Chapter 1
Classical Mechanics

A macroscopic object we encounter in our daily life consists of an enormously large number of atoms. While the behavior of these atoms is governed by the laws of quantum mechanics, it is often acceptable to describe them by means of classical mechanics. In this chapter, we familiarize ourselves with the basic concepts of classical mechanics. From the outset, we assume that concepts such as mass, time, displacement, and force are understood on the basis of our everyday experiences without further elaboration. The role of classical mechanics, then is to explore the precise relationship among these objects in mathematical terms.

1.1 Inertial Frame

When describing physical phenomena, we need a frame of reference from which we make our observations. The mathematical expression of the laws of mechanics takes the simplest form when the frame is the so-called **inertial frame of reference**.

To understand what this is, imagine that you are sitting in front of your desk in your office in your building, and suppose that the desk is firmly secured to the floor and the surface of the desk is perfectly smooth and horizontal. If you place a billiard ball on the desk and remove your hand without giving any nudge to the ball, it will just sit there. If you are to give a gentle push, it will start rolling on the desk. It will continue to move in a straight line with a constant speed until it falls off at the edge of the desk. (We shall formally define the word "speed" in Sect. 1.2.1.)

These observations remain valid even if your office is housed in a train moving at a constant speed on a perfectly straight and leveled railroad track. A coordinate system attached to the desk, secured either to the building or to the train, is an example of inertial frames of reference, in which an object when subjected to no net force, either stands still or keeps on moving along a straight line at a constant speed.

However, if the train housing your office is to suddenly change direction, the ball, which was previously standing still with respect to the desk, will suddenly start moving without you doing anything to it. It is certainly more difficult to describe

I. Kusaka, *Statistical Mechanics for Engineers*,
DOI 10.1007/978-3-319-13809-1_1

the behavior of the ball in such a frame of reference. In fact, the coordinate system attached to the desk, in this case, is no longer an inertial frame until the train resumes its motion along a straight line at a constant speed. Unless explicitly indicated otherwise, we shall always use an inertial frame of reference.

1.2 Mechanics of a Single Particle

Systems of our eventual interest consist of the order of 10^{24} atoms, if not more. Before we tackle the problem of describing their behavior, we shall first learn how to describe a motion of a single particle. This allows us to introduce the essential concepts and terminologies needed for the studies of many-particle systems.

1.2.1 Newton's Equation of Motion

Let us take some macroscopic object such as a ball, and imagine that the size of the object becomes smaller and smaller while maintaining its original mass. The limit of this imaginary process is a point carrying certain nonzero mass. Such a point is referred to as a **material point** or simply a **particle**.

The position $\boldsymbol{r}$ of a material point in space may change with time. This will be the case if, for example, you are to throw one such material point, for example, a ball. The time dependence of $\boldsymbol{r}$ is determined by **Newton's equation of motion**. In an inertial frame of reference, this equation takes a particularly simple form:

$$m\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2} = \boldsymbol{F}\,, \tag{1.1}$$

where t, m, and $\boldsymbol{F}$ denote, respectively, the time, the mass of the material point, and the net force acting on it. Conversely, if $\boldsymbol{r}$ of a particle is described by (1.1) in some frame of reference, the frame is an inertial frame.

In a Cartesian coordinate system, (1.1) can be written more explicitly as

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = F_x\,, \quad m\frac{\mathrm{d}^2y}{\mathrm{d}t^2} = F_y\,, \quad \text{and} \quad m\frac{\mathrm{d}^2z}{\mathrm{d}t^2} = F_z\,, \tag{1.2}$$

where x, y, z are, respectively, the x-, y-, and z-components of the vector $\boldsymbol{r}$. Likewise for F_x, F_y, and F_z.

Example 1.1. Motion under no force: If $\boldsymbol{F}$ is zero at all time, then (1.1) tells us that

$$\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2} = \boldsymbol{0}\,, \tag{1.3}$$

which may be integrated to give

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{c}_1 \quad \text{and} \quad \boldsymbol{r} = \boldsymbol{c}_1 t + \boldsymbol{c}_2 \,. \tag{1.4}$$

When no force is acting on the material point, the vector $\boldsymbol{v} := \mathrm{d}\boldsymbol{r}/\mathrm{d}t$, called the **velocity** of the material point, is seen to be independent of time. We also observe that the material point moves along a straight line, thus confirming what was said about inertial frames of reference in Sect. 1.1.

The constant vectors $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ may be determined from initial conditions. For example, if we know that the material point was at $\boldsymbol{r}_0$ at time $t = 0$ and that it was moving with the velocity $\boldsymbol{v}_0$ at $t = 0$, then substituting this information into (1.4), we obtain

$$\boldsymbol{c}_1 = \boldsymbol{v}_0 \quad \text{and} \quad \boldsymbol{c}_2 = \boldsymbol{r}_0 \,, \tag{1.5}$$

and hence

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{v}_0 \quad \text{and} \quad \boldsymbol{r} = \boldsymbol{v}_0 t + \boldsymbol{r}_0 \,. \tag{1.6}$$

Example 1.2. Motion under gravity: Let us turn to a slightly more interesting problem of throwing a ball. If we ignore the friction between the ball and the air, the only force acting on the ball after it left your hand is the force of gravity $m\boldsymbol{g}$, where the gravitational acceleration $\boldsymbol{g}$ is a vector pointing toward the center of the earth. In this case, (1.1) becomes

$$m\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2} = m\boldsymbol{g} \,. \tag{1.7}$$

Assuming that $\boldsymbol{g}$ is constant, which is an excellent approximation over the time and the length scales involved in this example, the equation can be readily integrated to give

$$\boldsymbol{v} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{g}t + \boldsymbol{c}_1 \quad \text{and} \quad \boldsymbol{r} = \frac{1}{2}\boldsymbol{g}t^2 + \boldsymbol{c}_1 t + \boldsymbol{c}_2 \,, \tag{1.8}$$

where $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ are constant vectors. Unless $\boldsymbol{g}$ is parallel to $\boldsymbol{c}_1$ or $\boldsymbol{c}_1 = \boldsymbol{0}$, the trajectory of the particle is *not* a straight line.

As before, let $\boldsymbol{r} = \boldsymbol{r}_0$ and $\boldsymbol{v} = \boldsymbol{v}_0$ at $t = 0$. Then, $\boldsymbol{c}_1 = \boldsymbol{v}_0$ and $\boldsymbol{c}_2 = \boldsymbol{r}_0$. With the vectors $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ so determined, (1.8) completely determines $\boldsymbol{r}$ and $\boldsymbol{v}$ at any other value of t.

The two examples we have just seen point to a general scheme of classical mechanics. That is, if $\boldsymbol{F}$ is at most a known function of $\boldsymbol{r}$, $\boldsymbol{v}$, and t, then $\boldsymbol{r}$ and $\boldsymbol{v}$ of a particle at any time t are completely determined once they are specified at $t = 0$ or at any other instant of time for that matter. The pair of vectors $(\boldsymbol{r}, \boldsymbol{v})$ associated

with a single particle at some instant of time is referred to as the **mechanical state** of that particle at that instant. It follows that the mechanical state of a particle is completely determined for all t if it is specified at some instant, say $t = 0$.[2]

The length $\|\boldsymbol{v}\| := \sqrt{\boldsymbol{v}\cdot\boldsymbol{v}} = \sqrt{{v_x}^2 + {v_y}^2 + {v_z}^2}$ of the velocity vector $\boldsymbol{v}$ is called the **speed**, where v_x, v_y, and v_z denote, respectively, the x-, y-, and z- components of the vector $\boldsymbol{v}$. We shall often write v instead of $\|\boldsymbol{v}\|$. Finally, $\mathrm{d}^2\boldsymbol{r}/\mathrm{d}t^2$ is the **acceleration**. For a brief review of vector algebra, see Appendix A.

We can now show that a frame of reference moving at a constant velocity with respect to an inertial frame is also an inertial frame. To see this, consider motion of a particle in two frames $\mathscr{O}$ and $\mathscr{O}'$, with respect to which the position vectors of the particle is $\boldsymbol{r}$ and $\boldsymbol{r}'$, respectively. Let $\boldsymbol{R}$ denote the position vector of the origin of $\mathscr{O}'$ with respect to that of $\mathscr{O}$. Then,

$$\boldsymbol{r} = \boldsymbol{r}' + \boldsymbol{R}\,. \tag{1.9}$$

If we suppose that $\mathscr{O}$ is an inertial frame, (1.1) holds for $\boldsymbol{r}$. Using (1.9), we find

$$m\frac{\mathrm{d}^2\boldsymbol{r}'}{\mathrm{d}t^2} = \boldsymbol{F} - m\frac{\mathrm{d}^2\boldsymbol{R}}{\mathrm{d}t^2}\,. \tag{1.10}$$

If the velocity $\mathrm{d}\boldsymbol{R}/\mathrm{d}t$ of $\mathscr{O}'$ with respect to $\mathscr{O}$ is constant, the second term on the right-hand side vanishes identically and we obtain

$$m\frac{\mathrm{d}^2\boldsymbol{r}'}{\mathrm{d}t^2} = \boldsymbol{F}\,, \tag{1.11}$$

indicating that $\mathscr{O}'$ is an inertial frame.

If $\mathscr{O}'$ is accelerating with respect to $\mathscr{O}$, then $\mathrm{d}^2\boldsymbol{R}/\mathrm{d}t^2$ must be retained in (1.10). At this point, however, our definition of inertial frames appears circular. In Newtonian mechanics, a force acting on a particle is defined by means of Newton's equation of motion in the form of (1.1). That is, the force is measured by the acceleration it produces.[3] Therefore, an observer in $\mathscr{O}'$ who is oblivious to the fact that his frame is not an inertial frame would simply compute the force $\boldsymbol{F}'$ acting on the particle as $m\mathrm{d}^2\boldsymbol{r}'/\mathrm{d}t^2$. This vector is, of course, equal to $\boldsymbol{F} - m\mathrm{d}^2\boldsymbol{R}/\mathrm{d}t^2$. But, this observer is under no obligation to separate $\boldsymbol{F}'$ into $\boldsymbol{F}$ and $-m\mathrm{d}^2\boldsymbol{R}/\mathrm{d}t^2$. Instead, he can insist that the frame $\mathscr{O}'$ is inertial and that the force acting on the particle is $\boldsymbol{F}'$, not $\boldsymbol{F}$. A tacit assumption made in Newtonian mechanics, therefore, is that forces can arise only through an effect a body produces onto another and that we can systematically either eliminate or account for such effects.

1.2.2 Work

If a particle experiences a displacement $\Delta\boldsymbol{r}$ under the influence of a force $\boldsymbol{F}$ and the force remained constant during the displacement, then the **work** W done *by* this

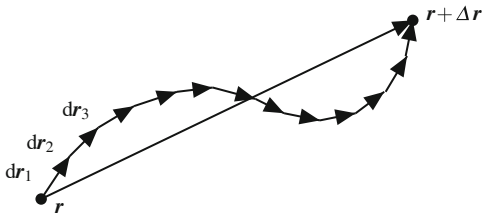

Fig. 1.1 The sum of infinitesimal displacements $\mathrm{d}\boldsymbol{r}_i$ results in a net displacement $\int \mathrm{d}\boldsymbol{r} = \sum_{i=1}^{n} \mathrm{d}\boldsymbol{r}_i = \Delta\boldsymbol{r}$.

force *on* the particle is, by definition, the dot product of $\boldsymbol{F}$ and $\Delta\boldsymbol{r}$:

$$W := \boldsymbol{F} \cdot \Delta\boldsymbol{r}\,, \tag{1.12}$$

This definition does not preclude the situation in which other forces are also acting on the particle. In that case, (1.12) simply gives the work done by this particular force $\boldsymbol{F}$ alone and does not include the work done by the other forces.

It may be that $\boldsymbol{F}$ actually changes during the displacement. To allow for this more general situation, we define the work by the equation

$$W := \int \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r}\,, \tag{1.13}$$

where the integration is along the path taken by the particle. When $\boldsymbol{F}$ remains constant along the path, (1.13) may be written as

$$W = \boldsymbol{F} \cdot \int \mathrm{d}\boldsymbol{r}\,. \tag{1.14}$$

If we regard the integral as a "sum" of infinitesimal displacements $\mathrm{d}\boldsymbol{r}_1, \mathrm{d}\boldsymbol{r}_2, \mathrm{d}\boldsymbol{r}_3, \ldots$, the integral is equal to $\Delta\boldsymbol{r}$ as seen from Fig. 1.1, and (1.14) reduces to (1.12). Thus, (1.13) contains (1.12) as a special case.

Let us suppose, as illustrated in Fig. 1.2, that we lift a particle of mass m from $z = h_1$ to $z = h_2$ under gravity. In order to do this, we have to apply a force $\boldsymbol{F}$ just enough to overcome the force exerted on the particle by gravity $m\boldsymbol{g}$. In terms of components with respect to the coordinate system shown in Fig. 1.2, $m\boldsymbol{g} \doteq (0, 0, -mg)$,[4] where the negative sign is needed because the force of gravity points in the negative z-direction. Thus, the force needed is $\boldsymbol{F} \doteq (0, 0, mg + \varepsilon)$, where $\varepsilon > 0$. Since the displacement is in the positive z-direction, $\mathrm{d}\boldsymbol{r} \doteq (0, 0, \mathrm{d}z)$. Combining everything,

$$\begin{aligned} W \;&=\; \int \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} = \int_{h_1}^{h_2} (mg + \varepsilon)\mathrm{d}z = (mg + \varepsilon)(h_2 - h_1) \\ &\to mg(h_2 - h_1) \quad \text{as} \quad \varepsilon \to 0\,. \end{aligned} \tag{1.15}$$

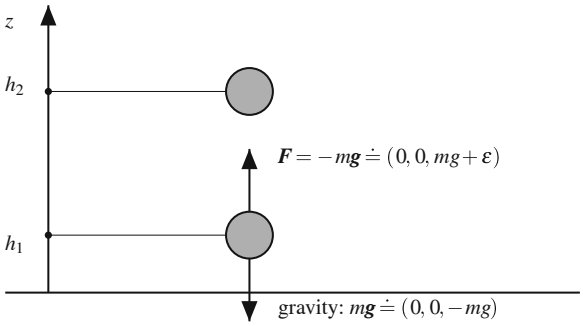

Fig. 1.2 A particle under gravity.

