1 Newtonian mechanics

1.1 Introduction

Newtonian mechanics is a mathematical model whose purpose is to account for the motions of the various objects in the universe. The general principles of this model were first enunciated by Sir Isaac Newton in a work titled *Philosophiae Naturalis Principia Mathematica* (Mathematical Principles of Natural Philosophy). This work, which was published in 1687, is nowadays more commonly referred to as the *Principia*.[1]

Until the beginning of the twentieth century, Newtonian mechanics was thought to constitute a *complete* description of all types of motion occurring in the universe. We now know that this is not the case. The modern view is that Newton's model is only an *approximation* that is valid under certain circumstances. The model breaks down when the velocities of the objects under investigation approach the speed of light in a vacuum, and must be modified in accordance with Einstein's *special theory of relativity*. The model also fails in regions of space that are sufficiently curved that the propositions of Euclidean geometry do not hold to a good approximation, and must be augmented by Einstein's *general theory of relativity*. Finally, the model breaks down on atomic and subatomic length scales, and must be replaced by *quantum mechanics*. In this book, we shall (almost entirely) neglect relativistic and quantum effects. It follows that we must restrict our investigations to the motions of *large* (compared with an atom), *slow* (compared with the speed of light) objects moving in *Euclidean* space. Fortunately, virtually all the motions encountered in conventional celestial mechanics fall into this category.

Newton very deliberately modeled his approach in the *Principia* on that taken in Euclid's *Elements*. Indeed, Newton's theory of motion has much in common with a conventional *axiomatic system*, such as Euclidean geometry. Like all axiomatic systems, Newtonian mechanics starts from a set of terms that are *undefined* within the system. In this case, the fundamental terms are *mass*, *position*, *time*, and *force*. It is taken for granted that we understand what these terms mean, and, furthermore, that they correspond to *measurable* quantities that can be ascribed to, or associated with, objects in the world around us. In particular, it is assumed that the ideas of position in space, distance in space, and position as a function of time in space are correctly described by conventional Euclidean vector algebra and vector calculus. The next component of an axiomatic system is a set of *axioms*. These are a set of *unproven* propositions,

[1] An excellent discussion of the historical development of Newtonian mechanics, as well as the physical and philosophical assumptions that underpin this theory, is given in Barbour 2001.

involving the undefined terms, from which all other propositions in the system can be derived via logic and mathematical analysis. In the present case, the axioms are called *Newton's laws of motion* and can be justified only via experimental observation. Note, incidentally, that Newton's laws, in their primitive form, are applicable only to *point objects* (i.e., objects of negligible spatial extent). However, these laws can be applied to extended objects by treating them as collections of point objects.

One difference between an axiomatic system and a physical theory is that, in the latter case, even if a given prediction has been shown to follow necessarily from the axioms of the theory, it is still incumbent on us to test the prediction against experimental observations. Lack of agreement might indicate faulty experimental data, faulty application of the theory (for instance, in the case of Newtonian mechanics, there might be forces at work that we have not identified), or, as a last resort, incorrectness of the theory. Fortunately, Newtonian mechanics has been found to give predictions that are in excellent agreement with experimental observations in all situations in which it would be expected to hold.

In the following, it is assumed that we know how to set up a rigid Cartesian frame of reference and how to measure the positions of point objects as functions of time within that frame. It is also taken for granted that we have some basic familiarity with the laws of mechanics.

1.2 Newton's laws of motion

Newton's laws of motion, in the rather obscure language of the *Principia*, take the following form:

1. Every body continues in its state of rest, or uniform motion in a straight line, unless compelled to change that state by forces impressed on it.
2. The change of motion (i.e., momentum) of an object is proportional to the force impressed on it, and is made in the direction of the straight line in which the force is impressed.
3. To every action there is always opposed an equal reaction; or, the mutual actions of two bodies on each other are always equal, and directed to contrary parts.

