k, \tag{8.173}$$

$$\{\tilde{\pi}_i, \tilde{\pi}_j\}_{D1} = -\frac{1}{a^2}\epsilon_{ijk} J_k, \qquad \{J_i, \tilde{\pi}_j\}_{D1} = \epsilon_{ijk}\tilde{\pi}_k, \tag{8.174}$$

while the Dirac bracket of s with any other coordinate vanishes. Since $s = 0$ on the constraint surface, it can be omitted from consideration. In short, the surface (8.166) is now described by coordinates $J_i, \tilde{\pi}_i$ that are constrained by

$$J\tilde{\pi} = 0, \tag{8.175}$$

$$J^2 = a^2\tilde{\pi}^2, \tag{8.176}$$

and obey the algebra[14] (8.173) and (8.174). Note also that the Dirac bracket of the constraints (8.175) and (8.176) with any phase-space quantity vanishes.

We already have the desired algebra (8.173), but in the system with four independent variables. To improve this, we impose two more second-class constraints and construct the corresponding Dirac bracket. To guarantee that the bracket does not modify its form for J_i, the Eq. (8.173), one of the constraints must give the vanishing $D1$-bracket with J_i. The only possibility[15] is to take (a function of) J^2. As another constraint, we can take any phase-space function that forms a second-class system with J^2. The ambiguity in choosing the second constraint suggests that the corresponding dynamical realization will be a locally-invariant theory with first-class constraints, see below. Let us take, for example, the constraints

$$J^2 - \frac{3\hbar^2}{4} = 0, \qquad \epsilon_{3jk}\tilde{\pi}_j J_k = 0, \text{ (that is, } \omega_3 = 0). \tag{8.177}$$

Using their bracket, $\{J^2 - \frac{3\hbar^2}{4}, \epsilon_{3jk}\tilde{\pi}_j J_k\}_{D1} = -2\tilde{\pi}_3 J^2$, we obtain the $D2$-Dirac bracket

$$\begin{aligned}\{A, B\}_{D2} =&\{A, B\}_{D1} - \{A, J^2\}_{D1}\frac{1}{2\tilde{\pi}_3 J^2}\{\epsilon_{3jk}\tilde{\pi}_j J_k, B\}_{D1} \\ &+ \{A, \epsilon_{3jk}\tilde{\pi}_j J_k\}_{D1}\frac{1}{2\tilde{\pi}_3 J^2}\{J^2, B\}_{D1}.\end{aligned} \tag{8.178}$$

In the result we have a two-dimensional surface determined by the constraints

$$J\tilde{\pi} = 0, \qquad \tilde{\pi}^2 = \frac{3\hbar^2}{4a^2}, \qquad \tilde{\pi}_1 J_2 - \tilde{\pi}_2 J_1 = 0, \tag{8.179}$$

---

[14] We point out that when $a^2 = 1$, it is precisely the Lorentz-group algebra written in terms of the rotation J and the Lorentz boost $\tilde{\pi}$ generators.

[15] $SO(3)$ -algebra has only one Casimir operator J^2.

$$J^2 = \frac{3\hbar^2}{4}. \tag{8.180}$$

Equations (8.179) can be used to exclude all $\tilde{\pi}_i$

$$\vec{\tilde{\pi}} = \left(-\frac{J_1 J_3}{a\sqrt{J_1^2 + J_2^2}}, -\frac{J_2 J_3}{a\sqrt{J_1^2 + J_2^2}}, \frac{1}{a}\sqrt{J_1^2 + J_2^2} \right). \tag{8.181}$$

Then we deal with the remaining variables J_i obeying the desired algebra (8.173), and subject to the constraint (8.180). As has been discussed above, their quantization leads to the spin operators (8.162), (8.163).

Semiclassical Lagrangian action for the spin one-half particle. We start with discussion of the spin part of a Lagrangian. The constraints (8.166) plus the second constraint from Eq. (8.179) prompt us to write[16]

$$L_{spin} = \frac{1}{2g}\dot{\omega}^2 + g\frac{b^2}{2a^2} + \frac{1}{\phi}(\omega^2 - a^2), \quad b^2 = \frac{3\hbar^2}{4}, \tag{8.182}$$

Here $g(t)$, $\phi(t)$ are auxiliary degrees of freedom. This leads to the complete Hamiltonian

$$H = \frac{g}{2}(\pi^2 - \frac{b^2}{a^2}) - \frac{1}{\phi}(\omega^2 - a^2) + v_g\pi_g + v_\phi\pi_\phi. \tag{8.183}$$

where π_g, π_ϕ are conjugate momenta for g, ϕ and v_g, v_ϕ represent the velocities for the primary constraints $\pi_g = 0$, $\pi_\phi = 0$.

The Dirac procedure results in a constraint system that can be described as follows. The variables g, ϕ are subject to second-class constraints $\pi_\phi = 0$, $\frac{gb^2}{a^2} + \frac{2a^2}{\phi} = 0$ and to the primary first class constraint $\pi_g = 0$. This is associated with the local symmetry of the action (8.182) (with the parameter being an arbitrary function $\alpha(t)$)

$$\delta\omega_i = \alpha\dot{\omega}_i, \qquad \delta g = (\alpha g)^{\cdot}, \qquad \delta\phi = \alpha\dot{\phi} - \dot{\alpha}\phi. \tag{8.184}$$

Imposing the gauge $g = 1$, the variables g, ϕ, π_g, π_ϕ can be omitted from consideration.

We also obtain the desired constraints

$$\omega^2 = a^2, \qquad \omega\pi = 0, \tag{8.185}$$

[16] The first two terms resemble the relativistic particle Lagrangian $\frac{1}{2e}\dot{x}^2 + \frac{em^2}{2}$, the last one implies $p^2 = m^2$.

$$\pi^2 - \frac{3\hbar^2}{4a^2} = 0. \tag{8.186}$$

The pair (8.185) is second-class, while (8.186) represents the first-class constraint.[17]

The complete action for the spin one-half particle is given by

$$S = \int dt \left[\frac{m}{2}\dot{x}^2 + \frac{e}{c}A_i\dot{x}^i + eA_0 \right.$$
$$\left. + \frac{1}{2g}(\dot{\omega}_i - \frac{e}{mc}\epsilon_{ijk}\omega_j B_k)^2 + g\frac{b^2}{2a^2} + \frac{1}{\phi}(\omega^2 - a^2) \right]. \tag{8.187}$$

Here x_i, $i = 1, 2, 3$, stands for the spatial coordinates of the particle, and $\mathbf{B} = \nabla \times \mathbf{A}$. The second and third terms represent the minimal interaction with the vector potential A_0, A_i of an external electromagnetic field, while the fourth term contains the interaction of the spin with the magnetic field. In the end, this produces the Pauli term in the quantum-mechanical Hamiltonian.

As has been discussed above, the constraints presented in the model allow us to describe it in terms of the variables x_i, its conjugate momenta p_i, and the spin vector $J_i = \epsilon_{ijk}\omega_j\pi_k$ (we point out that in contrast to ω_i, π_i, the spin vector J_i is a gauge-invariant quantity with respect to the local symmetry (8.184)).

Note that the action leads to a reasonable classical theory. Indeed, in terms of the gauge-invariant variables x, J, the classical dynamics is governed by the Lagrangian equations

$$m\ddot{x}_i = eE_i + \frac{e}{c}\epsilon_{ijk}\dot{x}_j B_k + \frac{e}{mc}J_k\partial_i B_k, \tag{8.188}$$
$$\dot{J}_i = \frac{e}{mc}\epsilon_{ijk}J_j B_k, \tag{8.189}$$

where $\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} + \nabla A_0$. Since $J^2 \approx \hbar^2$, the J-term disappears from Eq. (8.188) at the classical limit $\hbar \to 0$. Then Eq. (8.188) reproduces the classical motion in an external electromagnetic field. Note too that in the absence of interaction, the spinning particle does not experience an undesirable self-acceleration. Equation (8.189) describes the classical spin precession in an external magnetic field.

In the Hamiltonian formulation, taking into account the constraints that are present, we obtain the only non-vanishing Dirac brackets $\{x_i, p_j\} = \delta_{ij}$, $\{J_i, J_j\} = \epsilon_{ijk}J_k$, and the Hamiltonian

$$H = \frac{1}{2m}(\mathbf{p} - \frac{e}{c}\mathbf{A})^2 - \frac{e}{mc}\mathbf{BJ} - eA_0. \tag{8.190}$$

Hence canonical quantization of the model implies the Pauli equation

[17] More exactly, the first-class constraint is given by the combination $\pi^2 - \frac{3\hbar^2}{4a^2} + \frac{3\hbar^2}{4a^4}(v^2 - a^2) = 0$.

$$i\hbar\frac{\partial\Psi}{\partial t} = \left(\frac{1}{2m}\left(\hat{\mathbf{p}} - \frac{e}{c}\mathbf{A}\right)^2 - eA_0 - \frac{e}{mc}\mathbf{B}\hat{\mathbf{J}}\right)\Psi. \tag{8.191}$$

8.8 Local Symmetries and Constraints

As we have seen in various examples, local symmetry of a Lagrangian action and first-class constraints of the corresponding Hamiltonian formulation represent characteristic properties of a degenerate action. The aim of the following sections is to establish a detailed relationship between them for a general degenerate action. It is instructive to demonstrate the relation on an illustrative example. Consider the relativistic particle Lagrangian $L = \sqrt{(\dot{x}^\mu)^2}$. This implies the Hamiltonian constraint $T \equiv \frac{1}{2}(p^2 - 1)$, as well as the local symmetry of the Lagrangian action, $\delta x^\mu = \epsilon\frac{\dot{x}^\mu}{\sqrt{\dot{x}^2}}$. This can be rewritten as follows

$$\delta x^\mu = \epsilon\{x^\mu, T\}\Big|_{p_\mu \to \frac{\partial L}{\partial \dot{x}^\mu}}, \tag{8.192}$$

So the local symmetry can be constructed with the help of the constraint, and represents the Lagrangian counterpart of the canonical transformation generated by the constraint on the phase space. It would be interesting to find an appropriate generalization of this recipe for obtaining local symmetry in general case. Since the Hamiltonian constraints can be found in the course of the Dirac procedure, this would give a regular method for obtaining the symmetries.

The general form of infinitesimal local symmetry is

$$\delta q^B = \epsilon^a R_{(0)a}{}^B + \dot{\epsilon}^a R_{(1)a}{}^B + \ddot{\epsilon}^a R_{(2)a}{}^B + \ldots + \overset{(N-1)}{\epsilon}{}^a R_{(N-1)a}{}^B, \tag{8.193}$$

where $\overset{(k)}{\epsilon}{}^a \equiv \frac{d^k\epsilon^a}{d\tau^k}$. It will be called $\overset{(N-1)}{\epsilon}$-type symmetry. $\dot{\epsilon}$-type symmetry is called *gauge symmetry*. The set of functions $R_{(k)a}{}^B(q, \dot{q}, \ldots)$ is called the generator of the symmetry.

In this section we suppose that a symmetry is known, and discuss the restrictions that this fact implies on the Hamiltonian formulation of a theory (for a much more detailed analysis, see [43, 44]). In the subsequent sections we develop the so-called formalism of extended Lagrangian, which allows us to analyze the inverse task: how the local symmetries can be reconstructed from the known system of constraints.

Analysis of the general case (8.193) implies rather tedious algebraic manipulations, see [45]. So we restrict ourselves to the simplest case of $\dot{\epsilon}$-type symmetry. This is sufficient to illustrate all the affirmations that remain true for the general case as well. For the convenience of the reader, we first summarize the affirmations.

Consider infinitesimal local transformations with at most one derivative acting on the parameters $\epsilon^a(\tau)$

$$\delta q^A = \epsilon^a R_{0a}{}^A(q, \dot{q}) + \dot{\epsilon}^a R_{1a}{}^A(q, \dot{q}), \tag{8.194}$$

and suppose that an action is invariant

$$\delta S = \int d\tau (\epsilon^a \omega_{0a} + \dot{\epsilon}^a \omega_{1a})^{\cdot}, \tag{8.195}$$

where ω_{0a}, ω_{1a} are some functions. Then

1. The quantities $R_{1a}{}^A$ represent null-vectors of the Hessian matrix

$$R_{1a}{}^A M_{AB} \equiv 0, \tag{8.196}$$

   that is we are dealing with a singular theory.
2. There are the following identities among equations of motion

$$\left(\frac{\delta S}{\delta q^A} R_{1a}{}^A\right)^{\cdot} - \frac{\delta S}{\delta q^A} R_{0a}{}^A \equiv 0. \tag{8.197}$$

3. As any other Lagrangian quantity, the identities can be rewritten in terms of the coordinates q^A, p_B, v^α of the extended phase space. The result is the following system

$$\begin{gathered} R_{1a}{}^i = \{q^i, \Phi_\alpha\} R_{1a}{}^\alpha, \\ R_{0a}{}^i = \{q^i, \Phi_\alpha\} R_{0a}{}^\alpha - \left\{q^i,\ R_{1a}{}^\alpha \{\Phi_\alpha, H\}\right\}. \end{gathered} \tag{8.198}$$

$$\frac{\partial}{\partial v^\beta}(R_{1a}{}^\alpha \{\Phi_\alpha, H\}) \equiv 0, \tag{8.199}$$

$$R_{0a}{}^\alpha \{\Phi_\alpha, H\} - \left\{R_{1a}{}^\alpha \{\Phi_\alpha, H\}, H\right\} \equiv 0. \tag{8.200}$$

   The equations have a simple meaning. Equation (8.198) states that in arbitrary theory not all the generators are independent: the i-generators $R_{0a}{}^i$, $R_{1a}{}^i$ are expressed through the α-generators.