We define a new quantity

$$\psi(z) := mgz \tag{1.16}$$

and call it the **potential energy** of the particle due to gravity when it is at height z above the ground. With this definition, (1.15) becomes

$$W = \psi(h_2) - \psi(h_1) \,. \tag{1.17}$$

We also notice that

$$-\frac{\mathrm{d}\psi(z)}{\mathrm{d}z} = -mg \,, \tag{1.18}$$

that is, by taking the derivative of ψ with respect to z and putting a negative sign in front of it, we obtain the z-component of the force of gravity. This is a general feature of a potential energy as we shall see in the next subsection.

1.2.3 Potential Energy

More generally, potential energy can be introduced as follows. We suppose that the force $\boldsymbol{F}_c$ acting on a particle is a function of its position. As was the case with lifting of a particle under gravity in the previous section, the force $\boldsymbol{F} = -\boldsymbol{F}_c$ is needed to counter $\boldsymbol{F}_c$ and move the particle.[5] If the particle is made to move along a closed path, such as the one depicted in Fig. 1.3, the work done by the force $\boldsymbol{F}$ is given by

$$W = \oint \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} \,, \tag{1.19}$$

where $\oint$ signifies the integration along a *closed* path. If this quantity is identically zero for *any* closed path, including those that does not necessarily pass through the points A, B, C, and D in the figure, then $\boldsymbol{F}_c$ is said to be **conservative**.

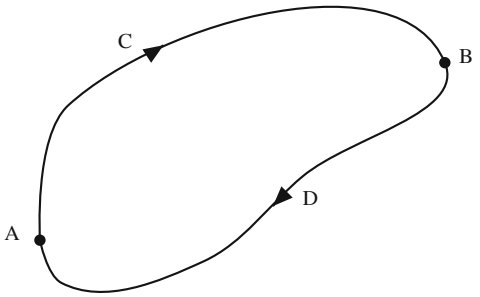

Fig. 1.3 A closed path along which the particle is moved.

The particular closed path *ACBDA* shown in Fig. 1.3 can be regarded as consisting of two parts, the path $A \to C \to B$ and the path $B \to D \to A$. Thus,

$$W = \int_{A\to C\to B} \boldsymbol{F}\cdot \mathrm{d}\boldsymbol{r} + \int_{B\to D\to A} \boldsymbol{F}\cdot \mathrm{d}\boldsymbol{r} =: W_{A\to C\to B} + W_{B\to D\to A}\ . \tag{1.20}$$

Now, suppose that we reverse the second part of the path and bring the particle from A to B through D. The infinitesimal displacement $\mathrm{d}\boldsymbol{r}$ on this reversed path $A \to D \to B$ will have the opposite sign to that on the original path $B \to D \to A$. But, the force $\boldsymbol{F}_c(\boldsymbol{r})$, and hence $\boldsymbol{F}(\boldsymbol{r})$, will remain unaffected by the reversal of the path. Thus,

$$W_{A\to D\to B} = -W_{B\to D\to A}\ , \tag{1.21}$$

and hence

$$W = W_{A\to C\to B} - W_{A\to D\to B}\ . \tag{1.22}$$

If $\boldsymbol{F}_c$ is conservative, then $W = 0$ and we have

$$W_{A\to C\to B} = W_{A\to D\to B}\ . \tag{1.23}$$

In other words, the work required to move a particle from point A to point B against the conservative force $\boldsymbol{F}_c$ is independent of the actual path taken and depends only on the positions of the end points A and B. Thus, the work can be written as

$$W_{A\to B} = \psi(\boldsymbol{r}_A, \boldsymbol{r}_B)\ , \tag{1.24}$$

without the superfluous references to C and D.

We refer to the quantity $\psi(\boldsymbol{r}_A, \boldsymbol{r}_B)$ as the **potential energy** of the particle at $\boldsymbol{r}_B$ with respect to the reference point $\boldsymbol{r}_A$. When $\psi(\boldsymbol{r}_A, \boldsymbol{r})$ is known for all $\boldsymbol{r}$ of interest, we can calculate the work required to move a particle between any pair of points. For example, we have

$$\begin{aligned} W_{X\to Y} &= W_{X\to A\to Y} = W_{X\to A} + W_{A\to Y} = -W_{A\to X} + W_{A\to Y} \\ &= -\psi(\boldsymbol{r}_A, \boldsymbol{r}_X) + \psi(\boldsymbol{r}_A, \boldsymbol{r}_Y)\ . \end{aligned} \tag{1.25}$$

Note that $\psi(\boldsymbol{r}_A,\boldsymbol{r})$ depends on our choice of $\boldsymbol{r}_A$. However, since $W_{X\to Y}$ depends only on $\boldsymbol{r}_X$ and $\boldsymbol{r}_Y$, any $\boldsymbol{r}_A$ dependence of ψ must cancel out in (1.25). Thus, as long as we are interested in computing the work required to move a particle between two points, the location of $\boldsymbol{r}_A$ is arbitrary.

Now, let us calculate the work required to move the particle by an infinitesimal distance from $\boldsymbol{r}$ to $\boldsymbol{r}+\mathrm{d}\boldsymbol{r}$. According to (1.25), the work is given by

$$\begin{aligned} W &= -\psi(\boldsymbol{r}_A,\boldsymbol{r})+\psi(\boldsymbol{r}_A,\boldsymbol{r}+\mathrm{d}\boldsymbol{r}) = \psi(\boldsymbol{r}_A,\boldsymbol{r}+\mathrm{d}\boldsymbol{r})-\psi(\boldsymbol{r}_A,\boldsymbol{r}) \\ &= \psi(\boldsymbol{r}_A,x+\mathrm{d}x,y+\mathrm{d}y,z+\mathrm{d}z)-\psi(\boldsymbol{r}_A,x,y,z)\,. \end{aligned} \tag{1.26}$$

Using equations from Appendix B.1, we expand the right-hand side into the Taylor series to obtain

$$W = \mathrm{d}\psi + \text{h.o.}, \tag{1.27}$$

where h.o. stands for the second- and higher order terms in $\mathrm{d}x$, $\mathrm{d}y$, and $\mathrm{d}z$, and

$$\mathrm{d}\psi = \frac{\partial\psi}{\partial x}\mathrm{d}x+\frac{\partial\psi}{\partial y}\mathrm{d}y+\frac{\partial\psi}{\partial z}\mathrm{d}z =: \mathrm{d}\boldsymbol{r}\cdot\nabla\psi\,. \tag{1.28}$$

In the last step, we introduced a new symbol $\nabla\psi$. This vector is known as the **gradient** of ψ and may also be written as $\partial\psi/\partial\boldsymbol{r}$. In a Cartesian coordinate system, in which

$$\mathrm{d}\boldsymbol{r} \doteq (\mathrm{d}x,\mathrm{d}y,\mathrm{d}z), \tag{1.29}$$

we have

$$\nabla\psi \equiv \frac{\partial\psi}{\partial\boldsymbol{r}} \doteq \left(\frac{\partial\psi}{\partial x},\frac{\partial\psi}{\partial y},\frac{\partial\psi}{\partial z}\right)\,. \tag{1.30}$$

In other coordinate systems, components of $\mathrm{d}\boldsymbol{r}$ are not $(\mathrm{d}x,\mathrm{d}y,\mathrm{d}z)$. Accordingly, the components of $\nabla\psi$ differ from those displayed here.

Comparing (1.12) and (1.28), we identify $\nabla\psi$ as the force $\boldsymbol{F}$ required to move the particle. It follows that

$$\boldsymbol{F}_c = -\nabla\psi\,, \tag{1.31}$$

which should be compared with (1.18).

In classical mechanics, we take a point of view that the physical reality at some instant t concerning a particle is defined completely by the pair of quantities $\boldsymbol{r}(t)$ and $\boldsymbol{v}(t)$. How these quantities evolve with time is determined by Newton's equation of motion, and hence by the force acting on the particle and initial conditions. It seems natural, then, to expect that the force $\boldsymbol{F}_c$ computed by (1.31) should not in any way depend on our *arbitrary* choice of $\boldsymbol{r}_A$.

Exercise 1.1. Show that

$$\nabla\psi(\boldsymbol{r}_{A'},\boldsymbol{r}) = \nabla\psi(\boldsymbol{r}_A,\boldsymbol{r})\,, \tag{1.32}$$

that is, the force $\boldsymbol{F}_c$ is indeed independent of the choice of $\boldsymbol{r}_A$. ///

In view of the indifference of both $W_{X\to Y}$ and $\boldsymbol{F}_c$ to our choice of $\boldsymbol{r}_A$, we shall suppress the explicit reference to $\boldsymbol{r}_A$ in the function ψ and denote the potential energy

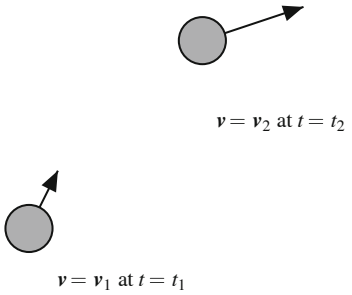

Fig. 1.4 Accelerating a particle.

simply as $\psi(\boldsymbol{r})$ even though its value certainly do depend on $\boldsymbol{r}_A$. From (1.24), we see that $\psi(\boldsymbol{r}_A, \boldsymbol{r}_A) = 0$. After dropping the first reference to $\boldsymbol{r}_A$ according to the convention just adopted, we obtain $\psi(\boldsymbol{r}_A) = 0$. We interpret this equation as identifying $\boldsymbol{r}_A$ as the reference point with respect to which $\psi(\boldsymbol{r})$ is computed.

1.2.4 Kinetic Energy

Let us calculate the work required to change the velocity of a particle of mass m from $\boldsymbol{v}_1$ at $t = t_1$ to $\boldsymbol{v}_2$ at time t_2 as illustrated in Fig. 1.4. The net force $\boldsymbol{F}$ exerted on it cannot be zero: Newton's equation of motion tells us that $\boldsymbol{v}$ remains constant otherwise. What we want to compute is the work done by this force, whether it is applied by you, some field of force such as gravity, or both. Using (1.1) in (1.13),

$$W = \int_{(t=t_1)}^{(t=t_2)} m \frac{\mathrm{d}^2 \boldsymbol{r}}{\mathrm{d}t^2} \cdot \mathrm{d}\boldsymbol{r} \,. \tag{1.33}$$

Because $\boldsymbol{r}$ is a function of t, we have $\mathrm{d}\boldsymbol{r} = (\mathrm{d}\boldsymbol{r}/\mathrm{d}t)\mathrm{d}t$, which allows us to change the integration variables from $\boldsymbol{r}$ to t as

$$W = \int_{t_1}^{t_2} m \frac{\mathrm{d}^2 \boldsymbol{r}}{\mathrm{d}t^2} \cdot \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \mathrm{d}t = \frac{1}{2} m \int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \cdot \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \right) \mathrm{d}t \,. \tag{1.34}$$

You can verify the second equality by evaluating the last member of the equation (with a help of the product rule of differentiation) to restore the second one. Since $\mathrm{d}\boldsymbol{r}/\mathrm{d}t = \boldsymbol{v}$,

$$W = \frac{1}{2} m \int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t} (\boldsymbol{v} \cdot \boldsymbol{v}) \, \mathrm{d}t = \frac{1}{2} m [\boldsymbol{v} \cdot \boldsymbol{v}]_{t_1}^{t_2} = \frac{1}{2} m \left({v_2}^2 - {v_1}^2 \right) \,. \tag{1.35}$$

The resulting expression for W is seen to depend only on the mass of the particle and its velocities at t_1 and t_2. It is independent of the actual path taken by the particle,

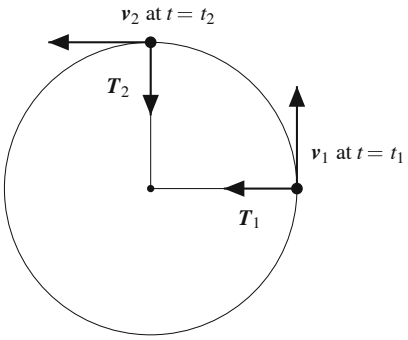

Fig. 1.5 In the case of a particle moving on a circle at a constant speed, the velocity $\boldsymbol{v}$ changes its direction without changing its length. A force is still needed to change the direction of $\boldsymbol{v}$ and is supplied, for example, by the tension $\boldsymbol{T}$ in a string to which the particle is attached. The work W done on the particle by $\boldsymbol{T}$, however, is zero.

that is, how $\boldsymbol{r}$ changed with t. In fact, we have made no assumption in this regard. In this sense, $mv^2/2$ is a reflection of an intrinsic property m of the particle and its instantaneous mechanical state, $\boldsymbol{v}$. The quantity $mv^2/2$ is called the **kinetic energy** of the particle.

From (1.34), we observe that $W \equiv 0$ if the acceleration $\mathrm{d}^2\boldsymbol{r}/\mathrm{d}t^2$ is always perpendicular to the velocity $\mathrm{d}\boldsymbol{r}/\mathrm{d}t$. Equation (1.35) then indicates that the length of the velocity vector remains constant. Such a motion can be observed directly if you attach a particle to one end of a (massless) string and swing it around the other end as shown in Fig. 1.5.