Let us now examine how these laws can be applied to a system of point objects.

1.3 Newton's first law of motion

Newton's first law of motion essentially states that a point object subject to zero net external force moves in a straight line with a constant speed (i.e., it does not accelerate). However, this is true only in special frames of reference called *inertial frames*. Indeed, we can think of Newton's first law as the definition of an inertial frame: an inertial frame of reference is one in which a point object subject to zero net external force moves in a straight line with constant speed.

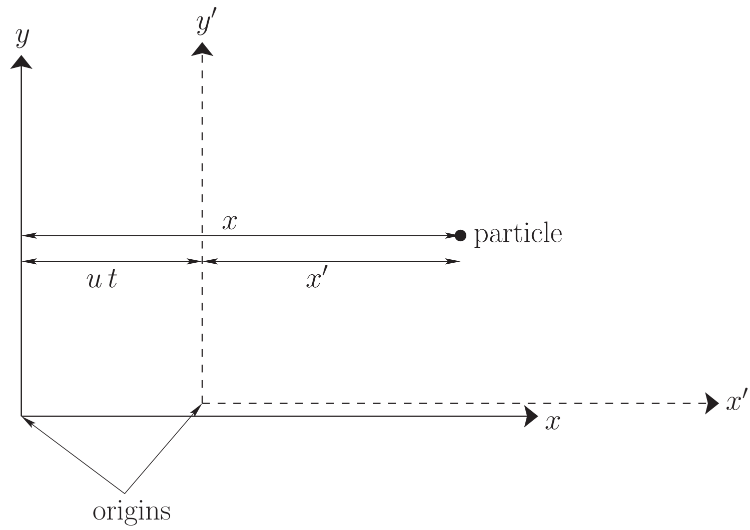

Fig. 1.1 A Galilean coordinate transformation.

Suppose that we have found an inertial frame of reference. Let us set up a Cartesian coordinate system in this frame. The motion of a point object can now be specified by giving its position vector, $\mathbf{r} \equiv (x, y, z)$, with respect to the origin of the coordinate system, as a function of time, t. Consider a second frame of reference moving with some constant velocity $\mathbf{u}$ with respect to the first frame. Without loss of generality, we can suppose that the Cartesian axes in the second frame are parallel to the corresponding axes in the first frame, that $\mathbf{u} \equiv (u, 0, 0)$, and, finally, that the origins of the two frames instantaneously coincide at $t = 0$. (See Figure 1.1.) Suppose that the position vector of our point object is $\mathbf{r}' \equiv (x', y', z')$ in the second frame of reference. It is evident, from Figure 1.1, that at any given time, t, the coordinates of the object in the two reference frames satisfy

$$x' = x - u\,t, \tag{1.1}$$

$$y' = y, \tag{1.2}$$

and

$$z' = z. \tag{1.3}$$

This coordinate transformation was first discovered by Galileo Galilei (1564–1642), and is nowadays known as a *Galilean transformation* in his honor.

By definition, the instantaneous velocity of the object in our first reference frame is given by $\mathbf{v} = d\mathbf{r}/dt \equiv (dx/dt, dy/dt, dz/dt)$, with an analogous expression for the velocity, $\mathbf{v}'$, in the second frame. It follows, from differentiation of Equations (1.1)–(1.3) with respect to time, that the velocity components in the two frames satisfy

$$v_x' = v_x - u, \tag{1.4}$$

$$v_y' = v_y, \tag{1.5}$$

and

$$v_z' = v_z. \tag{1.6}$$

These equations can be written more succinctly as

$$\mathbf{v}' = \mathbf{v} - \mathbf{u}. \tag{1.7}$$

Finally, by definition, the instantaneous acceleration of the object in our first reference frame is given by $\mathbf{a} = d\mathbf{v}/dt \equiv (dv_x/dt,\, dv_y/dt,\, dv_z/dt)$, with an analogous expression for the acceleration, $\mathbf{a}'$, in the second frame. It follows, from differentiation of Equations (1.4)–(1.6) with respect to time, that the acceleration components in the two frames satisfy