Remember that $\{\Phi_\alpha, H\} = 0$ is the second-stage algebraic system of the Dirac procedure, see Eq. (8.40). So, Eq. (8.199) states that the combinations $T_a \equiv R_{1a}{}^\alpha \{\Phi_\alpha, H\}$ do not depend on v^α and thus represent $[a]$ second-stage constraints.

Equation (8.200) involves the Poisson bracket of these constraints with H, so the resulting quantity $\{T_a, H\}$ is a part of the third-stage algebraic system of the Dirac procedure. Hence Eq. (8.200) states that this part of the third-stage system coincides with the combinations $R_{0a}{}^\alpha \{\Phi_\alpha, H\}$ of the second-stage system. That is, the constraints T_a do not produce new constraints or equations for determining the velocities.

The symmetry transformations (8.194) can be used to construct local symmetry of the first-order action (2.155)

$$\delta q^A = \epsilon^a \bar{R}_{0a}{}^A + \dot{\epsilon}^a \bar{R}_{1a}{}^A,$$
$$\delta p_A = \frac{\partial^2 \bar{L}}{\partial q^A \partial v^B} \delta q^B + \epsilon^a \frac{\partial}{\partial q^A} (\bar{K}_A \bar{R}_{1a}{}^A),$$
$$\delta v^A = (\delta_\epsilon q^A)^{\cdot}, \tag{8.201}$$

where $\bar{A} = A(q^A, \dot{q}^A)|_{\dot{q} \to v}$. They can be used to construct local symmetry of the Hamiltonian action $p\dot{q} - H$ as well:

$$\delta q^A = \{q^A, G\}, \quad \delta p_A = \{p_A, G\},$$
$$\delta v^\alpha = \{H, \epsilon^a R_{0a}{}^\alpha + \dot{\epsilon}^a R_{1a}{}^\alpha\}, \tag{8.202}$$

where

$$G = (\epsilon^a R_{0a}{}^\alpha + \dot{\epsilon}^a R_{1a}{}^\alpha) \Phi_\alpha - \epsilon^a T_a. \tag{8.203}$$

Hence the infinitesimal Lagrangian symmetry, rewritten in the Hamiltonian form, represents, in the sector q^A, p_A, a canonical transformation with the generator G constructed from the primary Φ_α and the secondary T_a constraints.

Analogous affirmations hold for the general case (8.193) as well. In particular, $\overset{(N-1)}{\epsilon}$-type symmetry implies the appearance of constraints at the N-th stage of the Dirac procedure.

In the rest of this section we demonstrate the affirmations made above.

Lagrangian identities in first-order formalism. To analyze Eq. (8.195), we write it in the form of a power series with respect to derivatives of ϵ^a

$$\int d\tau \left[\frac{\partial L}{\partial q^A} R_{0a}{}^A + \frac{\partial L}{\partial \dot{q}^A} \dot{R}_{0a}{}^A \right] \epsilon^a +$$
$$\left[\frac{\partial L}{\partial q^A} R_{1a}{}^A + \frac{\partial L}{\partial \dot{q}^A} \left(R_{0a}{}^A + \dot{R}_{1a}{}^A \right) \right] \dot{\epsilon}^a +$$
$$\ddot{\epsilon}^a \frac{\partial L}{\partial \dot{q}^A} R_{1a}{}^A = \int d\tau (\dot{\omega}_{0a} \epsilon^a + (\omega_{0a} + \dot{\omega}_{1a}) \dot{\epsilon}^a + \omega_{1a} \ddot{\epsilon}^a).$$

Since it is fulfilled for an arbitrary $\epsilon^a(\tau)$, we have

$$\frac{\partial L}{\partial \dot{q}^A} R_{1a}{}^A = \omega_{1a}, \tag{8.204}$$
$$\frac{\partial L}{\partial q^A} R_{1a}{}^A + \frac{\partial L}{\partial \dot{q}^A} R_{0a}{}^A + \frac{\partial L}{\partial \dot{q}^A} \dot{R}_{1a}{}^A = \omega_{0a} + \dot{\omega}_{1a}, \tag{8.205}$$
$$\frac{\partial L}{\partial q^A} R_{0a}{}^A + \frac{\partial L}{\partial \dot{q}^A} \dot{R}_{0a}{}^A = \dot{\omega}_{0a}. \tag{8.206}$$

Substitution of Eq. (8.204) into (8.205) gives the expression for ω_{0a}

$$\frac{\delta S}{\delta q^A} R_{1a}{}^A + \frac{\partial L}{\partial \dot{q}^A} R_{0a}{}^A = \omega_{0a}, \tag{8.207}$$

which can be used in Eq. (8.206) and gives the Noether identities in the form

$$\left(\frac{\delta S}{\delta q^A} R_{1a}{}^A\right)^{\cdot} - \frac{\delta S}{\delta q^A} R_{0a}{}^A \equiv 0. \tag{8.208}$$

This expression can be presented in the form of a power series with respect to derivatives of q^A. It is convenient to introduce the notation

$$K_{ia}(q, \dot{q}) \equiv R_{ia}{}^A K_A, \quad i = 1, 2, \tag{8.209}$$

where K_A is the right-hand side of Lagrangian equations, see (1.115). Then the series looks like

$$\left[K_{0a} - \dot{q}^C \frac{\partial}{\partial q^C} K_{1a}\right] - \ddot{q}^A \left[M_{AB} R_{0a}{}^B + \frac{\partial}{\partial \dot{q}^A} K_{1a} + \left(\dot{q}^C \frac{\partial}{\partial q^C} + \ddot{q}^C \frac{\partial}{\partial \dot{q}^C}\right) M_{AB} R_{1a}{}^B\right] + \overset{(3)}{q}{}^A \left[M_{AB} R_{1a}{}^B\right] \equiv 0. \tag{8.210}$$

Since this is true for any $q^A(\tau)$, the square brackets in Eq. (8.210) must vanish separately. This gives the final form of the Lagrangian identities. Since they are fulfilled for any $q^A(\tau)$, they will remain identities after the substitution $\dot{q}^A(\tau) \longrightarrow v^A(\tau)$. In the result we obtain identities of first-order formalism

$$\bar{M}_{AB}(q, v) \bar{R}_{1a}{}^B(q, v) \equiv 0, \tag{8.211}$$

$$\bar{M}_{AB} \bar{R}_{0a}{}^B + \frac{\partial}{\partial v^A} \bar{K}_{1a} \equiv 0, \tag{8.212}$$

$$\bar{K}_{0a} - v^B \frac{\partial}{\partial q^B} \bar{K}_{1a} \equiv 0. \tag{8.213}$$

Hamiltonian form of the identities. Let us obtain the Hamiltonian form of the identities, i.e. we perform substitution of the velocities $v^i(q^A, p_j, v^\alpha)$, see (8.30), into Eqs. (8.211), (8.212) and (8.213). We first mention an auxiliary formula

$$-\left.\frac{\partial^2 \bar{L}}{\partial q^B \partial v^A} v^B \bar{R}^A\right|_{v^i} = \frac{\partial H}{\partial p_A} \frac{\partial \Phi_\beta}{\partial q^A} R^\beta + v^B \frac{\partial v^i}{\partial q^B} M_{iA} R^A. \tag{8.214}$$

Here $\bar{R}^A(q, v)$ is any function. If it is a null vector of the matrix $\bar{M}_{AB}$: $\bar{M}_{AB} \bar{R}^B = 0$, the formula acquires the form

$$\bar{R}^A \left.\left(\frac{\partial \bar{L}}{\partial q^A} - \frac{\partial^2 \bar{L}}{\partial q^B \partial v^A} v^B\right)\right|_{v^i} = R^\alpha \{\Phi_\alpha, H\}. \tag{8.215}$$

In accordance with our division of the index: $A = (i, \alpha)$, Eq. (8.211) can be rewritten as

$$\bar{R}_{1a}{}^{i} = -\tilde{\bar{M}}^{ij}\bar{M}_{j\alpha}\bar{R}_{1a}{}^{\alpha}, \tag{8.216}$$

$$(\bar{M}_{\alpha\beta} - \bar{M}_{\alpha i}\tilde{\bar{M}}^{ij}\bar{M}_{j\beta})\bar{R}_{1a}{}^{\beta} = 0. \tag{8.217}$$

Substituting the velocities v^i into (8.216), it reads

$$R_{1a}{}^{i} = \{q^i, \Phi_\alpha\}R_{1a}{}^{\alpha}, \tag{8.218}$$

while Eq. (8.217) holds automatically, see Eq. (8.55). Similarly, Eq. (8.212) is equivalent to the pair

$$R_{0a}{}^{i} \equiv -\tilde{M}^{ij}M_{j\alpha}R_{0a}{}^{\alpha} - \frac{\partial}{\partial p_i}(K_{1a}), \tag{8.219}$$

$$\frac{\partial}{\partial v^\beta}(K_{1a}) \equiv 0, \tag{8.220}$$

where Eqs. (8.216) and (8.217) were used. By using Eqs. (8.71), (8.79) and (8.215) we find finally

$$R_{0a}{}^{i} \equiv \{q^i, \Phi_\alpha\}R_{0a}{}^{\alpha} - \{q^i, R_{1a}{}^{\alpha}\{\Phi_\alpha, H\}\}, \tag{8.221}$$

$$\frac{\partial}{\partial v^\beta}\left(R_{1a}{}^{\alpha}\{\Phi_\alpha, H\}\right) \equiv 0. \tag{8.222}$$

To substitute the multipliers $v^i(q^A, p_j, v^\alpha)$ into the first term of Eq. (8.213) we use Eqs. (8.78), (8.214) and (8.219), with the result being

$$\begin{aligned} K_{0a} =& R_{0a}{}^{\alpha}\{\Phi_\alpha, H\} + \frac{\partial H}{\partial q^A}\frac{\partial}{\partial p_A}(R_{1a}{}^{\alpha}\{\Phi_\alpha, H\}) \\ &+ v^B\frac{\partial v^i}{\partial q^B}M_{iA}R_{0a}{}^{A}. \end{aligned} \tag{8.223}$$

For the second term of Eq. (8.213) we obtain after some algebra

$$\begin{aligned} -\left(v^B\frac{\partial}{\partial q^B}\bar{K}_{1a}\right)\Bigg|_{v^i} &= -v^B\Bigg|_{v^i}\frac{\partial}{\partial q^B}K_{1a} + \\ & v^B\frac{\partial v^i}{\partial q^B}\left(\frac{\partial}{\partial v^i}\bar{K}_{1a}\right)\Bigg|_{v^i} = \\ -\frac{\partial H}{\partial p_A}\frac{\partial}{\partial q^A}(R_{1a}{}^{\alpha}\{\Phi_\alpha, H\}) &- v^B|_{v^i}\frac{\partial v^i}{\partial q^B}M_{iA}R_{0a}{}^{A}, \end{aligned} \tag{8.224}$$

where Eqs. (8.215) and (8.212) were used. Combining the Eqs. (8.223) and (8.224), we find the Hamiltonian form of Eq. (8.213)

$$R_{0a}{}^{\alpha}\{\Phi_\alpha, H\} - \left\{R_{1a}{}^{\alpha}\{\Phi_\alpha, H\}, H\right\} \equiv 0. \tag{8.225}$$

Bringing together all the results, we arrive at the Hamiltonian form of the identities, Eqs. (8.198), (8.199) and (8.200).

Local symmetry of the first-order action. Invariance of the first-order action (2.155) under the transformations (8.201) can be demonstrated by direct computation. Variation under δq, δv given in Eq. (8.201), and under some δp reads (disregarding the total derivatives)

$$\begin{aligned}\delta S_v = &\int d\tau\, \epsilon^a \bar{K}_{0a} + \dot{\epsilon}^a \bar{K}_{1a} - \dot{v}^A \bar{M}_{AB}(\epsilon^a \bar{R}_{0a}{}^B + \dot{\epsilon}^a \bar{R}_{1a}{}^B) \\ &+ \left(\delta p_A - \frac{\partial^2 \bar{L}}{\partial q^A \partial v^B}\delta_\epsilon q^B\right)(\dot{q}^A - v^A) \end{aligned}\tag{8.226}$$

$$\begin{aligned} = &\int d\tau\, \dot{\epsilon}^a v^A \bar{M}_{AB} \bar{R}_{1a}{}^B + \epsilon^a \left(\bar{K}_{0a} - v^A \frac{\partial}{\partial q^B}\bar{K}_{1a}\right) \\ &- \epsilon^a \dot{v}^A \left(\bar{M}_{AB}\bar{R}_{0a}{}^B + \frac{\partial}{\partial v^A}\bar{K}_{1a}\right) \\ &+ \left(\delta p_A - \frac{\partial^2 \bar{L}}{\partial q^A \partial v^B}\delta_\epsilon q^B - \epsilon^a \frac{\partial}{\partial q^A}\bar{K}_{1a}\right)(\dot{q}^A - v^A), \end{aligned}\tag{8.227}$$

where we have carried out integration by parts in the second term of Eq. (8.226). The first and the second lines in Eq. (8.227) vanish due to Eqs. (8.211), (8.212) and (8.213). Then the variation δS_v will be a total derivative, if we choose δp_A according to Eq. (8.201).