Exercise 1.2. In the derivation of (1.15), we tacitly assumed that the kinetic energy acquired by the particle is negligible. Validate this assumption in the limit of $\varepsilon \to 0$. You may assume that the particle is at rest initially at the height h_1. ///

1.2.5 Conservation of Energy

Let us consider a particle moving under the influence of a conservative force $-\nabla\psi$ and an additional external force $\boldsymbol{F}_{\mathrm{ext}}$. We calculate the work done on the particle by $\boldsymbol{F}_{\mathrm{ext}}$. In this case, Newton's equation of motion reads

$$m\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2} = -\nabla\psi(\boldsymbol{r}) + \boldsymbol{F}_{\mathrm{ext}}\,, \tag{1.36}$$

from which we have

$$W = \int_{(t=t_1)}^{(t=t_2)} \boldsymbol{F}_{\mathrm{ext}}\cdot\mathrm{d}\boldsymbol{r} = \int_{(t=t_1)}^{(t=t_2)} \left[m\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2} + \nabla\psi(\boldsymbol{r})\right]\cdot\mathrm{d}\boldsymbol{r}\,. \tag{1.37}$$

Evidently, the two terms in the integrand can be integrated separately. The integration of the first term, as before, yields $mv_2{}^2/2 - mv_1{}^2/2$. The integration of the second term is almost immediate if we recall (1.28):

$$\int_{(t=t_1)}^{(t=t_2)} \nabla\psi(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} = \int_{(t=t_1)}^{(t=t_2)} \mathrm{d}\psi = \psi(\boldsymbol{r}_2) - \psi(\boldsymbol{r}_1) \,. \tag{1.38}$$

Note that we were able to compute the integral without knowing the path taken by the particle. This should not surprise you since ψ is a potential energy.

Combining the results, we arrive at the expression

$$W = \left[\frac{1}{2}mv_2{}^2 + \psi(\boldsymbol{r}_2)\right] - \left[\frac{1}{2}mv_1{}^2 + \psi(\boldsymbol{r}_1)\right] \,. \tag{1.39}$$

The kinetic energy plus the potential energy of a particle, that is,

$$E := \frac{1}{2}mv^2 + \psi(\boldsymbol{r}) \,, \tag{1.40}$$

is called the **mechanical energy** or simply **energy** of the particle. According to (1.39), the change ΔE in the energy of a particle is equal to the work done on the particle by the external force $\boldsymbol{F}_{\text{ext}}$:

$$\Delta E = W \,. \tag{1.41}$$

From (1.37), we see that $W = 0$ if $\boldsymbol{F}_{\text{ext}} \equiv \boldsymbol{0}$. According to (1.41), this means that E remains constant. In other words, if a particle moves under the influence of a conservative force *alone*, then the energy of the particle does not change with time despite the fact that $\boldsymbol{r}$ and $\boldsymbol{v}$, upon which E depends, do change with time. Such a quantity is called a **constant of motion** or a **conserved quantity**. Mechanical energy is an example of constants of motion.

1.3 Mechanics of Many Particles

So far, we focused on the mechanics of a single particle, and treated a ball as if it is a material point. A more satisfactory description of the problem would treat the ball as a collection of many material points, each representing perhaps an atom making up the ball. These material points may also interact with each other as well as with an external field. **Newton's equation of motion** generalizes quite naturally to this case and we have

$$m_i \frac{\mathrm{d}^2\boldsymbol{r}_i}{\mathrm{d}t^2} = \boldsymbol{F}_i \,, \quad i = 1, \ldots, N \,, \tag{1.42}$$

where N is the number of material points. As before, $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ at any time t are determined completely by these equations and initial conditions, that is, $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ at time $t = 0$. In keeping with the case of a single particle system, by a **mechanical state** of a system at some instant t, we shall understand

the set of vectors $(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_N)$ referring to the positions and velocities of all the particles in the system at that moment. Thus, Newton's equations of motion, along with initial conditions, completely specify the mechanical state of a many-particle system at any other instant.

1.3.1 Mechanical Energy of a Many-Particle System

Let us calculate the work required to bring about a particular change in the mechanical state of a many-particle system.

If the system is subject to a conservative field ψ such as gravity, the ith particle experiences the force given by

$$-\nabla_i \psi(\boldsymbol{r}_i) \doteq \left(-\frac{\partial \psi}{\partial x_i}, -\frac{\partial \psi}{\partial y_i}, -\frac{\partial \psi}{\partial z_i}\right), \quad i = 1, \ldots, N \tag{1.43}$$

with x_i, y_i, and z_i denoting the x-, y-, and z-components of $\boldsymbol{r}_i$, respectively. The subscript i on ∇ indicates that the derivative is with respect to the position vector of the ith particle.

We assume that the interaction among the particles is characterized by a conservative force derivable from a single **potential energy** function $\phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$. That is

$$-\nabla_i \phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \doteq \left(-\frac{\partial \phi}{\partial x_i}, -\frac{\partial \phi}{\partial y_i}, -\frac{\partial \phi}{\partial z_i}\right), \quad i = 1, \ldots, N \tag{1.44}$$

gives the force exerted on the ith particle by all the other particles in the system. We arbitrarily set $\phi = 0$ when particles are separated sufficiently far away from each other that there is no interaction among them. If, in addition, the ith particle is subject to another external force $\boldsymbol{F}_{\text{ext},i}$, (1.42) becomes

$$m_i \frac{\mathrm{d}^2 \boldsymbol{r}_i}{\mathrm{d}t^2} = -\nabla_i \psi(\boldsymbol{r}_i) - \nabla_i \phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) + \boldsymbol{F}_{\text{ext},i}, \quad i = 1, \ldots, N. \tag{1.45}$$

Thus, the work W done on the many-particle system by the external forces $\boldsymbol{F}_{\text{ext},i}$ $(i = 1, \ldots, N)$ is given by

$$W = \sum_{i=1}^{N} \int_{(t=t_1)}^{(t=t_2)} \boldsymbol{F}_{\text{ext},i} \cdot \mathrm{d}\boldsymbol{r}_i = \sum_{i=1}^{N} \int_{(t=t_1)}^{(t=t_2)} \left[m_i \frac{\mathrm{d}^2 \boldsymbol{r}_i}{\mathrm{d}t^2} + \nabla_i \psi(\boldsymbol{r}_i) + \nabla_i \phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \right] \cdot \mathrm{d}\boldsymbol{r}_i . \tag{1.46}$$

The first two terms can be integrated as before and yield

$$\sum_{i=1}^{N} \left[\frac{1}{2} m_i {v_i}^2 + \psi(\boldsymbol{r}_i) \right]_{t_2} - \sum_{i=1}^{N} \left[\frac{1}{2} m_i {v_i}^2 + \psi(\boldsymbol{r}_i) \right]_{t_1}, \tag{1.47}$$

where we used the subscripts t_1 and t_2 to indicate that the expressions in the square brackets are to be computed at the initial state (at $t = t_1$) and the final state (at $t = t_2$), respectively. The last term of (1.46) can be computed by noting that

$$\sum_{i=1}^{N} \nabla_i \phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \cdot \mathrm{d}\boldsymbol{r}_i = \sum_{i=1}^{N} \left(\frac{\partial \phi}{\partial x_i} \mathrm{d}x_i + \frac{\partial \phi}{\partial y_i} \mathrm{d}y_i + \frac{\partial \phi}{\partial z_i} \mathrm{d}z_i \right) = \mathrm{d}\phi \,, \tag{1.48}$$

where the last step follows from (B.16) generalized to a function of $3N$ variables. Thus,

$$\sum_{i=1}^{N} \int_{(t=t_1)}^{(t=t_2)} \nabla_i \phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \cdot \mathrm{d}\boldsymbol{r}_i = \int_{(t=t_1)}^{(t=t_2)} \mathrm{d}\phi = \phi_2(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) - \phi_1(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \,, \tag{1.49}$$

where the subscripts 1 and 2 on ϕ refer to the initial and the final states, respectively. Combining everything, we arrive at

$$W = \Delta E \,, \tag{1.50}$$

where we defined the **mechanical energy** of the many-particle system by

$$E = \sum_{i=1}^{N} \left[\frac{1}{2} m_i {v_i}^2 + \psi(\boldsymbol{r}_i) \right] + \phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \,, \tag{1.51}$$

in which

$$\sum_{i=1}^{N} \frac{1}{2} m_i {v_i}^2 \tag{1.52}$$

is the **kinetic energy** of the many-particle system. If the system evolves under the influence of conservative forces (derivable from ψ and ϕ) *alone*, that is, if $\boldsymbol{F}_{\mathrm{ext},i} \equiv 0$ for all $i = 1, \ldots, N$, then its mechanical energy is a constant of motion.

1.4 Center of Mass

In dealing with a many-particle system, it is often convenient to work with its **center of mass**, which is defined by

$$\boldsymbol{R} := \frac{1}{M} \sum_{i=1}^{N} m_i \boldsymbol{r}_i \,, \tag{1.53}$$

where $M := \sum_{i=1}^{N} m_i$ is the total mass of the system.

Exercise 1.3. Consider a collection of N particles under gravity. If we take the coordinate system so that the z-axis points vertically upward, then

$$\psi(\boldsymbol{r}_i) = m_i g z_i \tag{1.54}$$

and (1.51) reduces to

$$E = \sum_{i=1}^{N} \left[\frac{1}{2} m_i {v_i}^2 + m_i g z_i \right] + \phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \,. \tag{1.55}$$

Show that this expression can be written as

$$E = \frac{1}{2} M V^2 + MgZ + \sum_i \frac{1}{2} m_i {v_i}'^2 + \phi(\boldsymbol{r}_1', \ldots, \boldsymbol{r}_N') \,, \tag{1.56}$$

where $\boldsymbol{V} := \mathrm{d}\boldsymbol{R}/\mathrm{d}t$ is the velocity of the center of mass, $V := \|\boldsymbol{V}\|$, Z is the z-component of $\boldsymbol{R}$, $\boldsymbol{r}_i' := \boldsymbol{r}_i - \boldsymbol{R}$, and $\boldsymbol{v}_i' := \mathrm{d}\boldsymbol{r}_i'/\mathrm{d}t$ for $i = 1, \ldots, N$. You will need to use the fact that ϕ depends on the relative positions of particles, not on their absolute position in space. ///

The first two terms of (1.56) is the mechanical energy the collection of N particles would have if all the particles were concentrated at $\boldsymbol{R}$ to form a single material point of mass M. The remaining term is the mechanical energy of the N particles computed in a (generally noninertial) coordinate system whose origin is attached to $\boldsymbol{R}$.

1.5 Hamilton's Principle

Mechanical behavior of a many-particle system is determined completely by (1.42) and initial conditions. However, for a system consisting of 10^{24} particles or more, which are of our principal interest, the actual solution of the equations of motion is practically impossible to obtain. More importantly, the solution contains far more detailed information than we possibly need. For example, you probably will not send your soup back in a restaurant on the basis that you do not like the position of a particular molecule in it. Our primary concern in this case will be in a much less detailed description of the system, such as its temperature, volume, and composition. Thus, our goal must be to extract such information on a system of many particles without ever solving (1.42). Newton's formulation of mechanics is not particularly useful for developing such a scheme. For that, we have to reformulate classical mechanics first, which is the subject of this section and, in fact, the rest of the chapter.

1.5.1 Lagrangian and Action

We define a quantity called **Lagrangian** L by

$$L := (\text{Kinetic Energy}) - (\text{Potential Energy}) \,. \tag{1.57}$$

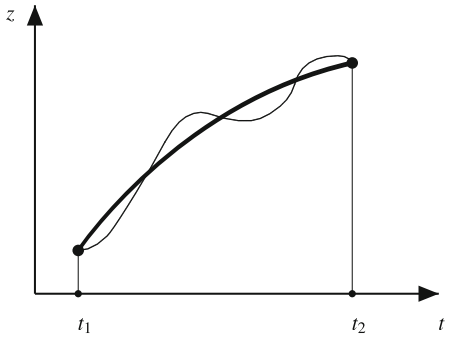

Fig. 1.6 The actual path $z(t)$ followed by the mechanical system (*thick curve*) and a varied path $z'(t) = z(t) + \delta z(t)$ (*thin wavy curve*). Note that $\delta z = 0$ at t_1 and t_2.

For example, the Lagrangian of a particle in a uniform gravitational field is given by

$$L = \frac{1}{2} m\dot{z}^2 - mgz \,, \tag{1.58}$$

where we took the z-axis vertically upward. We also introduced a short-hand notation $\dot{z} := \mathrm{d}z/\mathrm{d}t$ and assumed that the particle moves only in the z-direction. We see that, once z is known as a function of time t, we can find L as a function only of t. This function can then be integrated with respect to t over some interval of time $t_1 \le t \le t_2$, yielding the quantity called the **action integral** or action:

$$\mathscr{S}[z] := \int_{t_1}^{t_2} L(z, \dot{z}) \mathrm{d}t \,. \tag{1.59}$$

The value of the action $\mathscr{S}$ depends on the form of the function $z(t)$. We express this dependence by means of the notation $\mathscr{S}[z]$.

Clearly, the value of $\mathscr{S}$ can be calculated not only for the actual path $z(t)$ followed by the mechanical system under consideration but also for any other path $z'(t)$, for which we obtain $\mathscr{S}[z']$. (z' is not to be confused with $\dot{z}$.) You might wonder if computation of $\mathscr{S}$ for any path other than the actual one serves any useful purpose at all. As we shall find out, however, such a computation leads to a useful and surprisingly simple (at least conceptually) way to characterize the actual path.