$$a_x' = a_x, \tag{1.8}$$

$$a_y' = a_y, \tag{1.9}$$

and

$$a_z' = a_z. \tag{1.10}$$

These equations can be written more succinctly as

$$\mathbf{a}' = \mathbf{a}. \tag{1.11}$$

According to Equations (1.7) and (1.11), if an object is moving in a straight line with a constant speed in our original inertial frame (i.e., if $\mathbf{a} = \mathbf{0}$), then it also moves in a (different) straight line with a (different) constant speed in the second frame of reference (i.e., $\mathbf{a}' = \mathbf{0}$). Hence, we conclude that the second frame of reference is also an inertial frame.

A simple extension of the preceding argument allows us to conclude that there is an *infinite* number of different inertial frames moving with constant velocities with respect to one another. Newton thought that one of these inertial frames was special and defined an absolute standard of rest: that is, a static object in this frame was in a state of absolute rest. However, Einstein showed that this is not the case. In fact, there is no absolute standard of rest: in other words, all motion is relative—hence, the name *relativity* for Einstein's theory. Consequently, one inertial frame is just as good as another as far as Newtonian mechanics is concerned.

But what happens if the second frame of reference *accelerates* with respect to the first? In this case, it is not hard to see that Equation (1.11) generalizes to

$$\mathbf{a}' = \mathbf{a} - \frac{d\mathbf{u}}{dt}, \tag{1.12}$$

where $\mathbf{u}(t)$ is the instantaneous velocity of the second frame with respect to the first. According to this formula, if an object is moving in a straight line with a constant speed in the first frame (i.e., if $\mathbf{a} = \mathbf{0}$), then it does not move in a straight line with a constant speed in the second frame (i.e., $\mathbf{a}' \neq \mathbf{0}$). Hence, if the first frame is an inertial frame, then the second is *not*.

A simple extension of the preceding argument allows us to conclude that any frame of reference that accelerates with respect to a given inertial frame is not itself an inertial frame.

For most practical purposes, when studying the motions of objects close to the Earth's surface, a reference frame that is fixed with respect to this surface is approximately inertial. However, if the trajectory of a projectile within such a frame is measured to high precision, then it will be found to deviate slightly from the predictions of Newtonian mechanics. (See Chapter 5.) This deviation is due to the fact that the Earth is rotating, and its surface is therefore *accelerating* toward its axis of rotation. When studying the motions of objects in orbit around the Earth, a reference frame whose origin is the center of the Earth (or, to be more exact, the center of mass of the Earth–Moon system), and whose coordinate axes are fixed with respect to distant stars, is approximately inertial. However, if such orbits are measured to extremely high precision, then they will again be found to deviate very slightly from the predictions of Newtonian mechanics. In this case, the deviation is due to the Earth's orbital motion about the Sun. When studying the orbits of the planets in the solar system, a reference frame whose origin is the center of the Sun (or, to be more exact, the center of mass of the solar system), and whose coordinate axes are fixed with respect to distant stars, is approximately inertial. In this case, any deviations of the orbits from the predictions of Newtonian mechanics due to the orbital motion of the Sun about the galactic center are far too small to be measurable. It should be noted that it is impossible to identify an *absolute* inertial frame—the best approximation to such a frame would be one in which the cosmic microwave background appears to be (approximately) isotropic. However, for a given dynamic problem, it is always possible to identify an *approximate* inertial frame. Furthermore, any deviations of such a frame from a true inertial frame can be incorporated into the framework of Newtonian mechanics via the introduction of so-called fictitious forces. (See Chapter 5.)