Local symmetry of the Hamiltonian action. One may expect that the transformations (8.201) with the velocities $v^i(q^A, p_j, v^\alpha)$ substituted will be a symmetry of the Hamiltonian action $p\dot{q} - H$. Let us find their manifest form. Using Eqs. (8.198), we obtain for the variation $\delta_\epsilon q^i|_{v^i}$

$$\delta q^i|_{v^i} = \left(\epsilon^a R_{0a}{}^\beta + \dot{\epsilon}^a R_{1a}{}^\beta\right)\{q^i, \Phi_\beta\} - \epsilon^a \left\{q^i, R_{1a}{}^\beta\{\Phi_\beta, H\}\right\}. \tag{8.228}$$

The variation $\delta_\epsilon q^\alpha|_{v^i}$ can be identically rewritten in a similar form

$$\begin{aligned}\delta q^\alpha|_{v^i} &= \epsilon^a R_{0a}{}^\alpha + \dot{\epsilon}^a R_{1a}{}^\alpha \equiv \\ \left(\epsilon^a R_{0a}{}^\beta + \dot{\epsilon}^a R_{1a}{}^\beta\right)&\{q^\alpha, \Phi_\beta\} - \epsilon^a\{q^\alpha, R_{1a}{}^\beta\{\Phi_\beta, H\}\},\end{aligned}\tag{8.229}$$

since $\{q^\alpha, \Phi_\beta\} = \delta^\alpha_\beta$ and since the quantity $R_{1a}{}^\beta\{\Phi_\beta, H\}$ does not depend on p_α. For the variation $\delta p_A|_{v^i}$ we have

$$
\begin{aligned}
\delta p_A|_{v^i} = \left(-\frac{\partial^2 \bar{L}}{\partial q^A \partial v^B}\delta_\epsilon q^B + \epsilon^a \frac{\partial}{\partial q^A}\bar{K}_{1a}\right)\Bigg|_{v^i} = \\
-\left(\frac{\partial}{\partial q^A}\left(\frac{\partial \bar{L}}{\partial v^B}\Bigg|_{v^i}\right) - \frac{\partial^2 \tilde{L}}{\partial v^i \partial v^B}\Bigg|_{v^i}\frac{\partial v^i}{\partial q^A}\right)\delta_\epsilon q^B|_{v^i} + \\
\epsilon^a \frac{\partial K_{1a}}{\partial q^A} - \epsilon^a \frac{\partial v^i}{\partial q^A}\frac{\partial \bar{K}_{1a}}{\partial v^i}\Bigg|_{v^i} = \\
-\frac{\partial \Phi_\alpha}{\partial q^A}\delta_\epsilon q^\alpha - M_{Bi}|_{v^i}\frac{\partial v^i}{\partial q^A}\left(\epsilon^a R_{0a}{}^B + \dot{\epsilon}^a R_{1a}{}^B\right) + \\
\epsilon^a \frac{\partial}{\partial q^A}\left(R_{1a}{}^\alpha\{\Phi_\alpha, H\}\right) + \epsilon^a M_{Bi}\frac{\partial v^i}{\partial q^A}R_{0a}{}^B = \\
\left(\epsilon^a R_{0a}{}^\alpha + \dot{\epsilon}^a R_{1a}{}^\alpha\right)\{p_A, \Phi_\alpha\} - \epsilon^a\left\{p_A, R_{1a}{}^\alpha\{\Phi_\alpha, H\}\right\}.
\end{aligned}
\tag{8.230}
$$

where Eqs. (8.215), (8.211) and (8.212) were used.

The Hamiltonian action is invariant under these transformations, as a consequence of the identities (8.199) and (8.200). Disregarding total derivatives, the variation of $p\dot{q}$ can be expressed as follows

$$
\delta(p_A\dot{q}^A) = \Phi_\alpha(\delta q^\alpha)^{\cdot} - \dot{\epsilon}^a R_{1a}{}^\alpha\{\Phi_\alpha, H\} - \epsilon^a \dot{v}^\beta \frac{\partial}{\partial v^\beta}\left(R_{1a}{}^\alpha\{\Phi_\alpha, H\}\right),
$$

while for the variation of H we have

$$
\begin{aligned}
-\delta H = -\Phi_\alpha(\delta q^\alpha)^{\cdot} + \dot{\epsilon}^a R_{1a}{}^\alpha\{\Phi_\alpha, H\} \\
+\epsilon^a\left(R_{0a}{}^\alpha\{\Phi_\alpha, H\} - \{R_{1a}{}^\alpha\{\Phi_\alpha, H\}, H\}\right).
\end{aligned}
$$

combining these terms and using Eqs. (8.199) and (8.200) we have $\delta_\epsilon S_H =$div.

To find the final form of the symmetry, we identically rewrite the transformations obtained in the form

$$
\begin{aligned}
\delta q^A = \{q^A, G\} - \{q^A, \delta_\epsilon q^\alpha\}\Phi_\alpha, \quad \delta p_A = \{p_A, G\} - \{p_A, \delta_\epsilon q^\alpha\}\Phi_\alpha, \\
\delta v^\alpha = \{H, \delta q^\alpha\} - \left(-(\delta q^\alpha)^{\cdot} + \{H, \delta q^\alpha\}\right),
\end{aligned}
$$

where G is given in (8.203). We note that the transformations

$$
\begin{aligned}
\bar{\delta} q^A = \{q^A, \delta q^\alpha\}\Phi_\alpha, \quad \bar{\delta} p_A = \{p_A, \delta q^\alpha\}\Phi_\alpha, \\
\bar{\delta} v^\alpha = -(\delta q^\alpha)^{\cdot} + \{H, \delta q^\alpha\}.
\end{aligned}
$$

represent a trivial symmetry of the Hamiltonian action, and thus can be omitted. The remaining part is precisely Eq. (8.202).

8.9 Local Symmetry Does Not Imply a Conserved Charge

Since in the expression for local symmetry (8.194) the parameters ϵ are arbitrary functions, we can take ϵ =const. So, local symmetry implies global symmetry. Let us construct the corresponding Noether charge. When ϵ =const, the invariance condition (8.195), instead of Eqs. (8.204), (8.205) and (8.206) implies only one of them, Eq. (8.206). The latter can be identically rewritten as follows

$$\left(\frac{\partial L}{\partial \dot{q}^A} R_{0a}{}^A - \omega_{0a}\right)^{\cdot} = \frac{\delta S}{\delta q^A} R_{0a}{}^A, \tag{8.231}$$

So the Noether charge is

$$Q_a = \frac{\partial L}{\partial \dot{q}^A} R_{0a}{}^A - \omega_{0a}, \tag{8.232}$$

At the same time, in our theory there is the identity (8.207). Using this in the previous expression, we obtain

$$Q_a = -\frac{\delta S}{\delta q^A} R_{1a}{}^A. \tag{8.233}$$

Hence a Noether charge of a local symmetry vanishes on equations of motion and thus cannot be used to characterize physical states. The same is true for gauge field theories.

8.10 Formalism of Extended Lagrangian

To continue the analysis of local symmetries in a singular theory, we associate with the initial Lagrangian (8.27) the so-called extended Lagrangian [46, 47]. This is formulated on the extended configuration space (q^A, s^a), where s^a stand for auxiliary variables associated with all the higher-stage constraints Φ_a. One of the advantages of the extended formalism is that the Dirac procedure, being applied to the extended Lagrangian, always stops at the third stage. Hamiltonian equations of the extended formulation have a more symmetric form, which essentially simplifies the analysis of their structure. Here we formulate the extended Lagrangian formulation and demonstrate its equivalence with the initial one. Local symmetries will be discussed in the next two sections.

Construction of the extended Lagrangian. Let

$$\omega_i(q^A, \dot{q}^A, s^a), \tag{8.234}$$

be a solution to the following equation[18]

$$\dot{q}^i - v^i(q^A, \omega_j, \dot{q}^\alpha) - s^a \frac{\partial \Phi_a(q^A, \omega_j)}{\partial \omega_i} = 0. \tag{8.235}$$

Here the functions $v^i(q^A, \omega_j, \dot{q}^\alpha)$, $\Phi_a(q^A, \omega_j)$ are taken from the initial formulation, see (8.30) and (8.45). The *extended Lagrangian* $\tilde{L}(q^A, \dot{q}^A, s^a)$ for $L(q^A, \dot{q}^A)$ is defined by

$$\begin{aligned}\tilde{L}(q^A, \dot{q}^A, s^a) =& L\left(q^A, D_\tau q^i, \dot{q}^\alpha\right) \\ &+ s^a \left[\omega_i \frac{\partial \Phi_a(q^a, \omega_i)}{\partial \omega_i} - \Phi_a(q^A, \omega_i)\right],\end{aligned} \tag{8.236}$$

where $D_\tau q^i$ is a quantity similar to the covariant derivative

$$\partial_\tau q^i \longrightarrow D_\tau q^i = \partial_\tau q^i - s^a \frac{\partial \Phi_a(q^a, \omega_i)}{\partial \omega_i}. \tag{8.237}$$

The second line in (8.236) disappears when the higher-stage constraints are homogeneous on momenta. For example, for the constraints of the form $\Phi_a = p_a$, where p_a is a part of the momenta $p_i = (p_a, p'_i)$, the extended action acquires the form

$$\tilde{L} = L(q^A, \dot{q}^a - s^a, \dot{q}'^i, \dot{q}^\alpha). \tag{8.238}$$

For the case $\Phi_a = h_a{}^i(q) p_i$ the extended Lagrangian is

$$\tilde{L} = L(q^A, \dot{q}^i - s^a h_a{}^i, \dot{q}^\alpha). \tag{8.239}$$

Let us discuss some properties of the extended Lagrangian.

First, we confirm that Eq. (8.235) can be resolved with respect to ω in the vicinity of the point $s^a = 0$. Indeed, when $s^a = 0$, this equation coincides with Eq. (8.30) of the initial formulation, which can be resolved, see (8.29). Hence $\det \frac{\partial (Eq.(8.235))^i}{\partial \omega_j} \neq 0$ at the point $s^a = 0$. Then the same is true in the vicinity of this point, and Eq. (8.235) can thus be resolved.

Second, by construction, the following properties hold:

$$\omega_i(q^A, \dot{q}^A, s^a)\Big|_{s^a=0} = \frac{\partial L}{\partial \dot{q}^i}, \tag{8.240}$$

[18] As will be shown below, Eq. (8.234) represents a solution to the equation $\tilde{p}_j = \frac{\partial \tilde{L}}{\partial \dot{q}^j}$ defining the conjugate momenta $\tilde{p}_j$ of the extended formulation.

$$\tilde{L}(q^A, \dot{q}^A, s^a)|_{s^a=0} = L(q^A, \dot{q}^A), \tag{8.241}$$

$$\left.\frac{\partial \tilde{L}}{\partial \omega_i}\right|_{\omega(q,\dot{q},s)} = 0, \tag{8.242}$$

$$\frac{\partial \tilde{L}}{\partial \dot{q}^i} = \omega_i(q^A, \dot{q}^A, s^a), \tag{8.243}$$

$$\frac{\partial \tilde{L}}{\partial \dot{q}^\alpha} = \left.\frac{\partial L(q^A, v^i, \dot{q}^\alpha)}{\partial \dot{q}^\alpha}\right|_{v^i(q,\omega,\dot{q}^\alpha)} = f_\alpha(q^A, \omega_j(q, \dot{q}, s)). \tag{8.244}$$

In Eq. (8.242), $\tilde{L}$ is considered as a function of ω. This formula greatly simplifies computations in the extended formalism.

Exercise
Prove these properties.

Using Eq. (8.235), the extended Lagrangian can be rewritten in the equivalent form

$$\begin{aligned}\tilde{L}(q^A, \dot{q}^A, s^a) = L(q^A, v^i(q^A, \omega_j, \dot{q}^\alpha), \dot{q}^\alpha) + \\ \omega_i(\dot{q}^i - v^i(q^A, \omega_j, \dot{q}^\alpha)) - s^a \Phi_a(q^A, \omega_j),\end{aligned} \tag{8.245}$$

where the functions v^i, ω_i are specified by Eqs. (8.30) and (8.235).