Now, let $z'(t) = z(t) + \delta z(t)$ so that $z'(t)$ differs only infinitesimally from the actual path $z(t)$ and calculate $\mathscr{S}[z + \delta z]$. (See Fig. 1.6.) **Hamilton's principle** states that, for any $\delta z(t)$,

$$\mathscr{S}[z + \delta z] - \mathscr{S}[z] = 0 \,, \tag{1.60}$$

to the first order of $\delta z(t)$ *provided that*

$$z(t_1) = z'(t_1) \quad \text{and} \quad z(t_2) = z'(t_2) \,, \tag{1.61}$$

that is, the perturbed path $z'(t)$ coincides with the actual path $z(t)$ at the beginning and at the end of the time period over which we consider the time evolution of the system. This may be expressed as

$$\delta z(t_1) = \delta z(t_2) = 0\,. \tag{1.62}$$

The meaning of the phrase *to the first order of* δz is explained in Appendix B.1 and will be made clearer as we explore the consequence of Hamilton's principle.

First, we evaluate the difference

$$\mathscr{S}[z+\delta z] - \mathscr{S}[z] = \int_{t_1}^{t_2} \left[L\left(z+\delta z, \frac{\mathrm{d}}{\mathrm{d}t}(z+\delta z)\right) - L(z,\dot{z}) \right] \mathrm{d}t\,. \tag{1.63}$$

Recalling the definition of the short-hand notation we introduced earlier, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}(z+\delta z) = \frac{\mathrm{d}z}{\mathrm{d}t} + \frac{\mathrm{d}}{\mathrm{d}t}(\delta z) = \dot{z} + \dot{\delta} z\,, \tag{1.64}$$

and hence

$$\mathscr{S}[z+\delta z] - \mathscr{S}[z] = \int_{t_1}^{t_2} \left[L(z+\delta z, \dot{z}+\dot{\delta} z) - L(z,\dot{z}) \right] \mathrm{d}t\,. \tag{1.65}$$

To extract the first-order term, we expand the integrand in the Taylor series:

$$L(z+\delta z, \dot{z}+\dot{\delta} z) - L(z,\dot{z}) = \frac{\partial L}{\partial z}\delta z + \frac{\partial L}{\partial \dot{z}}\dot{\delta} z + \text{h.o.}\,, \tag{1.66}$$

Bringing (1.66) into (1.65), we find

$$\mathscr{S}[z+\delta z] - \mathscr{S}[z] = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial z}\delta z + \frac{\partial L}{\partial \dot{z}}\dot{\delta} z \right) \mathrm{d}t + \text{h.o.} \tag{1.67}$$

The development up to this point might make you a bit uneasy. On the one hand, once the entire path $z(t)$ is specified, so is $\dot{z}(t) = \mathrm{d}z/\mathrm{d}t$. The same remark applies to the varied path as we have seen in (1.64). On the other hand, in (1.66), we considered the partial derivative of L with respect to z taken *while holding* $\dot{z}$ *constant.* Likewise for the partial derivative with respect to $\dot{z}$. That is, we are treating z and $\dot{z}$ as if they are independent. *Is this legitimate?* Given a function L that is expressed in terms of z and $\dot{z}$, this surely is a well-defined procedure mathematically, and there should be no doubt about the validity of (1.66), which is nothing but a Taylor series expansion. The real question, then is in what sense z and $\dot{z}$ can be regarded as independent. Why is L a function of both z and $\dot{z}$ and not a function of z alone? The reason is that, for a given value of z at any instant t, we can evidently assign different values to $\dot{z}$, thus obtaining various different mechanical states. In this sense, z and $\dot{z}$ at each instant are independent variables necessary to specify the mechanical state of the system at that instant. The Lagrangian is a function of the instantaneous mechanical state of the system, while the action depends on the entire path taken by the system between t_1 and t_2.

Applying the product rule of differentiation, we note that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{z}}\delta z\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{z}}\right)\delta z + \frac{\partial L}{\partial \dot{z}}\dot{\delta} z\,. \tag{1.68}$$

Integrating both sides of the equation with respect to t, we find

$$\left[\frac{\partial L}{\partial \dot{z}}\delta z\right]_{t_1}^{t_2} = \int_{t_1}^{t_2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{z}}\right)\delta z\mathrm{d}t + \int_{t_1}^{t_2}\frac{\partial L}{\partial \dot{z}}\dot{\delta} z\mathrm{d}t\,, \tag{1.69}$$

which is just an integration by parts of the integrand $(\partial L/\partial \dot{z})\dot{\delta} z$. The left-hand side of this equation is identically zero because of (1.62). Thus, we may rewrite (1.67) as

$$\mathscr{S}[z+\delta z] - \mathscr{S}[z] = \int_{t_1}^{t_2}\left[\frac{\partial L}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{z}}\right)\right]\delta z\mathrm{d}t + \text{h.o.} \tag{1.70}$$

By definition, the variation of $\mathscr{S}$ up to the first order of δz refers to the quantity

$$\delta\mathscr{S} := \int_{t_1}^{t_2}\left[\frac{\partial L}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{z}}\right)\right]\delta z\mathrm{d}t\,. \tag{1.71}$$

Hamilton's principle demands that this expression vanish for *any* δz. Hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{z}}\right) - \frac{\partial L}{\partial z} = 0 \quad \text{for all } t_1 \le t \le t_2\,. \tag{1.72}$$

That (1.72) is sufficient in order for (1.71) to vanish should be obvious. To see that (1.72) is also necessary, suppose that (1.72) did not hold at some instant t^*. This means that the left-hand side of (1.72) is either positive or negative at t^*. Suppose that it is negative. Then, the integrand occurring in (1.71) will be positive at t^*. *Provided that the integrand is a continuous function of* t, it remains positive over some interval containing t^*. Then, we could fine-tune δz over this interval and make (1.71) false. If we demand that $\delta\mathscr{S}$ vanish for *any* δz, nonzero value of the integrand in (1.71) cannot be allowed. Equation (1.72) is called **Lagrange's equation of motion**.

Exercise 1.4. Show that, for the Lagrangian given by (1.58), (1.72) reduces to Newton's equation of motion. ///

1.5.2 Generalized Coordinates

In deriving Lagrange's equation of motion, no use was made of the fact that z is one of the Cartesian coordinates. All what we required was that the configuration of the system can be specified by z and that L is a (twice) differentiable function of z and $\dot{z}$. Any such variable will do.

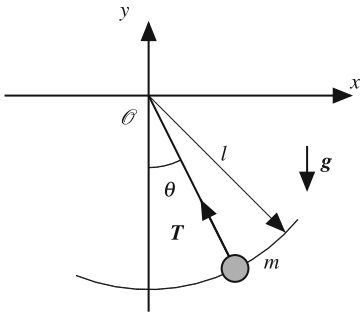

Fig. 1.7 A pendulum oscillating in the xy-plane. The length l of the rod is constant.

As an example, consider a simple pendulum in Fig. 1.7, which consists of a massless rigid rod of length l and a particle of mass m attached at the end of the rod. For simplicity, we assume that the particle is confined to the xy-plane. Once we specify θ, the coordinates (x, y) of the particle are completely determined:

$$x = l \sin\theta \quad \text{and} \quad y = -l \cos\theta \,. \tag{1.73}$$

Thus, we know the time evolution of the pendulum once we find θ as a function of t.

How do we find $\theta(t)$ then? If we can somehow express L as a function of θ and $\dot{\theta}$, we can use this $L(\theta, \dot{\theta})$ in (1.59) to find the action:

$$\mathscr{S}[\theta] := \int_{t_1}^{t_2} L(\theta, \dot{\theta}) \mathrm{d}t \,. \tag{1.74}$$

But, this is just (1.59) with z replaced by θ. So, by demanding that $\delta\mathscr{S} = 0$ with respect to arbitrary perturbation $\delta\theta$, we should arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = 0 \,, \tag{1.75}$$

which is an equation for $\theta(t)$.

We observe that the *form* of Lagrange's equation of motion as given by (1.72) remains unchanged by going from the Cartesian coordinate system to any other coordinate system. It should be carefully noted, however, that the functional *form* of L and hence the resulting equation of motion, in which partial derivatives of L are carried out explicitly, do depend on the variable used.

Example 1.3. Simple two-dimensional pendulum: To derive the differential equation for $\theta(t)$ from (1.75) in the case of the simple pendulum, we need to find the expressions for the kinetic and the potential energies of the system

first. From (1.73), we see that

$$\dot{x} = l\dot{\theta}\cos\theta \quad \text{and} \quad \dot{y} = l\dot{\theta}\sin\theta\,. \tag{1.76}$$

So, the kinetic energy of the particle is given by

$$\frac{1}{2}m\left(\dot{x}^2+\dot{y}^2\right) = \frac{1}{2}ml^2\dot{\theta}^2\,, \tag{1.77}$$

while the potential energy is

$$mgy = -mgl\cos\theta\,, \tag{1.78}$$

and hence

$$L = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta\,. \tag{1.79}$$

Taking the required partial derivatives, we find

$$\frac{\partial L}{\partial\dot{\theta}} = ml^2\dot{\theta} \quad \text{and} \quad \frac{\partial L}{\partial\theta} = -mgl\sin\theta\,. \tag{1.80}$$

Substituting these expressions in (1.75), we arrive at

$$ml^2\ddot{\theta} + mgl\sin\theta = 0\,. \tag{1.81}$$

Thus,

$$\ddot{\theta} = -\frac{g}{l}\sin\theta\,. \tag{1.82}$$

When θ is small, $\sin\theta \approx \theta$ and hence

$$\ddot{\theta} = -\frac{g}{l}\theta\,, \tag{1.83}$$

which is the equation of motion of a harmonic oscillator with the general solution given by

$$\theta(t) = c_1\cos\omega t + c_2\sin\omega t\,, \tag{1.84}$$

where $\omega = \sqrt{g/l}$. Under the initial conditions that $\theta = \theta_0$ and $\dot{\theta} = 0$ at $t = 0$, we obtain $c_1 = \theta_0$ and $c_2 = 0$.

One advantage of formulating mechanics in terms of Lagrange's equation of motion rather than Newton's is based on the fact that the choice of coordinates is arbitrary as long as they can uniquely specify the configuration of the system. To see why this is a good thing, let us try to analyze the same problem using Newton's equation of motion.

Since the pendulum is confined to the xy-plane, we need the x- and y-components of the equation:

$$m\ddot{x} = -T\sin\theta \quad \text{and} \quad m\ddot{y} = T\cos\theta - mg\,, \tag{1.85}$$

the solution of which must satisfy the constraint that the length l of the pendulum does not change:

$$x^2 + y^2 = l^2 = \text{const.} \tag{1.86}$$

In (1.85), $T = \|\boldsymbol{T}\|$ is the length of the unknown force vector $\boldsymbol{T}$ required to enforce the constraint. Thus, $\boldsymbol{T}$ must be found as a part of the solution. In contrast, $\boldsymbol{T}$ simply does not arise in Lagrange's equation of motion, because the constraint was automatically taken care of by choosing θ as a single variable to specify the configuration of the system. That is, x and y given by (1.73) automatically satisfy (1.86).

It is also worth pointing out that, in Newtonian mechanics, we work directly with vectors such as force, position, and velocity. In contrast, we deal primarily with scalars, such as kinetic and potential energies and Lagrangian in Lagrangian mechanics. This often leads to a simpler treatment of the same problem.

A variable needed to specify the configuration of the system, such as θ, is called a **generalized coordinate**. We refer to its time derivative, $\dot{\theta}$ here, as a **generalized velocity**.

Exercise 1.5. Show that (1.85) reduces to (1.82). ///

1.5.3 Many Mechanical Degrees of Freedom

The number of generalized coordinates necessary to specify the configuration of a system is called the number of **mechanical degrees of freedom**. In the above example, it was one. A **mechanical state** of the system having f mechanical degrees of freedom is specified by f generalized coordinates, $q_1, \ldots, q_f$, and f generalized velocities, $\dot{q}_1, \ldots, \dot{q}_f$. Accordingly, L is a function of these $2f$ variables and the action is given by

$$\mathscr{S}[q_1,\ldots,q_f] = \int_{t_1}^{t_2} L(q_1,\ldots,q_f,\dot{q}_1,\ldots,\dot{q}_f,t)\mathrm{d}t\,, \tag{1.87}$$

where we also allowed for the possibility that L changes even if $q_1, \ldots, q_f$ and $\dot{q}_1, \ldots, \dot{q}_f$ remained constant. Such time dependence is said to be *explicit*. In contrast, the time dependence of L through its dependence on $q_1, \ldots, q_f$ and $\dot{q}_1, \ldots, \dot{q}_f$ is said to be *implicit*. The **explicit time dependence** arises if, for example, the external field such as an electric field varies with time either according to some prescription or stochastically but in a manner that is *independent* of the mechanical state of the system. We exclude a field that changes with time in a manner that depends on the state of the system. In this latter case, the field and the system taken together must be studied as a single mechanical system.

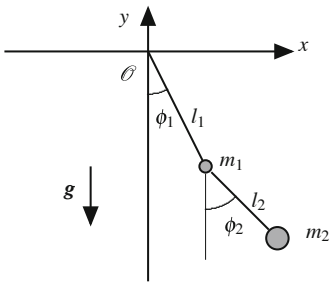

Fig. 1.8 A coplanar double pendulum oscillating in the *xy*-plane. Both l_1 and l_2 are constant.

The explicit time dependence of L does not play any role when considering the variations of $q_1, \ldots, q_f$ at each instant of time between given t_1 and t_2 because the field is common to both actual and varied paths.

When applying the stationarity condition of $\mathscr{S}$, that is, $\delta\mathscr{S} = 0$, the variations of $q_1, \ldots, q_f$ are entirely arbitrary except that the variations must vanish at $t = t_1$ and t_2. As one such variation, we may consider a variation, in which only δq_i is nonzero and $\delta q_j \equiv 0$ for $j \neq i$. Even for this rather special variation, Hamilton's principle demands that $\delta\mathscr{S} = 0$. This gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0\,. \tag{1.88}$$

Letting i vary from 1 to f, we arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0\,, \quad i = 1, \ldots, f\,. \tag{1.89}$$

What we have shown is that (1.89) is necessary in order for $\delta\mathscr{S}$ to vanish with respect to arbitrary variations of $q_1, \ldots, q_f$. Equation (1.89) is also sufficient for $\delta\mathscr{S} = 0$. (Take a moment to think about this.)