1.4 Newton's second law of motion

Newton's second law of motion essentially states that if a point object is subject to an external force, $\mathbf{f}$, then its equation of motion is given by

$$\frac{d\mathbf{p}}{dt} = \mathbf{f}, \tag{1.13}$$

where the momentum, $\mathbf{p}$, is the product of the object's inertial mass, m, and its velocity, $\mathbf{v}$. If m is not a function of time, then Equation (1.13) reduces to the familiar equation

$$m\,\frac{d\mathbf{v}}{dt} = \mathbf{f}. \tag{1.14}$$

This equation is valid only in an *inertial frame*. Clearly, the inertial mass of an object measures its reluctance to deviate from its preferred state of uniform motion in a straight line (in an inertial frame). Of course, the preceding equation of motion can be solved only if we have an independent expression for the force, $\mathbf{f}$ (i.e., a law of force). Let us suppose that this is the case.

An important corollary of Newton's second law is that force is a *vector quantity*. This must be the case, as the law equates force to the product of a scalar (mass) and a vector (acceleration).[2] Note that acceleration is obviously a vector because it is directly related to displacement, which is the prototype of all vectors. One consequence of force being a vector is that two forces, $\mathbf{f}_1$ and $\mathbf{f}_2$, both acting at a given point, have the same effect as a single force, $\mathbf{f} = \mathbf{f}_1 + \mathbf{f}_2$, acting at the same point, where the summation is performed according to the laws of vector addition. Likewise, a single force, $\mathbf{f}$, acting at a given point, has the same effect as two forces, $\mathbf{f}_1$ and $\mathbf{f}_2$, acting at the same point, provided that $\mathbf{f}_1 + \mathbf{f}_2 = \mathbf{f}$. This method of combining and splitting forces is known as the *resolution of forces*; it lies at the heart of many calculations in Newtonian mechanics.

Taking the scalar product of Equation (1.14) with the velocity, $\mathbf{v}$, we obtain

$$m\,\mathbf{v}\cdot\frac{d\mathbf{v}}{dt} = \frac{m}{2}\frac{d(\mathbf{v}\cdot\mathbf{v})}{dt} = \frac{m}{2}\frac{dv^2}{dt} = \mathbf{f}\cdot\mathbf{v}. \tag{1.15}$$

This can be written

$$\frac{dK}{dt} = \mathbf{f}\cdot\mathbf{v}, \tag{1.16}$$

where

$$K = \frac{1}{2}\,m\,v^2. \tag{1.17}$$

The right-hand side of Equation (1.16) represents the rate at which the force does work on the object—that is, the rate at which the force transfers energy to the object. The quantity K represents the energy that the object possesses by virtue of its motion. This type of energy is generally known as *kinetic energy*. Thus, Equation (1.16) states that any work done on a point object by an external force goes to increase the object's kinetic energy.

Suppose that under the action of the force, $\mathbf{f}$, our object moves from point P at time t_1 to point Q at time t_2. The net change in the object's kinetic energy is obtained by integrating Equation (1.16):

$$\varDelta K = \int_{t_1}^{t_2}\mathbf{f}\cdot\mathbf{v}\,dt = \int_P^Q \mathbf{f}\cdot d\mathbf{r}, \tag{1.18}$$

because $\mathbf{v} = d\mathbf{r}/dt$. Here, $d\mathbf{r}$ is an element of the object's path between points P and Q, and the integral in $\mathbf{r}$ represents the net *work* done by the force as the object moves along the path from P to Q.

As is well known, there are basically two kinds of forces in nature: first, those for which line integrals of the type $\int_P^Q\mathbf{f}\cdot d\mathbf{r}$ depend on the end points but not on the path taken between these points; second, those for which line integrals of the type $\int_P^Q\mathbf{f}\cdot d\mathbf{r}$ depend both on the end points and the path taken between these points. The first kind of force is termed *conservative*, whereas the second kind is termed *non-conservative*. It can be demonstrated that if the line integral $\int_P^Q \mathbf{f}\cdot d\mathbf{r}$ is *path independent*, for all choices of P and Q, then the force $\mathbf{f}$ can be written as the gradient of a scalar field. (See Section A.5.)