Hamiltonian formulation for the extended Lagrangian. According to Eqs. (8.243) and (8.244), the conjugate momenta $\tilde{p}_A, \pi_a$ for q^A, s^a are

$$\tilde{p}_i = \frac{\partial \tilde{L}}{\partial \dot{q}^i} = \omega_i(q^A, \dot{q}^A, s^a), \tag{8.246}$$

$$\tilde{p}_\alpha = \frac{\partial \tilde{L}}{\partial \dot{q}^\alpha} = f_\alpha(q^A, \omega_j),$$

$$\pi_a = \frac{\partial \tilde{L}}{\partial \dot{s}^a} = 0. \tag{8.247}$$

That is, $\tilde{p}_i$ is precisely the solution to our basic Eq. (8.235). Taking this into account, the system (8.246) and (8.247) is equivalent to the following one

$$\dot{q}^i = v^i(q^A, \tilde{p}_j, \dot{q}^\alpha) + s^a \frac{\partial \Phi_a(q^A, \tilde{p}_j)}{\partial \tilde{p}_i}, \tag{8.248}$$

$$\tilde{p}_\alpha - f_\alpha(q^A, \tilde{p}_j) = 0, \tag{8.249}$$

$$\pi_a = 0. \tag{8.250}$$

So, in the extended formulation the primary constraints (8.31) of the initial formulation are present. Besides, there are the trivial constraints (8.250) in a number equal to the number of all the higher-stage constraints of the initial formulation.

Using the definition (8.34), we obtain the Hamiltonian

$$\tilde{H}_0 = H_0 + s^a \Phi_a, \tag{8.251}$$

where H_0 is precisely the Hamiltonian of the initial formulation. Then the complete Hamiltonian for $\tilde{L}$ reads

$$\tilde{H} = H_0(q^A, \tilde{p}_j) + s^a \Phi_a(q^A, \tilde{p}_j) + v^\alpha \Phi_\alpha(q^A, \tilde{p}_B) + v^a \pi_a. \tag{8.252}$$

The complete Hamiltonian for $\tilde{L}$ sometimes is called the *extended Hamiltonian* for L.

Due to the very special structure of the Hamiltonian, Eq. (8.251), preservation in time of the primary constraints π_a, $\dot{\pi}_a = \{\pi_a, H_0 + s^a \Phi_a\} = -\Phi_a = 0$ implies the equations $\Phi_a = 0$. Hence *all the higher-stage constraints of the initial formulation are second-stage constraints of the extended theory*.

Preservation in time of the primary constraints Φ_α leads to the equations $\{\Phi_\alpha, \tilde{H}\} = \{\Phi_\alpha, H_0\} + \{\Phi_\alpha, \Phi_\beta\} v^\beta + \{\Phi_\alpha, \Phi_b\} s^b = 0$. In turn, preservation of the secondary constraints Φ_a leads to the equations $\{\Phi_a, \tilde{H}\} = \{\Phi_a, H_0\} + \{\Phi_a, \Phi_\beta\} v^\beta + \{\Phi_a, \Phi_b\} s^b = 0$. To continue the analysis, it is convenient to unify them as follows:

$$\{\Phi_I, H_0\} + \{\Phi_I, \Phi_J\} S^J = 0. \tag{8.253}$$

Here Φ_I are all the constraints of the initial formulation, and $S^J \equiv (v^\alpha, s^a)$. Using the matrix (8.91), the system (8.253) can be rewritten in the equivalent form

$$\{\tilde{\Phi}_{I_1}, H_0\} + O(\Phi_I) = 0, \tag{8.254}$$

$$\{\tilde{\Phi}_{I_2}, H_0\} + \{\tilde{\Phi}_{I_2}, \Phi_J\} S^J = O(\Phi_I). \tag{8.255}$$

Equation (8.254) does not contain any new information, since the first class constraints commute with the Hamiltonian, see Eq. (8.94). Let us analyze the system (8.255). First, note that due to the rank condition $\operatorname{rank}\{\tilde{\Phi}_{I_2}, \Phi_J\}\Big|_{\Phi_I} = [I_2] = \max$, exactly $[I_2]$ variables among S^I can be determined from the system. According to the Dirac prescription, we need to determine the maximal number of the multipliers v^α. To do this, let us restore v-dependence in Eq. (8.255): $\{\tilde{\Phi}_{I_2}, \Phi_\alpha\} v^\alpha + \{\tilde{\Phi}_{I_2}, H_0\} + \{\tilde{\Phi}_{I_2}, \Phi_b\} s^b = 0$. Since the matrix $\{\tilde{\Phi}_{I_2}, \Phi_\alpha\}$ is the same as in the initial formulation, from these equations we determine a group of variables v^{α_2} through the remaining variables v^{α_1}, where $[\alpha_2]$ is the number of second-class constraints among Φ_α. After substitution of the result into the remaining equations of the system (8.255), this acquires the form

$$v^{\alpha_2} = v^{\alpha_2}(q, \tilde{p}, s^a, v^{\alpha_1}), \qquad Q_{a_2 b}(q, \tilde{p}) s^b + P_{a_2}(q, \tilde{p}) = 0, \tag{8.256}$$

where $[a_2]$ is the number of higher-stage second-class constraints of the initial theory. It must be $P \approx 0$, since when $s^b = 0$, the system (8.255) is a subsystem of $\{\Phi_I, H\} = 0$, but the latter vanishes after substitution of the multipliers determined during the procedure; see the discussion after Eq. (8.46). Besides, note that rank $Q = [a_2] = \max$. Indeed, suppose that rank $Q = [a'] < [a_2]$. Then from Eq. (8.255) only $[\alpha_2] + [a'] < [I_2]$ variables among S^I can be determined, contradicting the conclusion reached before. In short, the system (8.253) for determining the second-stage and third-stage constraints and multipliers is equivalent to

$$v^{\alpha_2} = v^{\alpha_2}(q, \tilde{p}, s^{a_1}, v^{\alpha_1}), \tag{8.257}$$

$$s^{a_2} = \tilde{Q}^{a_2}{}_{b_1}(q, \tilde{p}) s^{b_1}, \tag{8.258}$$

Conservation in time of the constraints (8.258) leads to the equations for determining the multipliers

$$v^{a_2} = \{Q^{a_2}{}_{b_1}(q, \tilde{p}) s^{b_1}, \tilde{H}\}. \tag{8.259}$$

Since there are no new constraints, the Dirac procedure for $\tilde{L}$ stops at this stage. All the constraints of the theory have been revealed after completing the third stage.

Equivalence between the initial and the extended Lagrangian formulations. Now we are ready to compare theories $\tilde{L}$ and L. The dynamics of the theory $\tilde{L}$ is governed by the Hamiltonian equations

$$\begin{aligned} \dot{q}^A &= \{q^A, H\} + s^a\{q^A, \Phi_a\}, \qquad & \dot{\tilde{p}}_A &= \{\tilde{p}_A, H\} + s^a\{\tilde{p}_A, \Phi_a\}, \\ \dot{s}^a &= v^a, & \dot{\pi}_a &= 0, \end{aligned} \tag{8.260}$$

as well as by the constraints

$$\Phi_\alpha = 0, \qquad \Phi_a = 0, \tag{8.261}$$

$$\pi_{a_1} = 0, \tag{8.262}$$

$$\pi_{a_2} = 0, \qquad s^{a_2} = Q^{a_2}{}_{b_1}(q, \tilde{p}) s^{b_1}. \tag{8.263}$$

Here H is the complete Hamiltonian of the initial theory (8.33), and the Poisson bracket is defined on the phase space $q^A, s^a, \tilde{p}_A, \pi_a$. Note that each solution of the extended theory with $s^a = 0$ represents a solution of the initial theory as well.

The constraints $\pi_{a_1} = 0$ can be replaced by the combinations $\pi_{a_1} + \pi_{a_2} Q^{a_2}{}_{a_1}(q, \tilde{p}) = 0$, which represent a first-class subset. Let us make partial fixation of a gauge by imposing the equations $s^{a_1} = 0$ as gauge conditions for the subset. Then the (s^a, π_a)-sector of the theory disappears, whereas the Eqs. (8.260) and (8.261) coincide exactly with those of the initial theory L. Let us recall that $\tilde{L}$ has been constructed in the vicinity of the point $s^a = 0$. The gauge $s^{a_1} = 0$ implies

$s^a = 0$ due to the homogeneity of Eq. (8.258). This guarantees the self-consistency of the construction. Thus L represents a gauge for $\tilde{L}$, which proves the equivalence of the two formulations.

Comment Note that the variables π_b, s^a, belong to the (Ω, Q)–sector of special coordinates specified in Sect. 8.4. According to the ideology discussed there, we can omit the variables π_a, v^a, obtaining an equivalent theory. It reads

$$\dot{q}^A = \{q^A, H + s^a\Phi_a\}, \qquad \dot{\tilde{p}}_A = \{\tilde{p}_A, H + s^a\Phi_a\},$$
$$\Phi_\alpha = 0, \qquad \Phi_a = 0. \tag{8.264}$$

As we have seen above, equations for s^a are consequences of this system. So we have omitted them. Comparing this with the initial formulation (8.35) and (8.36), we can formulate the result of the present section as follows: *in a singular theory, all the higher-stage constraints (multiplied by their own "Lagrangian multipliers" s^a) can be added to the complete Hamiltonian H.*

8.11 Local Symmetries of the Extended Lagrangian: Dirac Conjecture

Since the initial Lagrangian is a gauge for the extended one, it is a matter of convenience which of them is used to describe a physical system of interest. Here we discuss one of advantages of the extended Lagrangian action: there is a closed formula for its local symmetries in terms of constraints.

According to the analysis carried out in the previous section, the primary constraints of the extended formulation are $\Phi_\alpha = 0$, $\pi_a = 0$. Among $\Phi_\alpha = 0$ first-class constraints are present, in a number equal to the number of primary first-class constraints of L. Among $\pi_a = 0$, we have found the first-class constraints $\pi_{a_1} - \pi_{a_2} Q^{a_2}{}_{a_1}(q, p) = 0$, in a number equal to the number of all the higher-stage first-class constraints of L. Thus the number of primary first-class constraints of $\tilde{L}$ coincides with the number $[I_1]$ of all the first-class constraints of L. We obtain now exact formula for $[I_1]$ local symmetries of the extended formulation $\tilde{L}$.

The symmetries are given by

$$\delta_{I_1} q^A = \epsilon^{I_1} \left.\{q^A, \tilde{\Phi}_{I_1}(q^A, \tilde{p}_B)\}\right|_{\tilde{p}_i \to \frac{\partial \tilde{L}}{\partial \dot{q}^i}}, \tag{8.265}$$

$$\delta_{I_1} s^a = \left.\left[\dot{\epsilon}^{I_1} K_{I_1}{}^a + \epsilon^{I_1}\left(b_{I_1}{}^a + s^b c_{I_1 b}{}^a + \dot{q}^\beta c_{I_1 \beta}{}^a\right)\right]\right|_{\tilde{p}_i \to \frac{\partial \tilde{L}}{\partial \dot{q}^i}}. \tag{8.266}$$

Here $\epsilon^{I_1}(\tau)$, $I_1 = 1, 2, \ldots, [I_1]$, are the local parameters, and K is the conversion matrix, see Eq. (8.91). Note that Eq. (8.265) represents an infinitesimal canonical transformation, with the generators being the first-class constraints of initial formulation. We point out that these formulas represent a direct generalization of our illustrative example, see Eq. (8.192).

According to Eq. (8.266) the variation of some s^a involves the derivative of parameters. Hence they can be identified with gauge fields for the symmetry. At this point, it is instructive to discuss what happens with local symmetries when we pass from L to $\tilde{L}$. As we have seen in Sect. 8.8, $\overset{(N-1)}{\epsilon}$-type symmetry implies N-th stage constraints in the Hamiltonian formulation for L. Replacing L with $\tilde{L}$, we arrive at the formulation with at most second-stage first-class constraints and the corresponding $\dot{\epsilon}$-type symmetries (8.265). That is each symmetry (8.193) of L "decomposes" into N gauge symmetries of $\tilde{L}$.

We now show that the variation of $\tilde{L}$ under the transformation (8.265) is proportional to the higher-stage constraints T_a. So, it can be canceled by appropriate variation of s^a, which is given by Eq. (8.266). In the subsequent computations we omit all the total derivatives. Besides, the notation $A|$ implies the substitution indicated in Eqs. (8.265) and (8.266).