In arriving at (1.89), we made use of the fact that q_i's are all capable of independent variations. If the number of q_i's exceeds f, they are not independent. Thus, by demanding that $\delta q_j \equiv 0$ for $j \neq i$, we may be imposing a constraint on the possible values of δq_i. In this case, Hamilton's principle does not necessarily lead to (1.89).

Example 1.4. Coplanar double pendulum: Let us find a Lagrangian of a coplanar double pendulum placed in a uniform gravitational field as depicted in Fig. 1.8.

The configuration of the system is uniquely determined if we specify the position of the particles (x_1, y_1) and (x_2, y_2). In terms of these variables, we have

$$L = \frac{1}{2}m_1(\dot{x_1}^2 + \dot{y_1}^2) - m_1 g y_1 + \frac{1}{2}m_2(\dot{x_2}^2 + \dot{y_2}^2) - m_2 g y_2 \ . \tag{1.90}$$

However, since the particles are attached to the rods, x_1, y_1, x_2, and y_2 are not independent and we cannot derive Lagrange's equations of motion directly from the Lagrangian given by (1.90).

Instead, we note that x_1, y_1, x_2, and y_2 are completely determined by specifying two variables ϕ_1 and ϕ_2. In other words, the mechanical degrees of freedom of the system is two and we should be able to find two Lagrange's equations of motion from a Lagrangian expressed in terms of ϕ_1, ϕ_2, $\dot{\phi}_1$, and $\dot{\phi}_2$. From Fig. 1.8, we note the following relations:

$$\begin{aligned} x_1 &= l_1 \sin\phi_1 \ , \\ y_1 &= -l_1 \cos\phi_1 \ , \\ x_2 &= l_1 \sin\phi_1 + l_2 \sin\phi_2 \ , \\ y_2 &= -l_1 \cos\phi_1 - l_2 \cos\phi_2 \ . \end{aligned} \tag{1.91}$$

Taking the time derivative,

$$\begin{aligned} \dot{x}_1 &= l_1 \dot{\phi}_1 \cos\phi_1 \ , \\ \dot{y}_1 &= l_1 \dot{\phi}_1 \sin\phi_1 \ , \\ \dot{x}_2 &= l_1 \dot{\phi}_1 \cos\phi_1 + l_2 \dot{\phi}_2 \cos\phi_2 \ , \\ \dot{y}_2 &= l_1 \dot{\phi}_1 \sin\phi_1 + l_2 \dot{\phi}_2 \sin\phi_2 \ . \end{aligned} \tag{1.92}$$

Using these relations in (1.90), we find

$$\begin{aligned} L = {} & \frac{1}{2}(m_1 + m_2) l_1^2 \dot{\phi}_1^2 + \frac{1}{2} m_2 l_2^2 \dot{\phi}_2^2 + m_2 l_1 l_2 \dot{\phi}_1 \dot{\phi}_2 \cos(\phi_1 - \phi_2) \\ & + (m_1 + m_2) g l_1 \cos\phi_1 + m_2 g l_2 \cos\phi_2 \ , \end{aligned} \tag{1.93}$$

where we used the identity:

$$\cos(\alpha + \beta) = \cos\alpha \cos\beta - \sin\alpha \sin\beta \ . \tag{1.94}$$

This L can be used in (1.89) to find Lagrange's equations of motion for $\phi_1(t)$ and $\phi_2(t)$.

Exercise 1.6. Consider a pendulum consisting of a massless harmonic spring and a particle of mass m attached at the end of the spring as shown in Fig. 1.9. The spring constant and the natural length of the spring are k and l_0, respectively.

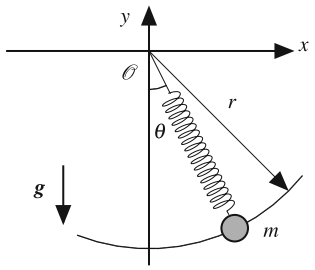

Fig. 1.9 A pendulum suspended by a harmonic spring.

a. Let $\boldsymbol{e}_r$ denote a unit vector pointing from the origin to the particle. Whether we are stretching the spring ($r > l_0$) or compressing it ($r < l_0$), the force exerted by the spring on the particle is given by

$$\boldsymbol{F} = -k(r - l_0)\boldsymbol{e}_r \,. \tag{1.95}$$

(Draw a diagram to convince yourself of this point.) Use this expression to show that the potential energy of the harmonic spring is given by

$$\psi(r) = \frac{1}{2}k(r - l_0)^2 \,. \tag{1.96}$$

b. Using θ and r as your generalized coordinates, find the Lagrangian of the system. You may assume that the motion of the spring and the particle is confined to the xy-plane. ///

Exercise 1.7. A system consists of two particles, one at $\boldsymbol{r}_1$ and the other at $\boldsymbol{r}_2$. They interact via a potential energy ϕ which depends only on their relative position $\boldsymbol{r} := \boldsymbol{r}_1 - \boldsymbol{r}_2$. Using the position vector $\boldsymbol{R}$ of the center of mass

$$\boldsymbol{R} = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{m_1 + m_2} \tag{1.97}$$

and $\boldsymbol{r}$ as your generalized coordinates, show that

$$L = \frac{1}{2}M\dot{R}^2 + \frac{1}{2}\mu\dot{r}^2 - \phi(\boldsymbol{r}) \,, \tag{1.98}$$

where $M := m_1 + m_2$, $\dot{R}^2 := \|\dot{\boldsymbol{R}}\|^2 = \dot{\boldsymbol{R}} \cdot \dot{\boldsymbol{R}}$, and likewise for $\dot{r}^2$. The quantity

$$\mu := \frac{m_1 m_2}{m_1 + m_2} \tag{1.99}$$

is the so-called **reduced mass**. Why is it more convenient to work with $\boldsymbol{R}$ and $\boldsymbol{r}$ than with $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$? ///

Given a mechanical system, its Lagrangian is determined only to an additive function $\mathrm{d}F/\mathrm{d}t$. In fact, if L_1 denotes a Lagrangian of a mechanical system with f degrees of freedom, the new function

$$L_2 := L_1 + \frac{\mathrm{d}F(q_1,\ldots,q_f,t)}{\mathrm{d}t} \tag{1.100}$$

also serves as a Lagrangian of the same system.

The validity of this claim is immediately obvious if we integrate (1.100) with respect to time. In fact, if

$$\mathscr{S}_i[q_1,\ldots,q_f] := \int_{t_1}^{t_2} L_i(q_1,\ldots,q_f,\dot{q}_1,\ldots,\dot{q}_f,t)\mathrm{d}t\ , \tag{1.101}$$

then

$$\begin{aligned}\mathscr{S}_2[q_1,\ldots,q_f] = \mathscr{S}_1[q_1,\ldots,q_f] + F(q_1(t_2),\ldots,q_f(t_2),t_2)\\ -F(q_1(t_1),\ldots,q_f(t_1),t_1)\ .\end{aligned} \tag{1.102}$$

Since $\delta q_i(t_1) = \delta q_i(t_2) = 0$ for all $i = 1, \ldots, f$, we see that $\delta\mathscr{S}_2 = \delta\mathscr{S}_1$. Thus, if $\delta\mathscr{S}_1 = 0$, then $\delta\mathscr{S}_2 = 0$ for *the same set of functions* $q_1(t), \ldots, q_f(t)$, which is then the solution of Lagrange's equations of motion whether they are derived from L_1 or L_2. Alternatively, you can validate the claim through a more explicit computation of various derivatives as illustrated in the next exercise.

Exercise 1.8. Use (1.100) to compute

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_2}{\partial\dot{q}_i}\right) - \frac{\partial L_2}{\partial q_i}$$

and show that it is zero if and only if

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_1}{\partial\dot{q}_i}\right) - \frac{\partial L_1}{\partial q_i} = 0\,.$$

This being the case for $i = 1,\ldots,f$, the same set of functions $q_1(t), \ldots, q_f(t)$ found by solving Lagrange's equations of motion obtained from L_1 also satisfies those derived from L_2. ///

1.6 Momentum and Energy: Definitions

The quantity

$$p_i := \frac{\partial L}{\partial\dot{q}_i} \tag{1.103}$$

is called the **generalized momentum** conjugate to the generalized coordinate q_i. When q_i is one of the Cartesian coordinates, then its conjugate momentum is called the **linear momentum**. Thus, the linear momentum of the ith particle in the system is given by

$$\boldsymbol{p}_i = \frac{\partial L}{\partial \dot{\boldsymbol{r}}_i} = \frac{\partial L}{\partial \boldsymbol{v}_i} \,. \tag{1.104}$$

We recall from Sect. 1.2.3 that this is just a compact way of writing

$$p_x = \frac{\partial L}{\partial v_x} \,, \quad p_y = \frac{\partial L}{\partial v_y} \,, \quad \text{and} \quad p_z = \frac{\partial L}{\partial v_z} \,, \tag{1.105}$$

where we omitted the subscript i for brevity. We define the **energy function**

$$h := \sum_{i=1}^{f} \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \,. \tag{1.106}$$

As seen in the following example, h is often equal to the mechanical energy, that is, the sum of the kinetic and the potential energies.

Example 1.5. Energy function of a particle: Consider a particle of mass m moving in an external field ψ. Using a Cartesian coordinate system, we have

$$L = \frac{1}{2} m v^2 - \psi(\boldsymbol{r}) \,. \tag{1.107}$$

From (1.104),

$$\boldsymbol{p} = \frac{\partial L}{\partial \boldsymbol{v}} = m\boldsymbol{v} \,. \tag{1.108}$$

According to (1.106),

$$h = \boldsymbol{p} \cdot \boldsymbol{v} - L = m v^2 - \frac{1}{2} m v^2 + \psi(\boldsymbol{r}) = \frac{1}{2} m v^2 + \psi(\boldsymbol{r}) \,. \tag{1.109}$$

At least in this example, therefore, the energy function h is nothing but the mechanical energy defined by (1.40).

If you prefer, you can work directly with the components of various vectors. Thus, we start by writing (1.107) as

$$L = \frac{1}{2} m \left({v_x}^2 + {v_y}^2 + {v_z}^2 \right) - \psi(\boldsymbol{r}) \,. \tag{1.110}$$

From (1.105),

$$p_x = \frac{\partial L}{\partial v_x} = m v_x \,, \quad p_y = \frac{\partial L}{\partial v_y} = m v_y \,, \quad \text{and} \quad p_z = \frac{\partial L}{\partial v_z} = m v_z \,, \tag{1.111}$$

which is just (1.108). According to (1.106),

$$\begin{aligned} h &= (p_x v_x + p_y v_y + p_z v_z) - L \\ &= m{v_x}^2 + m{v_y}^2 + m{v_z}^2 - \frac{1}{2}m\left({v_x}^2 + {v_y}^2 + {v_z}^2\right) + \psi(\boldsymbol{r}) \\ &= \frac{1}{2}m\left({v_x}^2 + {v_y}^2 + {v_z}^2\right) + \psi(\boldsymbol{r})\,. \end{aligned} \tag{1.112}$$

So, is h always equal to E? Disappointingly, the answer is no. As an example, consider the following exercise.

Exercise 1.9. For L_1 and L_2 in (1.100), show that

$$h_2 = h_1 - \frac{\partial F}{\partial t}\,. \tag{1.113}$$

That is, if $h_1 = E$, then $h_2 \neq E$ unless $\partial F/\partial t \equiv 0$. ///

In the next section, marked with a dagger (†) for an optional reading, we examine the condition under which $h = E$. The conclusion is that, for a broad range of applications of classical mechanics we are going to encounter, $h = E$ holds. Thus, unless stated otherwise, we shall assume this equality and speak only of the energy E.

To conclude this section, let us deduce one immediate consequence of (1.106). We consider two mechanical systems, A and B, and suppose that they are sufficiently far away from each other that they evolve independently. However, there is nothing in principle that prevents us from treating them as a single composite system. It is clear that Lagrange's equations of motion for the composite system can be derived from the Lagrangian $L := L_a + L_b$, in which L_a depends only on q_i's and $\dot{q}_i$'s pertaining to system A and does not depend on those pertaining to system B. Similarly for L_b. This means that Lagrangian is an additive quantity. From (1.106), it follows that energy also is an additive quantity.

1.7 †Energy Function and Energy

Consider a collection of N particles subject to an external field ψ. Using a Cartesian coordinate system, we have

$$E = \sum_{i=1}^{N}\left[\frac{1}{2}m_i{v_i}^2 + \psi(\boldsymbol{r}_i, t)\right] + \phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)\,, \tag{1.114}$$

where we allowed for the explicit time dependence of ψ. As in Exercise 1.3, ϕ is the potential energy due to the mutual interaction among particles.