[2] A *scalar* is a physical quantity that is invariant under rotation of the coordinate axes. A *vector* is a physical quantity that transforms in an analogous manner to a displacement under rotation of the coordinate axes.

In other words, all conservative forces satisfy

$$\mathbf{f}(\mathbf{r}) = -\nabla U \tag{1.19}$$

for some scalar field $U(\mathbf{r})$. [Incidentally, mathematicians, as opposed to physicists and astronomers, usually write $f(\mathbf{r}) = +\nabla U$.] Note that

$$\int_P^Q \nabla U \cdot d\mathbf{r} \equiv \Delta U = U(Q) - U(P), \tag{1.20}$$

irrespective of the path taken between P and Q. Hence, it follows from Equation (1.18) that

$$\Delta K = -\Delta U \tag{1.21}$$

for conservative forces. Another way of writing this is

$$E = K + U = \text{constant}. \tag{1.22}$$

Of course, we recognize Equation (1.22) as an *energy conservation equation*: E is the object's total energy, which is conserved; K is the energy the object has by virtue of its motion, otherwise known as its *kinetic energy*; and U is the energy the object has by virtue of its position, otherwise known as its *potential energy*. Note, however, that we can write energy conservation equations only for conservative forces. Gravity is an obvious example of such a force. Incidentally, potential energy is undefined to an arbitrary additive constant. In fact, it is only the *difference* in potential energy between different points in space that is well defined.

1.5 Newton's third law of motion

Consider a system of N mutually interacting point objects. Let the ith object, whose mass is m_i, be located at position vector $\mathbf{r}_i$. Suppose that this object exerts a force $\mathbf{f}_{ji}$ on the jth object. Likewise, suppose that the jth object exerts a force $\mathbf{f}_{ij}$ on the ith object. Newton's third law of motion essentially states that these two forces are equal and opposite, irrespective of their nature. In other words,

$$\mathbf{f}_{ij} = -\mathbf{f}_{ji}. \tag{1.23}$$

(See Figure 1.2.) One corollary of Newton's third law is that an object cannot exert a force on itself. Another corollary is that all forces in the universe have corresponding reactions. The only exceptions to this rule are the fictitious forces that arise in non-inertial reference frames (e.g., the centrifugal and Coriolis forces that appear in rotating reference frames—see Chapter 5). Fictitious forces do not generally possess reactions.

Newton's third law implies *action at a distance*. In other words, if the force that object i exerts on object j suddenly changes, then Newton's third law demands that there must be an *immediate* change in the force that object j exerts on object i. Moreover, this must be true irrespective of the distance between the two objects. However, we now know that Einstein's special theory of relativity forbids information from traveling through

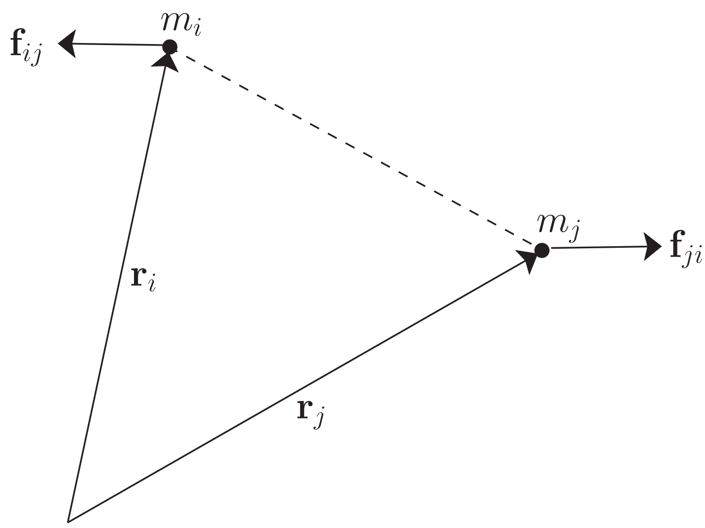

Fig. 1.2 Newton's third law.