To give a proof, it is convenient to represent the extended Lagrangian (8.236) in terms of the initial Hamiltonian H_0, instead of the initial Lagrangian L. Using Eq. (8.68) we write

$$\begin{aligned}\tilde{L}(q^A, \dot{q}^A, s^a) = {} & \omega_i \dot{q}^i + f_\alpha(q^A, \omega_j)\dot{q}^\alpha \\ & - H_0(q^A, \omega_j) - s^a T_a(q^A, \omega_j),\end{aligned} \tag{8.267}$$

where the functions $\omega_i(q, \dot{q}, s)$, $f_\alpha(q, \omega)$ are defined by Eqs. (8.234), (8.32). According to the identity (8.242), the variation of $\tilde{L}$ with respect to ω_i does not give any contribution. Taking this into account, the variation of Eq. (8.267) under the transformation (8.265) can be written in the form

$$\begin{aligned}\delta\tilde{L} = {} & -\dot{\omega}_i(q, \dot{q}, s)\left.\frac{\partial\tilde{\Phi}_{I_1}}{\partial\tilde{p}_i}\right|\epsilon^{I_1} - \dot{f}_\alpha(q, \omega(q, \dot{q}, s)\left.\frac{\partial\tilde{\Phi}_{I_1}}{\partial\tilde{p}_\alpha}\right|\epsilon^{I_1} \\ & -\left.\left(\frac{\partial H_0(q^A, \tilde{p}_j)}{\partial q^A} + \dot{q}^\alpha\frac{\partial\Phi_\alpha(q^A, \tilde{p}_B)}{\partial q^A} + s^a\frac{\partial\Phi_a(q^A, \tilde{p}_j)}{\partial q^A}\right)\right| \left.\{q^A, \tilde{G}_{I_1}\}\right|\epsilon^{I_1} \\ & -\delta_{I_1} s^a \Phi_a(q^A, \omega_j).\end{aligned}$$

To see that $\delta\tilde{L}$ is a total derivative, we add the following zero

$$\begin{aligned}0 \equiv {} & \left[\left.\frac{\partial\tilde{L}}{\partial\omega_i}\right|_{\omega_i}\{\tilde{p}_i, \tilde{\Phi}_{I_1}\}\right. \\ & \left.\left. - \left(\frac{\partial H_0}{\partial\tilde{p}_\beta} + \dot{q}^\alpha\frac{\partial\Phi_\alpha}{\partial\tilde{p}_\beta} + s^a\frac{\partial\Phi_a}{\partial\tilde{p}_\beta}\right)\{\tilde{p}_\beta, \tilde{\Phi}_{I_1}\} + \dot{q}^\alpha\{\tilde{p}_\alpha, \tilde{\Phi}_{I_1}\}\right]\right|\epsilon^{I_1},\end{aligned}$$

to the r.h.s. of the previous expression. It then reads

$$\delta\tilde{L} = \left[\dot{\epsilon}^{I_1}\tilde{\Phi}_{I_1} - \epsilon^{I_1}\left(\{H_0, \tilde{\Phi}_{I_1}\} + \dot{q}^{\alpha}\{\Phi_\alpha, \tilde{\Phi}_{I_1}\} + s^a\{\Phi_a, \tilde{\Phi}_{I_1}\}\right)\right]\Big|$$
$$-\delta_{I_1}s^a\Phi_a(q^A, \omega_j) =$$
$$\left[\dot{\epsilon}^{I_1}\tilde{\Phi}_{I_1} + \epsilon^{I_1}\left(b_{I_1}{}^I + \dot{q}^{\alpha}c_{I_1\alpha}{}^I + s^b c_{I_1 b}{}^I\right)\Phi_I\right]\Big| - \delta_{I_1}s^a\Phi_a(q^A, \omega_j),$$

where b, c are coefficient functions of the constraint algebra (8.92). Using the equalities $\Phi_I| = (0,\ \Phi_a(q^A, \omega_j))$, $\tilde{\Phi}_{I_1}\Big| = K_{I_1}{}^a\Phi_a(q^A, \omega_j)$, we finally obtain

$$\delta\tilde{L} = \left[\dot{\epsilon}^{I_1}K_{I_1}{}^a + \epsilon^{I_1}\left(b_{I_1}{}^a + \dot{q}^{\alpha}c_{I_1\alpha}{}^a + s^b c_{I_1 b}{}^a\right) - \delta_{I_1}s^a\right]\Big|_{\tilde{p}_i \to \frac{\partial\tilde{L}}{\partial\dot{q}^i}}\Phi_a.$$

Then the variation of s^a given in Eq. (8.266) implies $\delta\tilde{L}$ =div, as has been stated.

Example Consider a system with the configuration-space variables $x^\mu,\ e,\ g$ (where $\mu = 0, 1, 2, 3,\ \eta_{\mu\nu} = (-, +, +, +)$), and with the action being

$$S = \int d\tau\left(\frac{1}{2e}(\dot{x}^\mu - gx^\mu)^2 + \frac{g^2}{2e}\right), \qquad a = \text{const}. \tag{8.268}$$

This implies the complete Hamiltonian

$$H = \frac{1}{2}ep^2 + g(xp) - \frac{g^2}{2e} + v_e p_e + v_g p_g, \tag{8.269}$$

as well as the constraints

$$\Phi_1 \equiv p_e = 0, \qquad T_1 \equiv -\frac{1}{2}\left(p^2 + \frac{g^2}{e^2}\right) = 0; \tag{8.270}$$
$$\Phi_2 \equiv p_g = 0, \qquad T_2 \equiv \frac{g}{e} - (xp) = 0. \tag{8.271}$$

They can be reorganized with the aim of separating the first class constraints $\tilde{\Phi}_1$ and $\tilde{T}_1$

$$\tilde{\Phi}_1 \equiv p_e + \frac{g}{e}p_g = 0,$$
$$\tilde{T}_1 \equiv -\frac{1}{2}\left(p^2 - \frac{g^2}{e^2}\right) - \frac{g}{e}(xp) + \frac{g^2}{e}p_g = 0; \tag{8.272}$$
$$p_g = 0, \qquad \frac{g}{e} - (xp) = 0. \tag{8.273}$$

In this case, the solution to the basic Eq. (8.235) is given by

$$\omega^\mu = \frac{1}{e - s^2}(\dot{x}^\mu - (g - s^2)x^\mu). \tag{8.274}$$

Using the Eqs. (8.270), (8.271) and (8.274) we obtain the extended Lagrangian (8.236)

$$\tilde{L} = \frac{1}{2(e - s^1)}(\dot{x}^\mu - (g - s^2)x^\mu)^2 + \frac{g^2}{2e}\left(1 + \frac{s^1}{e}\right) - \frac{g}{e}s^2. \quad (8.275)$$

Two local symmetries of $\tilde{L}$ are obtained according to Eqs. (8.265) and (8.266), using the expression (8.272) for the first-class constraints. They read

$$\begin{aligned} &\delta_1 x^\mu = -\epsilon^1\left(\omega^\mu + \frac{g}{e}x^\mu\right), \qquad \delta_1 e = 0, \qquad \delta_1 g = \epsilon^1\frac{g^2}{e}, \\ &\delta_1 s^1 = \dot{\epsilon}^1 - 2\epsilon^1\left(\frac{gs^1}{e} - s^2\right), \qquad \delta_1 s^2 = \left(\epsilon^1\frac{g}{e}\right)^{\cdot} + \epsilon^1\frac{g^2}{e}; \end{aligned} \quad (8.276)$$

$$\begin{aligned} &\delta_2 x^\mu = 0, \qquad \delta_2 e = \epsilon^2, \qquad \delta_2 g = \epsilon^2\frac{g}{e}, \\ &\qquad\qquad \delta_2 s^1 = \epsilon^2, \qquad \delta_2 s^2 = \epsilon^2\frac{g}{e}. \end{aligned} \quad (8.277)$$

Invariance of $\tilde{L}$ under (8.277) can be easily verified. By tedious computations, the reader can confirm that it is invariant under (8.276) as well, $\delta_1\tilde{L} = -\frac{1}{2}(\epsilon^1(\omega^\mu)^2 + \epsilon^1\left(\frac{g}{e}\right)^2)^{\cdot}$.

Dirac conjecture. Consider a theory which involves only first-class constraints. This implies that in the total Hamiltonian, $H = H_0 + v^\alpha\Phi_\alpha$, all the velocities remain undetermined. According to Sect. 2.3 the solution to the Hamiltonian equations, in linear order with respect to $\delta\tau$, is

$$z(\delta\tau) = z(0) + \delta\tau\{z, H_0\} + \delta\tau v^\alpha\{z, \Phi_\alpha\}, \quad (8.278)$$

and depends on the arbitrary functions v^α. According to Sect. 8.4, solutions which correspond to different choices of v, $z_1(\delta\tau, v_1)$ and $z_2(\delta\tau, v_2)$, are equivalent, and describe the same physical state.

Dirac observed that, according to (8.278), the solutions z_1 and z_2 are related by canonical transformation with the generators being first-class constraints: $\delta z = \epsilon^\alpha\{z, \Phi_\alpha\}$, $\epsilon^\alpha = \delta\tau\triangle v^\alpha$. The Dirac conjecture is that the higher-state constraints also generate transformations that do not change physical states.

We point out that (8.265) can be considered as a proof of the Dirac conjecture formulated as follows: all first-class constraints of an initial Lagrangian are generators of local symmetry of the extended Lagrangian.

8.12 Local Symmetries of the Initial Lagrangian

When only first-class constraints are present in the formulation, symmetries of the extended Lagrangian can be used to restore those of the initial Lagrangian. In the absence of second-class constraints, Eqs. (8.265) and (8.266) acquire the form

$$\delta_I q^A = \epsilon^I \left.\{q^A, \Phi_I\}\right|_{\tilde{p}_i \to \frac{\partial \tilde{L}}{\partial \dot{q}^i}},$$
$$\delta_I s^a = \left.\left[\dot{\epsilon}^a \delta_{aI} + \epsilon^I \left(b_I{}^a + s^b c_{Ib}{}^a + \dot{q}^\beta c_{I\beta}{}^a\right)\right]\right|_{\tilde{p}_i \to \frac{\partial \tilde{L}}{\partial \dot{q}^i}}. \tag{8.279}$$

We note that the extended Lagrangian coincides with the original one for $s^a = 0$: $\tilde{L}(q, 0) = L(q)$, see Eq. (8.241). So the initial action will be invariant under any transformation

$$\delta q^A = \sum_{I_1} \left.\delta_I q^A\right|_{s=0}, \tag{8.280}$$

which obeys the system $\delta s^a|_{s=0} = 0$, that is

$$\dot{\epsilon}^I K_I{}^a + \epsilon^I \left(b_I{}^a + \dot{q}^\beta c_{I\beta}{}^a\right) = 0. \tag{8.281}$$

We have $[a]$ equations for $[\alpha] + [a]$ variables ϵ^I. In the work [48] it was demonstrated that these equations can be solved by pure algebraic methods, which give some $[a]$ of ϵ in terms of the remaining ϵ and their derivatives of order less than N. This allows us to find $[\alpha]$ local symmetries of L.

We present two examples of how it works.

Maxwell action. Consider the Maxwell action of an electromagnetic field

$$S = -\frac{1}{4} \int d^4x F_{\mu\nu} F^{\mu\nu} = \int d^4x \left[\frac{1}{2}(\partial_0 A_a - \partial_a A_0)^2 - \frac{1}{4}(F_{ab})^2\right]. \tag{8.282}$$

In this case, the functions v^i from Eq. (8.30) are given by $p_a + \partial_a A_0$. The action implies primary and secondary constraints

$$p_0 = 0, \qquad \partial_a p_a = 0. \tag{8.283}$$

Then the basic Eq. (8.235) acquires the form $\partial_0 A_a - \omega_a - \partial_a A_0 + \partial_a s = 0$, and the extended Lagrangian action is[19]

[19] In the transition from mechanics to a field theory, derivatives are replaced by variational derivatives. In particular, the last term in Eq. (8.235) reads $\frac{\delta}{\delta \omega_i(x)} \int d^3y s^a(x) T_a(q^A(y), \omega_i(y))$.

$$\tilde{S} = \int d^4x \left[\frac{1}{2}(\partial_0 A_a - \partial_a A_0 + \partial_a s)^2 - \frac{1}{4}(F_{ab})^2 \right]. \tag{8.284}$$

Its local symmetries can immediately be written according to Eqs. (8.279); the non-vanishing variations are

$$\begin{aligned} \delta_\beta A_0 &= \beta, \qquad & \delta_\beta s &= \beta, \\ \delta_\alpha A_b &= -\partial_b \alpha, \qquad & \delta_\alpha s &= \partial_0 \alpha. \end{aligned} \tag{8.285}$$

According to Eq. (8.280), the symmetry of the initial action appears as the following combination

$$\begin{aligned} (\delta_\beta + \delta_\alpha) A_b &= -\partial_b \alpha, \\ (\delta_\beta + \delta_\alpha) A_0 &= \beta, \end{aligned} \tag{8.286}$$

where the parameters obey the equation $\partial_0 \alpha + \beta = 0$. The substitution $\beta = -\partial_0 \alpha$ into Eq. (8.286) gives the standard form of $U(1)$ gauge symmetry

$$A'_\mu = A_\mu + \partial_\mu \alpha. \tag{8.287}$$

Example with fourth-stage constraints. Let us consider the Lagrangian

$$L = \frac{1}{2}(\dot{x})^2 + \xi (x)^2, \tag{8.288}$$

where $x^\mu(\tau), \xi(\tau)$ are configuration space variables, $\mu = 0, 1, \ldots, n$, $(x)^2 \equiv \eta_{\mu\nu} x^\mu x^\nu$, $\eta_{\mu\nu} = (-, +, \ldots, +)$.

Denoting the conjugate momenta for x^μ, ξ as p_μ, p_ξ, the complete Hamiltonian reads

$$H_0 = \frac{1}{2} p^2 - \xi (x)^2 + v_\xi p_\xi, \tag{8.289}$$

where v_ξ is the velocity for the primary constraint $p_\xi = 0$. The complete system of constraints is

$$\Phi_1 \equiv p_\xi = 0, \quad T_2 \equiv x^2 = 0, \quad T_3 \equiv xp = 0, \quad T_4 \equiv p^2 = 0. \tag{8.290}$$

In this case, the variable ξ plays the role of q^α, while x^μ play the role of q^i of the general formalism.