Suppose that we introduced a set of generalized coordinates $q_1,\ldots,q_f$ that are related to $\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N$ by the following set of equations:

$$\boldsymbol{r}_i = \boldsymbol{r}_i(q_1,\ldots,q_f,t)\,, \quad i=1,\ldots,N\,. \tag{1.115}$$

If there is no constraint on $\boldsymbol{r}_i$'s, then $f=DN$, where D is the dimensionality of the space in which the particles move. Otherwise, $f<DN$. In (1.73), for example, $f=1$, $D=2$, and $N=1$. Following the approach we took in Example 1.3, we proceed to express L in terms of q_i's and $\dot{q}_i$'s. First, we note that

$$\boldsymbol{v}_i = \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} = \sum_{j=1}^{f}\frac{\partial \boldsymbol{r}_i}{\partial q_j}\dot{q}_j + \frac{\partial \boldsymbol{r}_i}{\partial t}\,. \tag{1.116}$$

Thus,

$$v_i{}^2 = \boldsymbol{v}_i\cdot\boldsymbol{v}_i = \sum_{j=1}^{f}\sum_{k=1}^{f}A_i^{jk}\dot{q}_j\dot{q}_k + 2\sum_{j=1}^{f}B_i^j\dot{q}_j + C_i\,, \tag{1.117}$$

where we defined

$$A_i^{jk} := \frac{\partial \boldsymbol{r}_i}{\partial q_j}\cdot\frac{\partial \boldsymbol{r}_i}{\partial q_k}\,, \quad B_i^j := \frac{\partial \boldsymbol{r}_i}{\partial q_j}\cdot\frac{\partial \boldsymbol{r}_i}{\partial t}\,, \quad \text{and} \quad C_i := \frac{\partial \boldsymbol{r}_i}{\partial t}\cdot\frac{\partial \boldsymbol{r}_i}{\partial t}\,. \tag{1.118}$$

Using (1.117), we see that the kinetic energy of the system is given by

$$\frac{1}{2}\sum_{i=1}^{N}m_i v_i{}^2 = \frac{1}{2}\sum_{i=1}^{N}m_i\left[\sum_{j=1}^{f}\sum_{k=1}^{f}A_i^{jk}\dot{q}_j\dot{q}_k + 2\sum_{j=1}^{f}B_i^j\dot{q}_j + C_i\right]\,. \tag{1.119}$$

Because of (1.115), $\boldsymbol{r}_i$'s in ψ and ϕ are now functions of $q_1,\ldots,q_f$ and t. Thus, the total potential energy may be written as

$$\sum_{i=1}^{N}\psi(\boldsymbol{r}_i,t) + \phi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N) =: U(q_1,\ldots,q_f,t)\,. \tag{1.120}$$

So, the Lagrangian, when expressed in terms of $q_1,\ldots,q_f$ is

$$L = \frac{1}{2}\sum_{i=1}^{N}m_i\left[\sum_{j=1}^{f}\sum_{k=1}^{f}A_i^{jk}\dot{q}_j\dot{q}_k + 2\sum_{j=1}^{f}B_i^j\dot{q}_j + C_i\right] - U(q_1,\ldots,q_f,t)\,. \tag{1.121}$$

It follows that

$$\frac{\partial L}{\partial \dot{q}_n} = \sum_{i=1}^{N}m_i\left[\sum_{j=1}^{f}A_i^{jn}\dot{q}_j + B_i^n\right]\,, \tag{1.122}$$

where we used the fact that $A_i^{jk}=A_i^{kj}$. (Take a moment to convince yourself of (1.122). You can set $N=1$ and $f=2$, for example, and compute $\partial L/\partial\dot{q}_1$ explicitly. This should give you some idea about the essential steps involved.) Using (1.106),

we arrive at

$$h = \sum_{n=1}^{f} \frac{\partial L}{\partial \dot{q}_n} \dot{q}_n - L = \sum_{i=1}^{N} m_i \left[\sum_{j=1}^{f} \sum_{n=1}^{f} A_i^{jn} \dot{q}_j \dot{q}_n + \sum_{n=1}^{f} B_i^n \dot{q}_n \right] - L$$
$$= \frac{1}{2} \sum_{i=1}^{N} m_i \left[\sum_{j=1}^{f} \sum_{k=1}^{f} A_i^{jk} \dot{q}_j \dot{q}_k - C_i \right] + U(q_1, \ldots, q_f, t) \,. \tag{1.123}$$

But, according to (1.119) and (1.120), the total mechanical energy of the system is

$$E = \frac{1}{2} \sum_{i=1}^{N} m_i \left[\sum_{j=1}^{f} \sum_{k=1}^{f} A_i^{jk} \dot{q}_j \dot{q}_k + 2 \sum_{j=1}^{f} B_i^j \dot{q}_j + C_i \right] + U(q_1, \ldots, q_f, t) \,. \tag{1.124}$$

Evidently, *if* $B_i^j \equiv 0$ and $C_i \equiv 0$ for all $i = 1, \ldots, N$, then $h = E$. In view of (1.118), this implies that $\boldsymbol{r}_i$'s do not depend *explicitly* on t. So, if the relation between $\boldsymbol{r}_i$'s and q_i's is of the form of

$$\boldsymbol{r}_i = \boldsymbol{r}_i(q_1, \ldots, q_f) \,, \quad i = 1, \ldots, N \tag{1.125}$$

in place of (1.115), then the energy function coincides with the mechanical energy of the system.

1.8 Conservation Laws and Symmetry

As a mechanical system evolves according to the equations of motion, its generalized coordinates and velocities change. If we construct a function of these variables, its value will, in general, change with them. However, the value of a certain combination of the generalized coordinates and velocities may remain constant. Such a quantity is called a **constant of motion** or a **conserved quantity**. We have already encountered one such quantity, that is, energy, in Sects. 1.2.5 and 1.3.1. Constants of motions are of great interest since they allow us to extract certain information about the mechanical behavior of the system without actually solving the equations of motion.

In general, the Lagrangian of a mechanical system is a function of generalized coordinates and generalized velocities. It may also depend explicitly on time. In certain problems, however, it may so happen that L remains unchanged when these variables undergo certain changes. As we shall see below, this, in turn, points to an existence of some conserved quantity.

Let us start by looking at a somewhat trivial example. We suppose that L is independent of q_i. Then,

$$\frac{\partial L}{\partial q_i} \equiv 0 \,. \tag{1.126}$$

From Lagrange's equation of motion,

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} \equiv 0\,. \tag{1.127}$$

Thus, if L is independent of q_i, then, its conjugate momentum p_i is a constant of motion. Now you are ready to look at something a little more profound.

1.8.1 Conservation of Linear Momentum

Let us consider a mechanical system A consisting of N particles. In the absence of any constraint that reduces the mechanical degrees of freedom, its Lagrangian is given by

$$L^A = L(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N,\dot{\boldsymbol{r}}_1,\ldots,\dot{\boldsymbol{r}}_N,t)\,, \tag{1.128}$$

Given this system A, imagine that we created its *identical* copy B, and placed system B some distance away from system A. These systems are then related by a linear translation, say by $\Delta\boldsymbol{r}$. Thus, the position $\boldsymbol{r}_i'$ (with respect to the same origin as system A) of the ith particle in system B is given by

$$\boldsymbol{r}_i' = \boldsymbol{r}_i + \Delta\boldsymbol{r}\,, \quad i = 1,\ldots,N\,. \tag{1.129}$$

By "identical copy," we also mean that

$$\dot{\boldsymbol{r}}_i' = \dot{\boldsymbol{r}}_i\,, \quad i = 1,\ldots,N\,, \tag{1.130}$$

that is, $\dot{\boldsymbol{r}}_i$'s remain unaffected by the linear translation. Now the question is this: What is the relationship between the Lagrangians L^A and L^B of these two systems? Here, it should be clearly noted that the comparison is between L^A when system A alone is present and L^B when system B alone is present. Thus, there is no interaction between A and B.

Since these two systems are identical in all respect except for its position in space, we have no reason to evaluate L^B any differently from L^A. That is, we use the same function L to evaluate L^A and L^B, the only difference being in the values assumed by its arguments:

$$L^B = L(\boldsymbol{r}_1+\Delta\boldsymbol{r},\ldots,\boldsymbol{r}_N+\Delta\boldsymbol{r},\dot{\boldsymbol{r}}_1,\ldots,\dot{\boldsymbol{r}}_N,t)\,. \tag{1.131}$$

If these two systems are far a way from any source of external field, there will be nothing that can distinguish the region occupied by system A and that by system B. This is referred to as the **homogeneity of space**. Less formally, one may say that "here" is as good as any other place. Therefore, we should expect that $L^B = L^A$. That is, the value of L remains unaffected by a linear translation by $\Delta\boldsymbol{r}$. Such an L is said to be *translationally invariant*.

As an example, we recall the system considered in Sect. 1.3.1, which has a translationally invariant L if $\psi \equiv 0$, that is, in the absence of an external field. Clearly, ψ remains zero even after a displacement by $\Delta \boldsymbol{r}$ of the mechanical system. The kinetic energy of the system depends only on mass and velocity of each particle. So, it remains unaffected as well. Likewise for the potential energy ϕ due to interaction among particles, which depends only on their relative position. Thus, $L^A = L^B$.

We now explore the consequence of the translational invariance of L. It is sufficient to restrict ourselves to an infinitesimal translation and replace $\Delta \boldsymbol{r}$ by $\delta \boldsymbol{r}$. From (1.128) and (1.131), we find

$$L^B - L^A = \sum_{i=1}^{N} \left(\frac{\partial L}{\partial \boldsymbol{r}_i} \cdot \delta \boldsymbol{r} \right) + \text{h.o.}, \tag{1.132}$$

where we recall that

$$\frac{\partial L}{\partial \boldsymbol{r}_i} \doteq \left(\frac{\partial L}{\partial x_i}, \frac{\partial L}{\partial y_i}, \frac{\partial L}{\partial z_i} \right) \quad \text{and} \quad \delta \boldsymbol{r} \doteq (\delta x, \delta y, \delta z) , \tag{1.133}$$

and hence

$$\frac{\partial L}{\partial \boldsymbol{r}_i} \cdot \delta \boldsymbol{r} = \frac{\partial L}{\partial x_i} \delta x + \frac{\partial L}{\partial y_i} \delta y + \frac{\partial L}{\partial z_i} \delta z . \tag{1.134}$$

From Lagrange's equations of motion and (1.104), we have

$$\frac{\partial L}{\partial \boldsymbol{r}_i} = \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{\boldsymbol{r}}_i} \right) = \frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t} . \tag{1.135}$$

Thus, the first-order term of (1.132), which we denote by δL, is given by

$$\delta L := \sum_{i=1}^{N} \left(\frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t} \cdot \delta \boldsymbol{r} \right) = \delta \boldsymbol{r} \cdot \sum_{i=1}^{N} \frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t} = \delta \boldsymbol{r} \cdot \frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_{i=1}^{N} \boldsymbol{p}_i \right) , \tag{1.136}$$

where we pulled $\delta \boldsymbol{r}$ out of the summation since it is common to all particles and hence is independent of i.

If L is translationally invariant, δL must be identically zero for *any* $\delta \boldsymbol{r}$, from which we deduce

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_{i=1}^{N} \boldsymbol{p}_i \right) \equiv \boldsymbol{0} . \tag{1.137}$$

It follows that the **total linear momentum** defined by

$$\boldsymbol{P} := \sum_{i=1}^{N} \boldsymbol{p}_i \tag{1.138}$$

is a constant of motion. In this way, the conservation of the total linear momentum is seen to be a consequence of the translational invariance of L, and hence of the homogeneity of space.

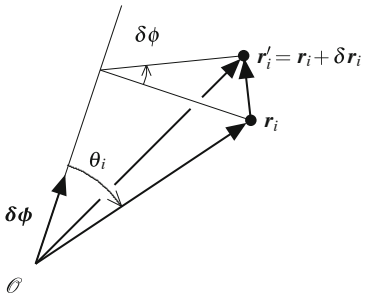

Fig. 1.10 Change in $\boldsymbol{r}_i$ upon rotation.

It may be that L is invariant with respect to translation in a direction parallel to the xy-plane, but not in the z-direction. This will be the case, for example, for a mechanical system moving in a uniform gravitational field if the z-axis is chosen to align with the gravitational acceleration $\boldsymbol{g}$. By considering the displacement $\boldsymbol{\delta r}$ that is parallel to the xy-plane, we observe that the x- and y- components of $\boldsymbol{P}$ are still constants of motion in this case. The z-component, however, is not. You can easily confirm this by allowing an object to free fall toward the floor: P_z of the object will continue to change with time.

1.8.2 ‡Conservation of Angular Momentum

Suppose that system B is related to system A by an infinitesimal rotation around some axis by $\delta\phi$. Again we expect that $L^A = L^B$. This is a consequence of the **isotropy of space**, that is, all directions in space are equivalent.