the universe faster than the velocity of light in vacuum. Hence, action at a distance is also forbidden. In other words, if the force that object i exerts on object j suddenly changes, then there must be a *time delay*, which is at least as long as it takes a light ray to propagate between the two objects, before the force that object j exerts on object i can respond. Of course, this means that Newton's third law is not, strictly speaking, correct. However, as long as we restrict our investigations to the motions of dynamical systems over timescales that are long compared with the time required for light rays to traverse these systems, Newton's third law can be regarded as being approximately correct.

In an inertial frame, Newton's second law of motion applied to the ith object yields

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = \sum_{j=1,N}^{j\neq i} \mathbf{f}_{ij}. \tag{1.24}$$

Note that the summation on the right-hand side of this equation excludes the case $j = i$, as the ith object cannot exert a force on itself. Let us now take this equation and sum it over all objects. We obtain

$$\sum_{i=1,N} m_i \frac{d^2\mathbf{r}_i}{dt^2} = \sum_{i,j=1,N}^{j\neq i} \mathbf{f}_{ij}. \tag{1.25}$$

Consider the sum over forces on the right-hand side of the preceding equation. Each element of this sum—$\mathbf{f}_{ij}$, say—can be paired with another element—$\mathbf{f}_{ji}$, in this case—that is equal and opposite, according to Newton's third law. In other words, the elements of the sum all cancel out in pairs. Thus, the net value of the sum is *zero*. It follows that Equation (1.25) can be written

$$M \frac{d^2\mathbf{r}_{cm}}{dt^2} = \mathbf{0}, \tag{1.26}$$

where $M = \sum_{i=1,N} m_i$ is the total mass. The quantity $\mathbf{r}_{cm}$ is the vector displacement of the *center of mass* of the system, which is an imaginary point whose coordinates are the mass weighted averages of the coordinates of the objects that constitute the system:

$$\mathbf{r}_{cm} = \frac{\sum_{i=1,N} m_i\,\mathbf{r}_i}{\sum_{i=1,N} m_i}. \tag{1.27}$$

According to Equation (1.26), the center of mass of the system moves in a uniform straight line, in accordance with Newton's first law of motion, irrespective of the nature of the forces acting between the various components of the system.

Now, if the center of mass moves in a uniform straight line, then the center of mass velocity,

$$\frac{d\mathbf{r}_{cm}}{dt} = \frac{\sum_{i=1,N} m_i\,d\mathbf{r}_i/dt}{\sum_{i=1,N} m_i}, \tag{1.28}$$

is a constant of the motion. However, the momentum of the ith object takes the form $\mathbf{p}_i = m_i\,d\mathbf{r}_i/dt$. Hence, the total momentum of the system is written

$$\mathbf{P} = \sum_{i=1,N} m_i\,\frac{d\mathbf{r}_i}{dt}. \tag{1.29}$$

A comparison of Equations (1.28) and (1.29) suggests that $\mathbf{P}$ is also a constant of the motion. In other words, the total momentum of the system is a *conserved* quantity, irrespective of the nature of the forces acting between the various components of the system. This result (which holds only if there is zero net external force acting on the system) is a direct consequence of Newton's third law of motion.