The constraints form the first-class system

$$\{\Phi_I, \Phi_J\} = c_{IJ}{}^K \Phi_K, \qquad \{\Phi_I, H_0\} = b_I{}^J \Phi_J, \tag{8.291}$$

with the non-vanishing coefficient functions being

$$c_{23}{}^2 = -c_{32}{}^2 = 2, \qquad c_{24}{}^3 = -c_{42}{}^3 = 4, \qquad c_{34}{}^4 = -c_{43}{}^4 = 2;$$
$$b_1{}^2 = 1, \qquad b_2{}^3 = 2, \qquad b_3{}^4 = 1, \qquad b_3{}^3 = 2\xi, \qquad b_4{}^3 = 4\xi.$$

In the present case, Eq. (8.235) acquires the form $\dot{x}^\mu - \omega^\mu - s^3 x^\mu - 2s^4\omega^\mu = 0$, so

$$\omega^\mu = \frac{1}{1+2s^4}(\dot{x}^\mu - s^3 x^\mu). \tag{8.292}$$

Then the extended Lagrangian (8.236) is given by

$$\tilde{L} = \frac{1}{2(1+2s^4)}(\dot{x}^\mu - s^3 x^\mu)^2 + (\xi - s^2)(x^\mu)^2. \tag{8.293}$$

Using the Eqs. (8.279) and the coefficient functions found before, four symmetries can immediately be written as follows

$$\begin{aligned} &\delta_1\xi = \epsilon^1, \qquad \delta_1 s^2 = \epsilon^1; \\ &\delta_2 s^2 = \dot{\epsilon}^2 + 2\epsilon^2 s^3, \qquad \delta_2 s^3 = 2\epsilon^2(1+2s^4); \\ &\delta_3 x^\mu = \epsilon^3 x^\mu, \ \delta_3 s^2 = 2\epsilon^3(\xi - s^2), \ \delta_3 s^3 = \dot{\epsilon}^3, \ \delta_3 s^4 = \epsilon^3(1+2s^4); \\ &\delta_4 x^\mu = 2\epsilon^4\frac{\dot{x}^\mu - s^3 x^\mu}{1+2s^4}, \qquad \delta_4 s^3 = 4\epsilon^4(\xi - s^2), \qquad \delta_4 s^4 = \dot{\epsilon}^4 - 2\epsilon^4 s^3. \end{aligned} \tag{8.294}$$

Since the initial Lagrangian L implies a unique chain of four first-class constraints, we expect that it has one local symmetry of the $\overset{(3)}{\epsilon}$-type. The symmetry can be found according to the defining Eqs. (8.281). In this case, they read

$$\begin{aligned} \epsilon^1 + \dot{\epsilon}^2 + 2\epsilon^3\xi &= 0, \\ 2\epsilon^2 + \dot{\epsilon}^3 + 4\epsilon^4\xi &= 0, \\ \epsilon^3 + \dot{\epsilon}^4 &= 0. \end{aligned} \tag{8.295}$$

This allows us to find $\epsilon^1, \epsilon^2, \epsilon^3$ in terms of $\epsilon^4 \equiv \epsilon$: $\epsilon^1 = -\frac{1}{2}\overset{(3)}{\epsilon} + 4\dot{\epsilon}\xi + 2\epsilon\dot{\xi}$, $\epsilon^2 = \frac{1}{2}\ddot{\epsilon} - 2\epsilon\xi$, $\epsilon^3 = -\dot{\epsilon}$. According to Eq. (8.280), the local symmetry of the initial Lagrangian (8.288) is given by

$$\delta x^\mu = -\dot{\epsilon}x^\mu + 2\epsilon\dot{x}^\mu, \qquad \delta\xi = -\frac{1}{2}\overset{(3)}{\epsilon} + 4\dot{\epsilon}\xi + 2\epsilon\dot{\xi}. \tag{8.296}$$

In the presence of second-class constraints, local symmetries of L can not generally be restored according to the trick (8.280) and (8.281). The reason is that the number of equations of the system (8.281) can be equal to or greater than the number of parameters ϵ^a.

The expression (8.293) for the extended Lagrangian suggests the following redefinition of variables: $1 + 2s^4 \equiv e$, $\xi - s^2 \equiv \xi_1$; then it can be written in the form

$$L(e, \xi_1) = \frac{1}{2e}(\dot{x}^\mu - s^3 x^\mu)^2 + \xi_1 (x^\mu)^2. \tag{8.297}$$

Let us write its symmetries. The symmetry (8.294) disappears, since $L(e, \xi_1)$ is constructed from gauge-invariant variables with respect to this symmetry. The remaining symmetries acquire the form

$$\delta_2 \xi_1 = -\dot{\epsilon}^2 - 2\epsilon^2 s^3, \qquad \delta_2 s^3 = 2e\epsilon^2; \tag{8.298}$$

$$\delta_3 x^\mu = \epsilon^3 x^\mu, \ \delta_3 \xi_1 = -2\epsilon^3 \xi_1, \ \delta_3 s^3 = \dot{\epsilon}^3, \ \delta_3 e = 2\epsilon^3 e; \tag{8.299}$$

$$\delta_4 x^\mu = \frac{2\epsilon^4}{e}(\dot{x}^\mu - s^3 x^\mu), \quad \delta_4 s^3 = 4\epsilon^4 \xi_1, \quad \delta_4 e = 2(\dot{\epsilon}^4 - 2\epsilon^4 s^3). \tag{8.300}$$

The δ_4-symmetry can be replaced by the combination $\delta_\epsilon \equiv \delta\left(\epsilon^4 = \frac{1}{2}\epsilon e\right) + \delta(\epsilon^3 = \epsilon s^3) + \delta(\epsilon^2 = -\epsilon \xi_1)$, which has a simpler form

$$\delta_\epsilon x^\mu = \epsilon \dot{x}^\mu, \quad \delta_\epsilon \xi_1 = (\epsilon \xi_1)^\cdot, \quad \delta_\epsilon s^3 = (\epsilon s^3)^\cdot, \quad \delta_\epsilon e = (\epsilon e)^\cdot, \tag{8.301}$$

and represents the reparametrization invariance. As the independent symmetries of $L(e, \xi_1)$, we can take either Eqs. (8.298), (8.299) and (8.300), or Eqs. (8.298), (8.299) and (8.301).

8.13 Conversion of Second-Class Constraints by Deformation of Lagrangian Local Symmetries

In this section we discuss invertible changes on the space of functions which have the following form $q(\tau) \to f(\tilde{q}(\tau), \dot{\tilde{q}}(\tau), \ldots)$. Owing to the invertibility (and under certain conditions that will be discussed below), changes of this kind lead to an equivalent Lagrangian. To understand their meaning, suppose that q enters into the initial Lagrangian without derivatives, which implies the primary constraint $p = 0$ in the Hamiltonian formulation. The transformed Lagrangian L' will contain derivatives of the new variable $\tilde{q}$, so the constraint generally does not appear in the formulation L'. As will be seen, in many cases a pair of second-class constraints of the initial Lagrangian is replaced by a first-class constraint in the transformed formulation. That is, the notion of first- and second-class constraints is not "invariant" under such a change. In this section we follow the work [49].

To illustrate how this works, we analyze the following dynamically trivial model defined on configuration space $x(\tau), \ y(\tau), \ z(\tau)$, with the Lagrangian action being

$$S = \int d\tau \left[\frac{1}{2}(\dot{x} - y)^2 + \frac{1}{2}z^2\right]. \tag{8.302}$$

This is invariant under the finite local symmetry with the parameter $\alpha(\tau)$

$$\delta x=\alpha, \qquad \delta y=\dot{\alpha}, \qquad \delta z=0. \tag{8.303}$$

So we have a formulation with $\dot{\alpha}$ -type symmetry. Moving on to Hamiltonian formalism, we obtain the following chains of constraints:

	Primary	Secondary	
First-class chain	$p_y=0,$	$p_x=0,$	(8.304)
Second-class chain	$p_z=0,$	$z=0.$	(8.305)

Consider the transformation $z=\tilde{z}+\dot{y}$. For the new variables $x,\ y,\ \tilde{z}$, the action acquires the form

$$S'=\int d\tau\left(\frac{1}{2}(\dot{x}-y)^2+\frac{1}{2}(\tilde{z}+\dot{y})^2\right), \tag{8.306}$$

and has $\ddot{\alpha}$ -type symmetry

$$\delta x=\alpha, \qquad \delta y=\dot{\alpha}, \qquad \delta\tilde{z}=-\ddot{\alpha}. \tag{8.307}$$

As we know, this implies the appearance of a constraint at the third stage of the Dirac procedure. In the new formulation, this replaces the second-class chain (8.305). Computing the constraints of the formulation (8.306), we find the following first-class chain

Primary	Second-stage	Third-stage	
$p_{\tilde{z}}=0,$	$p_y=0,$	$p_x=0.$	(8.308)

The reader can verify that the initial formulation is a gauge of the new one (it corresponds to the gauge $\tilde{z}=0$). Hence, using the change, two second-class constraints (8.305) have been replaced on the first-class constraint $\tilde{z}=0$. In the formulation S' only the first-class constraints are present. The procedure is called a *conversion of second-class constraints*. In the language of symmetries, the change raises the order of a symmetry, leading to deformation of the constraints structure.

Let us describe the conversion trick in further detail. Let $L(q^A,\dot{q}^A)$ be the Lagrangian of a theory with first- and second-class constraints. In the Lagrangian formulation, the first-class constraints manifest themselves in invariance of the action under some local symmetry transformations. Let

$$\delta q^A=\overset{(k)}{\epsilon} R^A(q,\dot{q})+\dots, \tag{8.309}$$

be an infinitesimal form of one of the symmetries. The dots stand for all terms with less then k-derivatives acting on a parameter. As we know, $\overset{(k)}{\epsilon}$ -type symmetry generally implies the appearance of some constraint at the $(k+1)$-stage of the Dirac procedure.

Let us divide coordinates q^A into two groups: $q^A = (q^i, q^\alpha)$. We change the parametrization of the configuration space: $q^A \longrightarrow \tilde{q}^A$ according to the transformation which involves derivatives of q^α

$$q^i = q^i(\tilde{q}^A, \dot{\tilde{q}}^\alpha), \qquad q^\alpha = q^\alpha(\tilde{q}^\beta). \tag{8.310}$$

We suppose that the transformation is invertible in the following sense

$$\det \frac{\partial q^i}{\partial \tilde{q}^j} \neq 0, \qquad \det \frac{\partial q^\alpha}{\partial \tilde{q}^\beta} \neq 0. \tag{8.311}$$

This implies that $\tilde{q}^A$ can be determined from (8.310): $\tilde{q}^i = \tilde{q}^i(q^A, \dot{q}^\alpha)$, $\tilde{q}^\alpha = \tilde{q}^\alpha(q^\beta)$. So, our theory can be equally analyzed in terms of the Lagrangian $L' \equiv L(q(\tilde{q}), \dot{q}(\tilde{q}))$. We further suppose that the transformation (8.310) has been chosen in such a way that $\tilde{L}$ does not involve higher derivatives, disregarding the total-derivative terms (we show below that this is possible in a singular theory)

$$L'(\tilde{q}, \dot{\tilde{q}}, \ddot{\tilde{q}}) = \tilde{L}''(\tilde{q}, \dot{\tilde{q}}) + \frac{dF(\tilde{q}, \dot{\tilde{q}})}{d\tau}. \tag{8.312}$$

Let us see what we can say about the structure of Hamiltonian constraints of our theory in the new parametrization $\tilde{L}$, as compared with L. We note that the local symmetry for the set $\tilde{q}$ is generally of $\overset{(k+1)}{\epsilon}$-type: $\delta\tilde{q}^i = \overset{(k+1)}{\epsilon} \frac{\partial \tilde{q}^i}{\partial \dot{\tilde{q}}^\alpha} \tilde{R}^\alpha(\tilde{q}^A, \dot{\tilde{q}}^A, \ddot{\tilde{q}}^\alpha) + \ldots$. Since the order of the symmetry has been raised by one unit, at the $(k+2)$-stage of the Dirac procedure an extra constraint appears. On the other hand, the physical sector of $\tilde{L}$ is the same as for L. If the order of other symmetries (if any) was not lowered, the only possibility[20] is that the extra $(k+2)$-stage constraint is first-class, and it replaces a pair of second-class constraints of the initial formulation. In short, an appropriate parametrization (8.310), (8.311) and (8.312) of the configuration space implies a deformation of local symmetries which, in turn, can result in the conversion of second-class constraints. Clearly, Eqs. (8.311) and (8.312) represent only necessary conditions for the conversion.

Note that we can consider more general transformations: $q^i = q^i(\tilde{q}^A, \dot{\tilde{q}}^\alpha, \ddot{\tilde{q}}^\alpha, \ldots, \overset{(s)}{\tilde{q}^\alpha})$, $q^\alpha = q^\alpha(\tilde{q}^\beta)$, which involve higher derivatives of $\tilde{q}^\alpha$. These generally increase the order of symmetry by s units, and $2s$ second-class constraints can be converted. For an example of this kind, see [49].