We need to relate $\boldsymbol{r}_i$'s and $\dot{\boldsymbol{r}}_i$'s of system B to those of system A. From Fig. 1.10, we see that

$$\|\delta\boldsymbol{r}_i\| = r_i \sin\theta_i \delta\phi \ , \tag{1.139}$$

where θ_i is the angle between the axis of rotation and $\boldsymbol{r}_i$. Let us denote by $\boldsymbol{\delta\phi}$ the vector of length $\delta\phi$ that points toward the direction a right-hand screw advances upon rotation by $\delta\phi$. Note that $\boldsymbol{\delta\phi}$ is *not* a variation of some vector $\boldsymbol{\phi}$. Rather, it is a vector of infinitesimal length $\|\boldsymbol{\delta\phi}\|$, where the "length" is an angle measured in radian in this case. Since $\delta\boldsymbol{r}_i$ is perpendicular to the plane including $\boldsymbol{r}_i$ and $\boldsymbol{\delta\phi}$, we see that $\delta\boldsymbol{r}_i$ is parallel to the vector $\boldsymbol{\delta\phi} \times \boldsymbol{r}_i$, where "$\times$" denotes the cross product of two vectors. (See Appendix A.6 for a brief review.) Combining this observation with (1.139), we arrive at

$$\delta\boldsymbol{r}_i = \boldsymbol{\delta\phi} \times \boldsymbol{r}_i \ , \tag{1.140}$$

The same argument applies for $\dot{\boldsymbol{r}}_i$ as well, leading to

$$\delta\dot{\boldsymbol{r}}_i = \boldsymbol{\delta\phi} \times \dot{\boldsymbol{r}}_i \ . \tag{1.141}$$

Thus, to the first order of $\delta\phi$, $L^B - L^A$ is given by

$$\delta L = \sum_{i=1}^{N} \left[\frac{\partial L}{\partial \boldsymbol{r}_i} \cdot \delta \boldsymbol{r}_i + \frac{\partial L}{\partial \dot{\boldsymbol{r}}_i} \cdot \delta \dot{\boldsymbol{r}}_i \right] = \sum_{i=1}^{N} \left[\frac{\partial L}{\partial \boldsymbol{r}_i} \cdot \boldsymbol{\delta\phi} \times \boldsymbol{r}_i + \frac{\partial L}{\partial \dot{\boldsymbol{r}}_i} \cdot \boldsymbol{\delta\phi} \times \dot{\boldsymbol{r}}_i \right] . \quad (1.142)$$

Using (1.89) and (1.104),

$$\delta L = \sum_{i=1}^{N} \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{\boldsymbol{r}}_i} \right) \cdot \boldsymbol{\delta\phi} \times \boldsymbol{r}_i + \frac{\partial L}{\partial \dot{\boldsymbol{r}}_i} \cdot \boldsymbol{\delta\phi} \times \dot{\boldsymbol{r}}_i \right] = \sum_{i=1}^{N} \left(\dot{\boldsymbol{p}}_i \cdot \boldsymbol{\delta\phi} \times \boldsymbol{r}_i + \boldsymbol{p}_i \cdot \boldsymbol{\delta\phi} \times \dot{\boldsymbol{r}}_i \right) . \quad (1.143)$$

With the help of a vector identity (A.32), we may rewrite this as

$$\delta L = \boldsymbol{\delta\phi} \cdot \sum_{i=1}^{N} \left(\boldsymbol{r}_i \times \dot{\boldsymbol{p}}_i + \dot{\boldsymbol{r}}_i \times \boldsymbol{p}_i \right) = \boldsymbol{\delta\phi} \cdot \frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_{i=1}^{N} \boldsymbol{r}_i \times \boldsymbol{p}_i \right) . \quad (1.144)$$

Since $\delta L \equiv 0$ for any $\boldsymbol{\delta\phi}$, we conclude that the **total angular momentum** defined by

$$\boldsymbol{M} := \sum_{i=1}^{N} \boldsymbol{r}_i \times \boldsymbol{p}_i \quad (1.145)$$

is a constant of motion. We see that the conservation of total angular momentum follows from the *rotational invariance* of L, and hence from the isotropy of space.

Exercise 1.10. Motion of a particle subject to a central field, that is, one in which the potential energy depends only on the distance from some particular point in space, is confined to a plane containing that point. Why? ///

1.8.3 Conservation of Energy

Returning now to a general mechanical system with f mechanical degrees of freedom, we suppose that its L does not depend *explicitly* on time:

$$L = L(q_1, \ldots, q_f, \dot{q}_1, \ldots, \dot{q}_f) . \quad (1.146)$$

Since q_i and $\dot{q}_i$ are functions of t, L still depends on t. But such a time dependence is said to be *implicit*. Using the chain rule,

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \sum_{i=1}^{f} \frac{\partial L}{\partial q_i} \dot{q}_i + \sum_{i=1}^{f} \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i . \quad (1.147)$$

We can rewrite the first term on the right-hand side using Lagrange's equation of motion (1.89) and write

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \sum_{i=1}^{f} \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i + \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i \right] = \sum_{i=1}^{f} \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) , \tag{1.148}$$

from which we find

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_{i=1}^{f} \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \right) = 0 . \tag{1.149}$$

Recalling (1.106) and our assumption that $h = E$, we see that the energy is a conserved quantity.

What will happen if we allowed for an *explicit* time dependence of L? This will be the case if there is a time-dependent external field. Carrying out the same analysis while retaining $\partial L/\partial t$, we observe that (1.149) is replaced by

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\frac{\partial L}{\partial t} . \tag{1.150}$$

Clearly, E is a constant of motion if and *only if* $\partial L/\partial t \equiv 0$.

Exercise 1.11. Derive (1.150). ///

The procedure we used to deduce the law of conservation of energy admits the following curious interpretation. Suppose you conducted a certain experiment today and plan on repeating the *identical* experiment tomorrow. Within mechanics, this means that you start the experiment from the identical initial conditions as specified by the generalized coordinates and velocities. You should naturally expect to find the identical outcome.[6] So, you have no reason to change the form of the function L between today (t^A) and tomorrow (t^B). You would also expect that its value remains the same for an identical mechanical state. In other words, each instant of time is equivalent to any other. This is referred to as the **homogeneity of time**. We demand, therefore, that the relation

$$L(q_1, \ldots, q_f, \dot{q}_1, \ldots, \dot{q}_f, t^A) = L(q_1, \ldots, q_f, \dot{q}_1, \ldots, \dot{q}_f, t^B) , \tag{1.151}$$

hold for *any* t^A and t^B. It follows that

$$\frac{\partial L}{\partial t} \equiv 0 . \tag{1.152}$$

Equation (1.150) then indicates that E is a constant of motion. Energy is a constant of motion because "now" is as good as any other instant of time.

One might argue that the outcome of the two experiments should be identical even if there is a time-dependent external field. But, this is so only if timing of the two experiments is properly chosen to make sure that they are in sync with the time-dependent external field in an identical manner. In other words, (1.151) holds only for properly chosen pairs of t^A and t^B. From this, we *cannot* deduce (1.152).

Example 1.6. Symmetry and constants of motion: Suppose that a system consisting of many particles moves in an external field generated by fixed sources uniformly distributed on an infinite plane, say, the xy-plane (on which $z = 0$). In this case, L is invariant with respect to (1) translation in t, (2) translation parallel to the xy-plane, and, if you read Sect. 1.8.2, (3) rotation around any axis perpendicular to the xy-plane. So, the energy, the x- and y-components of the linear momentum, and the z-component of the angular momentum are constants of motion.

If the source of the field is confined to $x > 0$, the system is invariant only with respect to (1) translation in t and (2) translation along the y-axis. So, constants of motion are the energy and the y-component of the linear momentum.

1.9 Hamilton's Equations of Motion

Lagrange's equations of motion are the relationships among the partial derivatives of L with respect to its independent variables $q_1, \ldots, q_f$ and $\dot{q}_1, \ldots, \dot{q}_f$. When supplemented with initial conditions, these equations of motion tell us how $q_1, \ldots, q_f$ and $\dot{q}_1, \ldots, \dot{q}_f$ evolve with time. Thus, we may say that the function L, *if* expressed in terms of $q_1, \ldots, q_f, \dot{q}_1, \ldots, \dot{q}_f$, and t, contains within it the mechanical behavior of the system in its entirety.

Suppose now that, for whatever reason, we wanted to describe mechanics of the system using $p_1, \ldots, p_f$ instead of $\dot{q}_1, \ldots, \dot{q}_f$ as our independent variables. This may require a new function to replace L. We would certainly want our new approach to be equally useful as the Lagrangian-based approach. We can ensure this by demanding that this new function *contains the same information as the Lagrangian.* In other words, we demand not only that we can construct this function from the Lagrangian, but also that we can recover the original Lagrangian from it. That way, even in the worst case scenario in which equations of motion cannot be derived *directly* from the new function, we can at least derive Lagrange's equations of motion by first finding the Lagrangian from our new function.

Our goal then is to replace $\dot{q}_i$ by $p_i = \partial L/\partial \dot{q}_i$ while preserving the information content of L. There is a mathematical procedure called **Legendre transformation** designed just for this purpose. A brief introduction to this method is given in Appendix C, in which we present the Legendre transformation of a function of a single variable. In contrast, we are now interested in replacing multiple variables, $\dot{q}_1, \ldots, \dot{q}_f$, by $p_1, \ldots, p_f$. In the next section, we show in detail how this can be accomplished by replacing one variable at a time, thus performing the Legendre transformation f times. Here, we take for granted that the transformation described in Appendix C generalizes naturally to the case of simultaneous replacement of multiple variables and proceed as follows:

Step 1: Given the Lagrangian

$$L(q_1,\ldots,q_f,\dot{q}_1,\ldots,\dot{q}_f,t)\,, \tag{1.153}$$

calculate p_i by (1.103) for $i=1,\ldots,f$, thus finding $p_1,\ldots,p_f$ as functions of $q_1,\ldots,q_f$ and $\dot{q}_1,\ldots,\dot{q}_f$:

$$p_i = p_i(q_1,\ldots,q_f,\dot{q}_1,\ldots,\dot{q}_f)\,,\ i=1,\ldots,f\,. \tag{1.154}$$

Step 2: Solve these f equations for $\dot{q}_1,\ldots,\dot{q}_f$ to express them as functions of $q_1,\ldots,q_f$ and $p_1,\ldots,p_f$:

$$\dot{q}_i = \dot{q}_i(q_1,\ldots,q_f,p_1,\ldots,p_f)\,,\ i=1,\ldots,f\,. \tag{1.155}$$

Step 3: Express

$$H := \sum_{i=1}^{f} p_i\dot{q}_i - L \tag{1.156}$$

as a function of $q_1,\ldots,q_f$ and $p_1,\ldots,p_f$. This makes H the negative of the **Legendre transform** of L. The function

$$H(q_1,\ldots,q_f,p_1,\ldots,p_f,t) \tag{1.157}$$

is called the **Hamiltonian** of the system. From (1.103) and (1.106), we see that Hamiltonian is the energy (function) of the system expressed as a function of $q_1,\ldots,q_f$ and $p_1,\ldots,p_f$.

By construction, L and H contain the same information. But, before we can accept the Hamiltonian as a useful concept, we must be able to derive from H the equations of motion that relate $q_1,\ldots,q_f$ and $p_1,\ldots,p_f$ to their time derivatives. From (1.156),

$$\mathrm{d}H = \sum_{i=1}^{f}(p_i\mathrm{d}\dot{q}_i + \dot{q}_i\mathrm{d}p_i) - \mathrm{d}L\,. \tag{1.158}$$

From (1.153),

$$\mathrm{d}L = \sum_{i=1}^{f}\left(\frac{\partial L}{\partial q_i}\mathrm{d}q_i + \frac{\partial L}{\partial \dot{q}_i}\mathrm{d}\dot{q}_i\right) + \frac{\partial L}{\partial t}\mathrm{d}t = \sum_{i=1}^{f}(\dot{p}_i\mathrm{d}q_i + p_i\mathrm{d}\dot{q}_i) + \frac{\partial L}{\partial t}\mathrm{d}t\,, \tag{1.159}$$

where we used (1.89) and (1.103). Combining (1.158) and (1.159)

$$\mathrm{d}H = \sum_{i=1}^{f}(-\dot{p}_i\mathrm{d}q_i + \dot{q}_i\mathrm{d}p_i) - \frac{\partial L}{\partial t}\mathrm{d}t\,. \tag{1.160}$$

On the other hand, (1.157) indicates that

$$\mathrm{d}H = \sum_{i=1}^{f} \left(\frac{\partial H}{\partial q_i} \mathrm{d}q_i + \frac{\partial H}{\partial p_i} \mathrm{d}p_i \right) + \frac{\partial H}{\partial t} \mathrm{d}t \,. \tag{1.161}$$

Comparing (1.160) and (1.161), we conclude that

$$\frac{\partial H}{\partial p_i} = \dot{q}_i \,, \quad \frac{\partial H}{\partial q_i} = -\dot{p}_i \,, \quad \text{and} \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \,, \quad i = 1, \ldots, f \,. \tag{1.162}$$

The first two of these equations, expressing $\dot{q}_1, \ldots, \dot{q}_f$ and $\dot{p}_1, \ldots, \dot{p}_f$ in terms of $q_1, \ldots, q_f$, $p_1, \ldots, p_f$, and t are called **Hamilton's equations of motion**. The last equation is an identity relating the partial derivatives of H and L with respect to t but plays no role in determining the time evolution of $q_1, \ldots, q_f$ and $p_1, \ldots, p_f$.

Because $H = E$, it seems we have two symbols referring to the same quantity. We emphasize, however, that the Hamiltonian H refers to the energy of a mechanical system *expressed as a function of generalized coordinates* $q_1, \ldots, q_f$ *and their conjugate momenta* $p_1, \ldots, p_f$. In contrast, the energy expressed as a function of other variables is not a Hamiltonian. Such a function will be denoted by E. On occasion, we also use E to refer to the *value* of H when $q_1, \ldots, q_f$ and $p_1, \ldots, p_f$ assume certain values.

Example 1.7. Particle moving in a conservative field: Consider a particle of mass m moving in a conservative potential field $\psi(\boldsymbol{r})$ with velocity $\boldsymbol{v}$. The kinetic energy of the particle is given by

$$\frac{1}{2} m \boldsymbol{v} \cdot \boldsymbol{v} \,, \tag{1.163}$$

and hence

$$L = \frac{1}{2} m \boldsymbol{v} \cdot \boldsymbol{v} - \psi(\boldsymbol{r}) \,. \tag{1.164}$$

By definition, the momentum $\boldsymbol{p}$ that is conjugate to $\boldsymbol{v}$ is

$$\boldsymbol{p} = \frac{\partial L}{\partial \boldsymbol{v}} = m\boldsymbol{v} \,, \tag{1.165}$$

and hence

$$\boldsymbol{v} = \frac{\boldsymbol{p}}{m} \,. \tag{1.166}$$

The Hamiltonian is obtained as follows:

$$H = \boldsymbol{p} \cdot \boldsymbol{v} - L = \boldsymbol{p} \cdot \frac{\boldsymbol{p}}{m} - \frac{1}{2} m \frac{\boldsymbol{p}}{m} \cdot \frac{\boldsymbol{p}}{m} + \psi(\boldsymbol{r}) = \frac{\boldsymbol{p} \cdot \boldsymbol{p}}{2m} + \psi(\boldsymbol{r}) \,. \tag{1.167}$$

From Hamilton's equations of motion,

$$\dot{\boldsymbol{r}} = \frac{\partial H}{\partial \boldsymbol{p}} = \frac{\boldsymbol{p}}{m} \quad \text{and} \quad \dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{r}} = -\frac{\partial \psi}{\partial \boldsymbol{r}} \,. \tag{1.168}$$

Eliminating $\boldsymbol{p}$ from these two equations, we arrive at

$$m\ddot{\boldsymbol{r}} = -\frac{\partial \psi}{\partial \boldsymbol{r}} \,, \tag{1.169}$$

which is Newton's equation of motion of the particle.