Taking the vector product of Equation (1.24) with the position vector $\mathbf{r}_i$, we obtain

$$m_i\,\mathbf{r}_i \times \frac{d^2\mathbf{r}_i}{dt^2} = \sum_{j=1,N}^{j\neq i} \mathbf{r}_i \times \mathbf{f}_{ij}. \tag{1.30}$$

The right-hand side of the this equation is the net *torque* about the origin that acts on object i as a result of the forces exerted on it by the other objects. It is easily seen that

$$m_i\,\mathbf{r}_i \times \frac{d^2\mathbf{r}_i}{dt^2} = \frac{d}{dt}\left(m_i\,\mathbf{r}_i \times \frac{d\mathbf{r}_i}{dt}\right) = \frac{d\mathbf{l}_i}{dt}, \tag{1.31}$$

where

$$\mathbf{l}_i = m_i\,\mathbf{r}_i \times \frac{d\mathbf{r}_i}{dt} \tag{1.32}$$

is the *angular momentum* of the ith object about the origin of our coordinate system. Moreover, the total angular momentum of the system (about the origin) takes the form

$$\mathbf{L} = \sum_{i=1,N} \mathbf{l}_i. \tag{1.33}$$

Hence, summing Equation (1.30) over all particles, we obtain

$$\frac{d\mathbf{L}}{dt} = \sum_{i,j=1,N}^{i\neq j} \mathbf{r}_i \times \mathbf{f}_{ij}. \tag{1.34}$$

Consider the sum on the right-hand side of Equation (1.34). A general term, $\mathbf{r}_i \times \mathbf{f}_{ij}$, in this sum can always be paired with a matching term, $\mathbf{r}_j \times \mathbf{f}_{ji}$, in which the indices

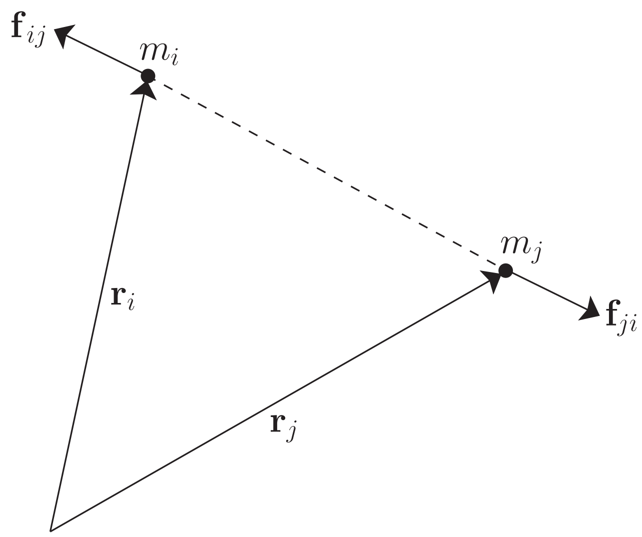

Fig. 1.3 Central forces.

have been swapped. Making use of Equation (1.23), we can write the sum of a general matched pair as

$$\mathbf{r}_i \times \mathbf{f}_{ij} + \mathbf{r}_j \times \mathbf{f}_{ji} = (\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{f}_{ij}. \tag{1.35}$$

Let us assume that the forces acting between the various components of the system are *central* in nature, so that $\mathbf{f}_{ij}$ is parallel to $\mathbf{r}_i - \mathbf{r}_j$. In other words, the force exerted on object j by object i either points directly toward, or directly away from, object i, and vice versa. (See Figure 1.3.) This is a reasonable assumption, as virtually all the forces that we encounter in celestial mechanics are of this type (e.g., gravity). It follows that if the forces are central, then the vector product on the right-hand side of the above expression is zero. We conclude that

$$\mathbf{r}_i \times \mathbf{f}_{ij} + \mathbf{r}_j \times \mathbf{f}_{ji} = \mathbf{0} \tag{1.36}$$

for all values of i and j. Thus, the sum on the right-hand side of Equation (1.34) is zero for any kind of central force. We are left with

$$\frac{d\mathbf{L}}{dt} = \mathbf{0}. \tag{1.37}$$

In other words, the total angular momentum of the system is a *conserved* quantity, provided that the different components of the system interact via *central* forces (and there is zero net external torque acting on the system).

1.6 Nonisolated systems

Up to now, we have considered only *isolated* dynamical systems, in which all the forces acting on the system originate from within the system itself. Let us now generalize