As the example discussed earlier shows, the condition (8.312) can be easily satisfied if some variable enters into the action without a derivative. In this respect,

[20] Here the condition (8.312) is important. A theory with higher derivatives, being equivalent to the initial one, has more degrees of freedom than the number of variables q^A, see Sect. 2.12. So the extra constraints would be responsible for ruling out these hidden degrees of freedom. Our condition (8.312) precludes the appearance of the hidden degrees of freedom.

let us point out that for a singular theory $L(q, \dot{q})$, there is an equivalent formulation, $L'(q', \dot{q}')$, with the desired property. Actually, starting from the singular L, we construct the Hamiltonian $H = H_0(q^A, p_j) + v^\alpha \Phi_\alpha$, where $\Phi_\alpha(q^A, p_B) = p_\alpha - f_\alpha(q^A, p_j)$ are primary constraints. As we know, the functions $H_0,\ f_\alpha$ do not depend on p_α. We further separate a phase-space pair which corresponds to some fixed α, for example $\alpha = 1$: $\alpha = (1, \alpha')$, $(q^A, p_A) = (q^1, p_1, z)$. According to Sect. 4.4.3, there is a canonical transformation $(q^1, p_1, z) \to (q'^1, p'_1, z')$, such that the Hamiltonian acquires the form $H' = H'_0(q'^1, z') + v^1 p'_1 + v^{\alpha'} \Phi_{\alpha'}(q'^1, z')$. We can restore the Lagrangian $L'(q', \dot{q}')$ which reproduces H' in Hamiltonian formalism. By construction, L' does not depend on $\dot{q}'^1$.

We finish this section with three examples of application of the conversion trick.

8.13.1 Conversion in a Theory with Hidden $SO(1, 4)$ Global Symmetry

In this example, the initial formulation implies a non-linear realization of a global symmetry, therefore is not convenient. The conversion reveals this hidden symmetry that is present in the theory. Besides, the extra gauge freedom of the converted version is used to find a parametrization which linearizes equations of motion.

Consider a theory on the configuration space $x^\mu,\ e,\ g$ (where $\mu = 0, 1, 2, 3$, $\eta_{\mu\nu} = (-, +, +, +)$), and with the action

$$S = \int d\tau \left(\frac{1}{2e}(\dot{x}^\mu - g x^\mu)^2 + \frac{1}{2e^2} g^2 - a g \right), \qquad a = \text{const.} \tag{8.313}$$

The model has a manifest $SO(1, 3)$ global symmetry. The only local symmetry is the reparametrization invariance, which represents $\dot{\alpha}$ -type symmetry

$$\delta\tau = 0, \quad \delta x^\mu = -\alpha \dot{x}^\mu, \quad \delta e = -(\alpha e)^{\cdot}, \quad \delta g = -(\alpha g)^{\cdot} \tag{8.314}$$

Moving on to Hamiltonian formalism we obtain the complete Hamiltonian (v_e, v_g stand for velocities associated with the primary constraints)

$$H = \frac{e}{2} p^2 + g(xp) - \frac{g^2}{2e^2} + a g + v_e p_e + v_g p_g, \tag{8.315}$$

as well as the constraints (the initial constraints have been reorganized with the aim of separating the first-class ones)

$$p_e + (xp + a) p_g = 0, \qquad p^2 + (xp + a)^2 + 2 e p^2 p_g = 0; \tag{8.316}$$

$$p_g = 0, \qquad g - e(xp + a) = 0. \tag{8.317}$$

Equation (8.316) represent first-class constraints. The equations of motion for the (e, x)-sector can be written as follow

$$\begin{aligned} \dot{e} &= v_e, & \dot{p}_e &= 0, \\ \dot{x}^\mu &= e(p^\mu + (xp + a)x^\mu), & \dot{p}^\mu &= -e(xp + a)p^\mu. \end{aligned} \tag{8.318}$$

In terms of variables

$$\mathcal{X}^\mu = \frac{ax^\mu}{xp + a}, \qquad \mathcal{P}^\mu = \frac{ap^\mu}{xp + a}, \tag{8.319}$$

they acquire a form similar to those of a free relativistic particle, namely

$$\dot{\mathcal{X}}^\mu = e\mathcal{P}^\mu, \qquad \dot{\mathcal{P}}^\mu = 0, \qquad \mathcal{P}^2 = -a^2. \tag{8.320}$$

The presence of the conserved charge $\dot{\mathcal{P}}^\mu = 0$ indicates a hidden global symmetry related with the homogeneity of the configuration space. As will be seen below, the conversion reveals the symmetry and allows us to find a manifestly-invariant formulation of the theory.

To convert a pair of second-class constraints (8.317) we need to raise the order of symmetry (8.314) by one unit. From Eq. (8.314) we note that this can be achieved by performing a shift of a variable on $\dot{e}$. Since the variable g enters into the action without a derivative, a shift of the type $g = \tilde{g} + \dot{e}$ does not lead to higher-derivative terms in the action and thus realizes the conversion. It is convenient to accompany the shift by an appropriate change of variables. Namely, let us make the invertible transformation $(x^\mu, e, g) \longrightarrow (\tilde{x}^A = (\tilde{x}^\mu, \tilde{x}^4),\ \tilde{g})$, where

$$\tilde{x}^\mu = e^{-\frac{1}{2}} x^\mu, \qquad \tilde{x}^4 = e^{-\frac{1}{2}}, \qquad \tilde{g} = g - \frac{\dot{e}}{2e}. \tag{8.321}$$

In terms of these variables the action (8.313) acquires the form

$$S' = \int d\tau \left(\frac{1}{2}(\dot{\tilde{x}}^A - \tilde{g}\tilde{x}^A)^2 - a\tilde{g} \right), \qquad \eta_{AB} = (-, +, +, +, +), \tag{8.322}$$

The resulting action has a manifest $SO(1, 4)$ global symmetry. The conserved current $\mathcal{P}^\mu$ then corresponds to the symmetry under rotations in $(\tilde{x}^\mu, \tilde{x}^4)$-planes. The local symmetry of the action (8.322) can be obtained from Eqs. (8.314) and (8.321), and is of $\ddot{\alpha}$-type

$$\delta\tau = 0, \qquad \delta\tilde{x}^A = \frac{1}{2}\dot{\alpha}\tilde{x}^A - \alpha\dot{\tilde{x}}^A, \qquad \delta\tilde{g} = \frac{1}{2}\ddot{\alpha} - \dot{\alpha}\tilde{g} - \alpha\dot{\tilde{g}}. \tag{8.323}$$

Moving on to Hamiltonian formulation we obtain the Hamiltonian

$$H = \frac{1}{2}\tilde{p}^2 + \tilde{g}\tilde{x}^A\tilde{p}_A + a\tilde{g} + v_{\tilde{g}}p_{\tilde{g}}, \tag{8.324}$$

and the first-class constraints

$$\tilde{p}_{\tilde{g}} = 0, \qquad \tilde{x}^A\tilde{p}_A + a = 0, \qquad \tilde{p}^A\tilde{p}_A = 0, \tag{8.325}$$

Thus S' represents the converted version of the action (8.313). Let us write equations of motion for the x^A-sector

$$\dot{\tilde{x}}^A = \tilde{p}^A + \tilde{g}\tilde{x}^A, \qquad \dot{\tilde{p}}^A = -\tilde{g}\tilde{p}^A. \tag{8.326}$$

In the gauge $\tilde{g} = \tilde{x}^\mu\tilde{p}_\mu + a,\ \ \tilde{p}_4 = \tilde{x}^\mu\tilde{p}_\mu + a$ for the theory (8.322) we reproduce the initial dynamics (8.318) (taken in the gauge $e = 1$). Going over to the gauge $\tilde{g} = 0,\ \ \tilde{p}_4 = a$, we obtain the free Eqs. (8.320). Hence the extra gauge freedom, resulting from the conversion of second-class constraints, can be used to search for the parametrization which implies the linear equations of motion.

8.13.2 Classical Mechanics Subject to Kinematic Constraints as a Gauge Theory

The conversion trick can be carried out in a theory which involves only second-class constraints, that is in a theory without local symmetries in the initial formulation. To begin with, we note that a given theory without local symmetry can be treated as a gauge theory on appropriately extended configuration space. For instance, a theory with the action $S(q^A)$ can be equally considered as a theory on the space q^A, a, where a is one more configuration-space variable, with local transformations defined by $q'^A = q^A$, $a' = a + \alpha$. Since a does not enter into the action, the latter is invariant under local transformations![21] This trivial gauge symmetry of the extended formulation can be further used for conversion of second-class constraints according to our procedure.[22]

Let us see how this works on an example of classical mechanics with kinematic constraints. In Sect. 8.6 we discussed this as a theory with the action

[21] This is a general situation: given a locally-invariant action, there are special coordinates such that the action does not depend on some of them [10].

[22] There are other possibilities for creating trivial local symmetries. For example, in a given Lagrangian action with one of the variables being q, let us make the substitution $q = ab$, where $a,\ b$ represent new configuration space variables. The resulting action is equivalent to the initial one, an auxiliary character of one of the new degrees of freedom is guaranteed by the trivial gauge symmetry: $a \to a' = \alpha a,\ b \to b' = \alpha^{-1}b$. Another simple possibility is to write $q = a + b$, which implies the symmetry $a \to a' = a + \alpha,\ b \to b' = b - \alpha$. The well-known example of this kind transformation is einbein formulation in gravity theory: $g_{\mu\nu} = e^a_\mu e^a_\nu$, which implies local Lorentz invariance.

$$S = \int d\tau \left[L_0(q, \dot{q}) + \lambda^i G_i(q) \right], \tag{8.327}$$

which implies $4[i]$ second-class constraints

$$p_{\lambda i} = 0, \quad G_i = 0, \quad f^a G_{ia} = 0, \quad \lambda^i - \tilde{\triangle}^{ij}\{F_j, H'\} = 0. \tag{8.328}$$

Now we present it as a locally-invariant theory which involves only first-class constraints.

Conversion can be carried out by making the following transformation in the action (8.327)

$$\lambda^i = \tilde{\lambda}^i + \ddot{e}^i, \tag{8.329}$$

where the auxiliary variable $e^i(\tau)$ has been introduced. The modified action

$$S' = \int d\tau \left[L_0(q, \dot{q}) - \dot{e}^i G_{ia}\dot{q}^a + \tilde{\lambda}^i G_i(q) \right], \tag{8.330}$$

does not contain higher-derivative terms and is invariant under the local transformations $\tilde{\lambda}^i \to \tilde{\lambda}'^i = \tilde{\lambda}^i + \ddot{\alpha}^i$, $e^i \to e'^i = e^i - \alpha^i$. Due to this $\ddot{\alpha}$-symmetry we expect the appearance of $3[i]$ first class constraints in the Hamiltonian formulation for the theory (8.330). To confirm this, let us write defining equations for conjugate momenta

$$p_a \equiv \frac{\partial L}{\partial \dot{q}^a} = \frac{\partial L_0}{\partial \dot{q}^a} - \dot{e}^i G_{ia}, \quad p_{ei} \equiv \frac{\partial L}{\partial \dot{e}^i} = -G_{ia}\dot{q}^a, \quad p_{\tilde{\lambda}i} = 0. \tag{8.331}$$

The last equation represents $[i]$ primary constraints. The remaining equations can be resolved with respect to the velocities $\dot{q}^a, \dot{e}^i$, since the corresponding block of the Hessian matrix is non-degenerate. It can easily be seen in special coordinates chosen as follows. The initial coordinates q^a can be reordered in such a way that the rank minor of the matrix $\frac{\partial G_i}{\partial q^a}$ is placed on the right: $q^a = (q^\alpha, q^i)$, $\det \frac{\partial G_i}{\partial q^j} \neq 0$. Now, let us make the invertible change of variables $q^a \to \tilde{q}^a$, where $\tilde{q}^\alpha = q^\alpha$, $\tilde{q}^i = G_i(q^a)$. In these variables our Lagrangian is

$$L' = L_0(\tilde{q}, \dot{\tilde{q}}) - \dot{e}^i \dot{\tilde{q}}^i + \tilde{\lambda}^i \tilde{q}^i. \tag{8.332}$$

From this expression we immediately find the determinant of the Hessian matrix: $\det \frac{\partial^2 \tilde{L}}{\partial^2(\tilde{q}, e)} = \det \frac{\partial^2 L_0}{\partial \dot{\tilde{q}}^\alpha \partial \dot{\tilde{q}}^\beta}$. It does not vanish since in classical mechanics the quadratic form $\frac{\partial^2 L_0}{\partial \dot{\tilde{q}}^a \partial \dot{\tilde{q}}^b}$ is positive defined.

Let us return to the analysis of the action (8.330). The complete Hamiltonian is

$$H = p_a\dot{q}^a + p_{ei}\dot{e}^i - L_0(q, \dot{q}) + \dot{e}^i G_{ia}q^a - \tilde{\lambda}^i G_i(q) + v^i_{\tilde{\lambda}} p_{\tilde{\lambda}i}, \tag{8.333}$$

where $\dot{q}^a$, $\dot{e}^i$ are solutions to Eqs. (8.331). As before, the second-stage constraints are $G_i(q) = 0$. Their conservation in time can be easily computed by using of Eq. (8.331): $\dot{G}_i = \{G_i, H\} = -p_{ei}$, which gives the third-stage constraints $p_{ei} = 0$. Then the complete constraint system is composed by $3[i]$ first class constraints

$$p_{\tilde{\lambda}i} = 0, \qquad G_i = 0, \qquad p_{ei} = 0. \tag{8.334}$$

The first-class constraints $p_{ei} = 0$ state that the variables e^i are pure gauge degrees of freedom, as was expected. They can be removed from the formulation if we choose the gauge $e^i = 0$. The remaining $2[i]$ first-class constraints in Eq. (8.334) replace $4[i]$ second class-constraints (8.328) of the initial formulation.