Exercise 1.12. Translate the above example using components of the vectors involved. ///

Exercise 1.13. By taking the partial derivative of (1.156) with respect to t while holding $q_1, \ldots, q_f$ and $p_1, \ldots, p_f$ constant, directly establish the third equation in (1.162). ///

Exercise 1.14. We obtained (1.162) from (1.153), (1.156), (1.157), and Lagrange's equations of motion. Since Lagrangian and Hamiltonian are equivalent ways of encoding the mechanical behavior of the system, it should be possible to reverse this process. Combining (1.162) with (1.153), (1.156), and (1.157), derive Lagrange's equations of motion. ///

1.10 †Routhian

Given the Lagrangian as in (1.153), we have

$$p_1 = \frac{\partial L}{\partial \dot{q}_1} = p_1(q_1, \ldots, q_f, \dot{q}_1, \ldots, \dot{q}_f, t) \,, \tag{1.170}$$

which may be solved for $\dot{q}_1$ to yield

$$\dot{q}_1 = \dot{q}_1(q_1, \ldots, q_f, p_1, \dot{q}_2, \ldots, \dot{q}_f, t) \,. \tag{1.171}$$

Using this equation, we express a new function

$$R_1 := L - p_1 \dot{q}_1 \tag{1.172}$$

as a function of $q_1, \ldots, q_f$, p_1, $\dot{q}_2, \ldots, \dot{q}_f$, and t:

$$R_1 = R_1(q_1, \ldots, q_f, p_1, \dot{q}_2, \ldots, \dot{q}_f, t) \,. \tag{1.173}$$

A partial Legendre transformation, such as R_1, is called a **Routhian** and is an equally valid encoding of the mechanical behavior of the system as L and H.

Before we can proceed to the next round of Legendre transformation, in which we replace $\dot{q}_2$ by p_2, we need to evaluate $\partial R_1/\partial \dot{q}_2$. From (1.172), we see that

$$\begin{aligned} \mathrm{d}R_1 &= \mathrm{d}L - p_1 \mathrm{d}\dot{q}_1 - \dot{q}_1 \mathrm{d}p_1 \\ &= \sum_{i=1}^{f} \frac{\partial L}{\partial q_i} \mathrm{d}q_i + \sum_{i=1}^{f} \frac{\partial L}{\partial \dot{q}_i} \mathrm{d}\dot{q}_i + \frac{\partial L}{\partial t} \mathrm{d}t - p_1 \mathrm{d}\dot{q}_1 - \dot{q}_1 \mathrm{d}p_1 \\ &= \sum_{i=1}^{f} \frac{\partial L}{\partial q_i} \mathrm{d}q_i + \sum_{i=2}^{f} \frac{\partial L}{\partial \dot{q}_i} \mathrm{d}\dot{q}_i + \frac{\partial L}{\partial t} \mathrm{d}t - \dot{q}_1 \mathrm{d}p_1 , \end{aligned} \tag{1.174}$$

where we used (1.103). It follows that

$$\frac{\partial R_1}{\partial \dot{q}_2} = \frac{\partial L}{\partial \dot{q}_2} = p_2 . \tag{1.175}$$

It is important to note that the partial derivative on the left is evaluated by holding $q_1, \ldots, q_f, p_1, \dot{q}_3, \ldots, \dot{q}_f$, and t constant. In contrast, that on the right is for constant $q_1, \ldots, q_f, \dot{q}_1, \dot{q}_3, \ldots, \dot{q}_f$, and t.

The left most expression of (1.175) gives p_2 as a function of $q_1, \ldots, q_f, p_1, \dot{q}_2, \ldots, \dot{q}_f$, and t. When this is solved for $\dot{q}_2$, we obtain

$$\dot{q}_2 = \dot{q}_2(q_1, \ldots, q_f, p_1, p_2, \dot{q}_3, \ldots, \dot{q}_f, t) . \tag{1.176}$$

Using this equation, we now express

$$R_2 := R_1 - p_2 \dot{q}_2 = L - p_1 \dot{q}_1 - p_2 \dot{q}_2 \tag{1.177}$$

as a function of $q_1, \ldots, q_f, p_1, p_2, \dot{q}_3, \ldots, \dot{q}_f$, and t. The resulting function is another Routhian. Continuing in this way, we finally arrive at

$$R_f := L - \sum_{i=1}^{f} p_i \dot{q}_i = R_f(q_1, \ldots, q_f, p_1, \ldots, p_f, t) . \tag{1.178}$$

Now that all of $\dot{q}_1, \ldots, \dot{q}_f$ are replaced by $p_1, \ldots, p_f$, we do not refer to R_f as a Routhian. Instead, it is the negative of the Hamiltonian H.

Exercise 1.15. Derive the following set of equations of motion from the Routhian R_n:

$$\frac{\partial R_n}{\partial q_i} = \dot{p}_i , \quad \frac{\partial R_n}{\partial p_i} = -\dot{q}_i , \quad i = 1, \ldots, n , \tag{1.179}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial R_n}{\partial \dot{q}_i} \right) - \frac{\partial R_n}{\partial q_i} = 0 , \quad i = n+1, \ldots, f . \tag{1.180}$$

///

Equations of motion derived from a Routhian find practical applications in an analysis of the stability of a steady motion.

1.11 Poisson Bracket

Quantities such as the energy and the total linear momentum of a system are called **dynamical variables**. More generally, we define a dynamical variable A as a function of the instantaneous state of the mechanical system, which is specified by q^f and p^f:

$$A(q^f, p^f, t) \,. \tag{1.181}$$

Here we allowed for an explicit time dependence of A. We also introduced a new notation in which q^f and p^f stand for $q_1, \ldots, q_f$ and $p_1, \ldots, p_f$, respectively. Note carefully the distinction between q^f and q_f. The former is the collective notation just introduced, while the latter refers to the fth generalized coordinate.

As the coordinates and momenta change in accordance with the laws of mechanics, the value of A will, in general, change as well. It is straightforward to find the expression for the time rate of change of A. Differentiating (1.181) with respect to time, we find

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \sum_{i=1}^{f} \left(\frac{\partial A}{\partial q_i} \dot{q}_i + \frac{\partial A}{\partial p_i} \dot{p}_i \right) + \frac{\partial A}{\partial t} \,. \tag{1.182}$$

Using Hamilton's equations of motion, we may rewrite the right-hand side as

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \sum_{i=1}^{f} \left(\frac{\partial A}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial q_i} \frac{\partial A}{\partial p_i} \right) + \frac{\partial A}{\partial t} \,. \tag{1.183}$$

We introduce the **Poisson bracket** of two dynamical variables A and B, which is defined by

$$\{A, B\} := \sum_{i=1}^{f} \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial B}{\partial q_i} \frac{\partial A}{\partial p_i} \right) \,, \tag{1.184}$$

and write (1.183) more compactly as

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \{A, H\} + \frac{\partial A}{\partial t} \,. \tag{1.185}$$

By definition, A is a **constant of motion** if $\mathrm{d}A/\mathrm{d}t \equiv 0$. It should be noted that the definition does *not* demand $\partial A/\partial t$ to be zero. Thus, a dynamical variable can be a constant of motion even if it has an explicit time dependence. The Hamiltonian is an exception to this rule. In fact, if we set $A = H$ in (1.185),

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \{H, H\} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} \,. \tag{1.186}$$

So, H is a constant of motion if and only if H does not depend explicitly on time.

As another example of dynamical variables, let us consider q_k. Using (1.185), we find

$$\dot{q}_k = \{q_k, H\} + \frac{\partial q_k}{\partial t} \,. \tag{1.187}$$

As in (1.183), the partial derivative with respect to t is taken while holding q^f and p^f constant. Thus, $\partial q_k/\partial t = 0$ and we obtain

$$\dot{q}_k = \{q_k, H\} \,. \tag{1.188}$$

Similarly, by setting $A = p_k$, we find

$$\dot{p}_k = \{p_k, H\} \,. \tag{1.189}$$

As seen in Exercise 1.16, (1.188) and (1.189) are nothing but Hamilton's equations of motion.

We note that, since q's and p's are independent variables, we have $\partial p_k/\partial q_i \equiv 0$ and $\partial q_k/\partial p_i \equiv 0$ regardless of the values of i and k. On the other hand, $\partial q_k/\partial q_i$ is unity if $i = k$ and zero if $i \neq k$, that is,

$$\frac{\partial q_k}{\partial q_i} = \delta_{ik} \quad \text{and} \quad \frac{\partial p_k}{\partial p_i} = \delta_{ik} \,, \tag{1.190}$$

where δ_{ik} is the **Kronecker delta**, which takes the value unity if $i = k$ and zero otherwise.

Exercise 1.16. Let A be a dynamical variable of a mechanical system with f degrees of freedom, that is, $A = A(q^f, p^f, t)$. Evaluate $\{q_j, A\}$ and $\{p_j, A\}$. Then, show that (1.188) and (1.189) are Hamilton's equations of motion. ///

Exercise 1.17. Using (1.190), show that

$$\{q_i, q_j\} = 0 \,, \quad \{q_i, p_j\} = \delta_{ij} \,, \quad \text{and} \quad \{p_i, p_j\} = 0 \,. \tag{1.191}$$

///

The Poisson bracket satisfies a number of identities:

$$\{A, A\} = 0 \,, \tag{1.192}$$

$$\{c, A\} = 0 \,, \tag{1.193}$$

$$\{A, B + C\} = \{A, B\} + \{A, C\} \,, \tag{1.194}$$

$$\{A, BC\} = B\{A, C\} + \{A, B\}C \,, \tag{1.195}$$

where c is a real number. Through a straightforward but lengthy computation, one can prove **Jacobi's identity**:

$$\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0 \,. \tag{1.196}$$

With the help of this identity, we can show that, if A and B are constants of motion, then so is $\{A, B\}$. This is known as the **Poisson theorem**, which, on occasion, allows us to generate a new constant of motion from known constants of motion.

Exercise 1.18. Prove the Poisson theorem.

a. For simplicity, first assume that neither A nor B depends explicitly on time.
b. Generalize the theorem when A and B may depend explicitly on time. ///

Finally, there appears to be no general agreement on the sign of the Poisson bracket. For example, Ref. [3] defines it as

$$\{A,H\} = \sum_{i=1}^{f} \left(\frac{\partial A}{\partial p_i}\frac{\partial H}{\partial q_i} - \frac{\partial H}{\partial p_i}\frac{\partial A}{\partial q_i} \right) , \tag{1.197}$$

which has the opposite sign to what we defined.

1.12 Frequently Used Symbols

$A := B$, the symbol A is defined by the expression B.
$A =: B$, the expression A defines the symbol B.
$\boldsymbol{a} \doteq (a_x, a_y, a_z)$, the x-, y-, z-components of the vector $\boldsymbol{a}$ are a_x, a_y, and a_z, respectively.
$\|\boldsymbol{a}\|$, length of the vector $\boldsymbol{a}$.
$\mathrm{d}\phi$, the first order term of $\Delta\phi$.
$\Delta\phi$, change in ϕ including the higher order terms.

f , the number of mechanical degrees of freedom.
$\boldsymbol{g}$, gravitational acceleration.
h , energy function.
h.o. , higher order terms of Taylor series expansion, typically second order and higher.
m , mass of a particle.
p_i , generalized momentum conjugate to q_i.
p^f , collective notation for $p_1, \ldots, p_f$.
$\boldsymbol{p}_i$, linear momentum of the ith particle.
q_i , the ith generalized coordinate.
q^f , collective notation for $q_1, \ldots, q_f$.
$\boldsymbol{r}_i$, position vector of the ith particle.
t , time.
$\boldsymbol{v}_i$, velocity vector of the ith particle.

E , mechanical energy.
$\boldsymbol{F}$, force.
H , Hamiltonian.
L , Lagrangian.
M , total mass of a many-particle system.
$\boldsymbol{M}$, total angular momentum of a many-particle system.
$\boldsymbol{P}$, total linear momentum of a many-particle system.

R , Routhian.
$\boldsymbol{R}$, position of the center of mass.
$\boldsymbol{V}$, velocity of the center of mass.
W , work.

$\mathscr{S}$, action.

$\delta q(t)$, infinitesimal variation of some function $q(t)$ at time t.
δ_{ij} , Kronecker delta.
μ , reduced mass.
ϕ , potential energy due to interparticle interactions.
ψ , potential energy due to an external field.

References and Further Reading

1. Goldstein H (1980) Classical Mechanics, 2nd edn. Addison-Wesley, Reading, Massachusetts
   The book has much to recommend and it is in fact one of the standard references on classical mechanics. For a detailed discussion of what we saw in this chapter, see Chaps. 1, 2, and 8. For a discussion on the relation between the energy function and the mechanical energy, see pp. 60-61. The corresponding discussion regarding the Hamiltonian and the mechanical energy is in pp. 349-351. For Routhian and its application, see Sect. 8.3. Note that his Routhian is the negative of ours. Various properties of the Poisson bracket are given in Chap. 9.
2. Lanczos C (1986) The variational principles of mechanics. Dover, New York
   A very detailed, yet highly accessible and engaging, discussion of the variational principles including Hamilton's principle.
3. Landau L D, Lifshitz E M (1973) Mechanics, 3rd edn. Pergamon Press, New York
   A very short account on classical mechanics. The book is full of unusual insights, though not everything they say is immediately obvious. Their Poisson bracket is the negative of ours.
4. Tolman R C (1979) The principles of statistical mechanics. Dover, New York
   Chapter 2 gives a short review of classical mechanics including topics we have omitted.