As a particular example, we consider a **particle on a 2-sphere of radius** c, with the action being

$$S = \int d^3x \left[\frac{1}{2}m\dot{\vec{x}}^2 + \lambda(\vec{x}^2 - c^2)\right]. \tag{8.335}$$

This implies the following chain of four second-class constraints

$$p_\lambda = 0, \qquad \vec{x}^2 - c^2 = 0, \qquad \vec{x}\vec{p} = 0, \qquad \vec{p}^2 + 2mc^2\lambda = 0. \tag{8.336}$$

Conversion is achieved by the transformation $\lambda = \tilde{\lambda} + \frac{1}{2}m\ddot{e}$, which generates the symmetry $\tilde{\lambda} \to \tilde{\lambda}' = \tilde{\lambda} + \frac{1}{2}m\ddot{\alpha}$, $e \to e' = e - \alpha$. The transformed action

$$S' = \int d^3x \left[\frac{1}{2}m\dot{\vec{x}}^2 - m\dot{e}\vec{x}\dot{\vec{x}} + \tilde{\lambda}(\vec{x}^2 - c^2)\right]. \tag{8.337}$$

implies first-class constraints only, namely

$$p_{\tilde{\lambda}} = 0, \qquad \vec{x}^2 - c^2 = 0, \qquad p_e = 0. \tag{8.338}$$

$O(N)$**-invariant non-linear sigma model**

$$S = \int d^Dx \left[\frac{1}{2}(\partial_\mu\phi^a)^2 + \lambda((\phi^a)^2 - 1)\right], \tag{8.339}$$

represents an example of field theory with a similar structure of second-class constraints. Hence the transformation $\lambda = \tilde{\lambda} + \partial_\mu\partial^\mu e$ gives the formulation with first class constraints only

$$S' = \int d^Dx \left[\frac{1}{2}(\partial_\mu\phi^a)^2 - 2\partial_\mu e\phi^a\partial^\mu\phi^a + \tilde{\lambda}((\phi^a)^2 - 1)\right]. \tag{8.340}$$

8.13.3 Conversion in Maxwell–Proca Lagrangian for Massive Vector Field

As one more example of the conversion in a theory with second-class constraints only, we consider the massive vector field $A^\mu(x^\nu)$ in Minkowski space. It is described by the following action:

$$S = \int d^4x \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} m^2 A^\mu A_\mu \right], \quad F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{8.341}$$

Moving on to Hamiltonian formulation we find the Hamiltonian

$$H = \int d^3x \left[\frac{1}{2} p_a^2 - p_a \partial_a A_0 + \frac{1}{4} F_{ab}^2 - \frac{1}{2} m^2 A^\mu A_\mu + v_0 p_0 \right], \tag{8.342}$$

as well as the primary and secondary constraints

$$p^0 = 0, \qquad \partial_a p_a - m^2 A_0 = 0. \tag{8.343}$$

The system is second-class, with the Poisson bracket algebra being

$$\{\partial_a p_a - m^2 A_0, p_0\} = -m^2 \delta^3(x - y). \tag{8.344}$$

Conservation in time of the secondary constraint determines the velocity $v_0 = -\partial_k A_k$. Equations of motion for the propagating modes are

$$\partial_0 A_a = -p_a + \partial_a A_0, \qquad \partial_0 p_a = -\partial_b F_{ba} - m^2 A_a, \tag{8.345}$$

while the modes $A_0,\ p^0$ are determined by the algebraic Eqs. (8.343). In a converted version these modes turn into the gauge degrees of freedom. In this case, a transformation which creates the desirable $\dot{\alpha}$ - symmetry consists of introducing the *Stuckelberg field* $\phi(x^\mu)$

$$A_\mu = \tilde{A}_\mu - \partial_\mu \phi. \tag{8.346}$$

According to our philosophy, we can think that, from the beginning, we have a theory on configuration space A_μ, ϕ, with the local symmetry being $A'^\mu = A^\mu$, $\phi' = \phi + \alpha$, and the action given by Eq. (8.341). The field ϕ does not enter into the action. In terms of the variables $\tilde{A}_\mu, \phi$, the transformed action reads

$$S' = \int d^4x \left[-\frac{1}{4} \tilde{F}_{\mu\nu} \tilde{F}^{\mu\nu} + \frac{1}{2} m^2 (\tilde{A}^\mu - \partial^\mu \phi)(\tilde{A}_\mu - \partial_\mu \phi) \right],$$
$$\tilde{F}_{\mu\nu} = \partial_\mu \tilde{A}_\nu - \partial_\nu \tilde{A}_\mu. \tag{8.347}$$

This is invariant under the local transformations

$$\phi \to \phi^{'} = \phi + \alpha, \qquad \tilde{A}_\mu \to \tilde{A}^{'}_\mu = \tilde{A}_\mu + \partial_\mu \alpha, \tag{8.348}$$

that is, $\tilde{A}_\mu$ transforms as an electromagnetic field. Due to this $\dot{\alpha}$-symmetry, we expect the appearance of two first-class constraints in the modified formulation. Indeed, the primary constraint of the theory (8.347) is the same as before: $\tilde{p}^0 = 0$. Then the Hamiltonian is

$$H = \int d^3x \left[\frac{1}{2}\tilde{p}_a^2 - \tilde{p}_a \partial_a \tilde{A}_0 + \frac{1}{4}\tilde{F}_{ab}^2 + \frac{1}{2m^2} p_\phi^2 + p_\phi \tilde{A}_0 + \right.$$
$$\left. \frac{1}{2} m^2 (\tilde{A}_a + \partial_a \phi)^2 + v_0 \tilde{p}^0 \right], \tag{8.349}$$

and implies the secondary constraint $\partial_a \tilde{p}_a + p_\phi = 0$. The complete constraint system

$$\tilde{p}_0 = 0, \qquad \partial_a \tilde{p}_a + p_\phi = 0, \tag{8.350}$$

is first-class. The last constraint in Eq. (8.350) states that ϕ is an auxiliary degree of freedom. It can be removed by the gauge $\phi = 0$. The first-class constraint $\tilde{p}^0 = 0$ replaces two second-class constraints (8.343) of the initial formulation, and states that A_0 is a gauge degree of freedom in the modified formulation (8.347). Equations of motion for the propagating modes in the modified theory are slightly different

$$\partial_0 \tilde{A}_a = -\tilde{p}_a + \partial_a \tilde{A}_0, \qquad \partial_0 \tilde{p}_a = -\partial_b \tilde{F}_{ba} - m^2 (\tilde{A}_a - \partial_a \phi). \tag{8.351}$$

Nevertheless, in the gauge $\phi = 0$ they coincide with the Eqs. (8.345) of the initial formulation.

Bibliography

1. E. Cartan, *Leçons sur les Invariants Intégraux* (Hermann, Paris, 1922)
2. V.I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd edn. (Springer, New York, NY, 1989)
3. A.A. Kirillov, *Elements of the Theory of Representations* (Springer, Berlin, 1976)
4. V.P. Maslov, M.V. Fedoruk, *Semiclassical Approximation in Quantum Mechanics* (D. Reidel Publishing Company, Dordrecht, 1981)
5. A.T. Fomenko, *Symplectic Geometry* (Gordon and Breach, New York, NY, 1988)
6. J.M. Souriau, *Structure des systémes dynamiques* (Dund, Paris, 1970)
7. J.E. Marsden, R.H. Abraham, *Foundations of Mechanics*, 2nd edn. (Benjamin-Cummings Publishing Company, Inc., Reading, MA, 1978)
8. P.A.M. Dirac, Can. J. Math. **2**,129 (1950); *Lectures on Quantum Mechanics* (Yeshiva University, New York, NY, 1964)
9. A.A. Slavnov, L. D. Faddeev, *Introduction in Quantum Theory of Gauge Fields* (Nauka, Moscow, 1978)
10. D.M. Gitman, I.V. Tyutin, *Quantization of Fields with Constraints* (Springer, Berlin, 1990)
11. M. Henneaux, C. Teitelboim, *Quantization of Gauge Systems* (Princeton University Press, Princeton, NJ, 1992)
12. H. Goldstein, *Classical Mechanics*, 2nd edn. (Addison-Wesley, Reading, MA, 1980)
13. L.D. Landau, E.M. Lifshits, *Mechanics* (Pergamon Press, Oxford, 1976)
14. F.R. Gantmacher, *Lectures on Analytical Mechanics* (MIR, Moscow, 1970)
15. S. Weinberg, *Gravitation and Cosmology* (Willey, New York, NY, 1972)
16. L.D. Landau, E.M. Lifshits, *The Classical Theory of Fields* (Pergamon Press, Oxford, 1980)
17. W. Pauli, *Theory of Relativity* (Pergamon Press, Oxford, 1958)
18. P.G. Bergmann, *Introduction to the Theory of Relativity* (Academic Press, New York, NY, 1967)
19. V.A. Ugarov, *Special Theory of Relativity* (Mir Publishers, Moscow, 1979)
20. R. Feynman, P. Leighton, M. Sands, *The Feynman Lectures on Physics: Commemorative Issue*, vol. 2 (Addison-Wesley, Reading, MA, 1989)
21. H. Hertz, *The Principles of Mechanics Presented in a New Form* (Dover Publications, New York, NY, 1956)
22. P.S. Wesson, *Five-Dimensional Physics: Classical and Quantum Consequences of Kaluza-Klein Cosmology* (World Scientific, Singapore, 2006)
23. V.S. Vladimirov, *Equations of Mathematical Physics*, 3rd edn. (Izdatel'stvo Nauka, Moscow, 1976), 528p. In Russian. (English translation: Equations of Mathematical Physics [by] V.S. Vladimirov. Publicado em 1971, M. Dekker, New York)
24. A.A. Deriglazov, Phys. Lett. B **626** 243–248, (2005)
25. W. Ehrenberg, R.E. Siday, Proc. R. Soc. Lond. B **62**, 8 (1949)
26. Y. Aharonov, D. Bohm, Phys. Rev. **115**, 485 (1959)
27. E. Schrödinger, Ann. Physik **81**, 109 (1926); see also letters by Shrödinger to Lorentz in: K. Przibram, Briefe zür Wellenmechanik, (Wien, 1963)

A. Deriglazov, *Classical Mechanics*, DOI 10.1007/978-3-642-14037-2

28. H. von Helmholtz, J. für Math., **C**, 151 (1886)
29. K.S. Stelle, Phys. Rev. **D16**, 953–969 (1977)
30. R.P. Woodard, *How Far Are We from the Quantum Theory of Gravity?* arXiv:0907.4238 [gr-qc]
31. M.V. Ostrogradsky, Mem. Ac. St. Petersbourg **VI 4**, 385 (1850)
32. D. Bohm, Phys. Rev. 85, 166 (1952); Phys. Rev. **85**, 180 (1952)
33. F. Mandl, *Introduction to Quantum Field Theory* (Interscience Publishers Inc., New York, 1959)
34. W. Yourgrau, S. Mandelstam, *Variational Principles in Dynamics and Quantum Theory* (Pitman, London; W. B. Sanders, Philadelphia, PA, 1968)
35. R.M. Wald, *General Relativity* (The University of Chicago Press, Chicago, IL; London, 1984)
36. J.D. Bjorken, S.D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill Book Company, New York, NY, 1964)
37. P.J. Olver, *Applications of Lie Groups to Differential Equations* (Springer, New York, NY, 1986)
38. J.L. Anderson, P.G. Bergmann, Phys. Rev. **83**, 1018 (1951); P.G. Bergmann, I. Goldberg, Phys. Rev. **98**, 531 (1955)
39. A.A. Deriglazov, Phys. Lett. A **373** 3920–3923, (2009)
40. D.J. Griffiths, *Introduction to Quantum Mechanics*, 2nd edn. (Pearson Prentice Hall, Upper Saddle River, NJ, 2005)
41. A.A. Deriglazov, SIGMA **6**, 016 (2010)
42. F.A. Berezin, M.S. Marinov, JETP Lett. **21**, 320 (1975); Ann. Phys. **104**, 336 (1977)
43. V.A. Borokhov, I.V. Tyutin, Phys. At. Nucl. **61**, 1603 (1998); Phys. At. Nucl. **62**, 10 (1999)
44. D.M. Gitman, I.V. Tyutin, Int. J. Mod. Phys. A **21**, 327 (2006)
45. A.A. Deriglazov, K.E. Evdokimov, Int. J. Mod. Phys. A **15**, 4045 (2000)
46. A.A. Deriglazov, J. Math. Phys. **50**, 012907 (2009)
47. A.A. Deriglazov, J. Phys. A **40**, 11083–11092 (2007)
48. M. Henneaux, C. Teitelboim, J. Zanelli Nucl. Phys. B **332**, 169 (1990)
49. A.A. Deriglazov, Z. Kuznetsova Phys. Lett. B **646**, 47 (2007)
50. S. Weinberg, *The Quantum Theory of Fields*, vol. 1 (Cambridge University Press, Cambridge, 1995)

Index