3

Classical mechanics

In this chapter we study the use of geometric algebra in classical mechanics. We will assume that readers already have a basic understanding of the subject, as a complete presentation of classical mechanics with geometric algebra would require an entire book. Such a book has been written, *New Foundations for Classical Mechanics* by David Hestenes (1999), which looks in detail at many of the topics discussed here. Our main focus in this chapter is to areas where geometric algebra offers some immediate benefits over traditional methods. These include motion in a central force and rigid-body rotations, both of which are dealt with in some detail. More advanced topics in Lagrangian and Hamiltonian dynamics are covered in chapter 12, and relativistic dynamics is covered in chapter 5.

Classical mechanics was one of the areas of physics that prompted the development of many of the mathematical techniques routinely used today. This is particularly true of vector analysis, and it is now common to see classical mechanics described using an abstract vector notation. Many of the formulae in this chapter should be completely familiar from such treatments. A key difference comes in adopting the outer product of vectors in place of the cross product. This means, for example, that angular momentum and torque both become bivectors. The outer product is clearer conceptually, but on its own it does not bring any calculational advantages. The main new computational tool we have at our disposal is the geometric product, and here we highlight a number of examples of its use.

In this chapter we have chosen to write all vectors in a bold font. This is conventional for three-dimensional physics and many of the formulae presented below look unnatural if this notation is not followed. Bivectors and other general multivectors are left in regular font, which helps to distinguish them from vectors.

3.1 Elementary principles

We start by considering a point particle with a trajectory $\boldsymbol{x}(t)$ described as a function of time. Here $\boldsymbol{x}$ is the position vector relative to some origin and the time t is taken as some absolute 'Newtonian' standard on which all observers agree. The particle has velocity

$$\boldsymbol{v} = \dot{\boldsymbol{x}} = \frac{d\boldsymbol{x}}{dt}, \tag{3.1}$$

where the overdot denotes differentiation with respect to time t. If the particle has mass m, then the momentum $\boldsymbol{p}$ is defined by $\boldsymbol{p} = m\boldsymbol{v}$. Newton's second law of motion states that

$$\dot{\boldsymbol{p}} = \boldsymbol{f}, \tag{3.2}$$

where the vector $\boldsymbol{f}$ is the force acting on the particle. Usually the mass m is constant and we recover the familiar expression $\boldsymbol{f} = m\boldsymbol{a}$, where $\boldsymbol{a}$ is the acceleration

$$\boldsymbol{a} = \frac{d^2\boldsymbol{x}}{dt^2}. \tag{3.3}$$

The case of constant mass is assumed throughout this chapter. The path for a single particle is then determined by a second-order differential equation (assuming $\boldsymbol{f}$ does not depend on higher derivatives).

The work done by the force $\boldsymbol{f}$ on a particle is defined by the line integral

$$W_{12} = \int_{t_1}^{t_2} \boldsymbol{f}\cdot\boldsymbol{v}\, dt = \int_1^2 \boldsymbol{f}\cdot d\boldsymbol{s}. \tag{3.4}$$

The final form here illustrates that the integral is independent of how the path is parameterised. From Newton's second law we have

$$W_{12} = m\int_{t_1}^{t_2} \dot{\boldsymbol{v}}\cdot\boldsymbol{v}\, dt = \frac{m}{2}\int_{t_1}^{t_2} \frac{d}{dt}(v^2)\, dt, \tag{3.5}$$

where $v = |\boldsymbol{v}| = \surd(\boldsymbol{v}^2)$. It follows that the work done is equal to the change in kinetic energy T, where

$$T = \tfrac{1}{2}mv^2. \tag{3.6}$$

In the case where the work is independent of the path from point 1 to point 2 the force is said to be *conservative*, and can be written as the gradient of a potential:

$$\boldsymbol{f} = -\boldsymbol{\nabla}V. \tag{3.7}$$

For conservative forces the work also evaluates to

$$W_{12} = -\int_1^2 d\boldsymbol{s}\cdot\boldsymbol{\nabla}V = V_1 - V_2 \tag{3.8}$$

and the total energy $E = T + V$ is conserved.

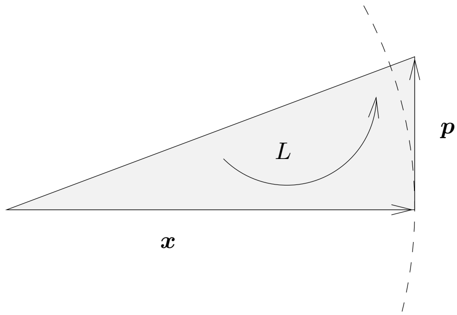

Figure 3.1 *Angular momentum.* The particle sweeps out the plane $L = \boldsymbol{x}\wedge\boldsymbol{p}$. The angular momentum should be directly related to the area swept out (cf. Kepler's second law), so is naturally encoded as a bivector. The position vector $\boldsymbol{x}$ depends on the choice of origin.

3.1.1 Angular momentum

Angular momentum is traditionally discussed in terms of the cross product, even though it is quite clear that what is required is a way of encoding the area swept out by a particle as it moves relative to some origin (see figure 3.1). We saw in chapter 2 that the exterior product provides this, and that the more traditional cross product is a derived concept based on the three-dimensional result that every directed plane has a unique normal. We therefore have no hesitation in dispensing with the traditional definition of angular momentum as an axial vector, and replace it with a bivector. So, if a particle has momentum $\boldsymbol{p}$ and position vector $\boldsymbol{x}$ from some origin, we define the angular momentum of the particle about the origin as the bivector

$$L = \boldsymbol{x}\wedge\boldsymbol{p}. \tag{3.9}$$

This definition does not alter the steps involved in computing L since the components are the same as those of the cross product. We will see, however, that the freedom we have to now use the geometric product can speed up derivations. The definition of angular momentum as a bivector maintains a clear distinction with vector quantities such as position and velocity, removing the need for the rather awkward definitions of polar and axial vectors. The definition of L as a bivector also fits neatly with the rotor description of rotations, as we shall see later in this chapter.

If we differentiate L we obtain

$$\frac{dL}{dt} = \boldsymbol{v}\wedge(m\boldsymbol{v}) + \boldsymbol{x}\wedge(m\boldsymbol{a}) = \boldsymbol{x}\wedge\boldsymbol{f}. \tag{3.10}$$

We define the torque N about the origin as the *bivector*

$$N = \boldsymbol{x} \wedge \boldsymbol{f}, \tag{3.11}$$

so that the torque and angular momentum are related by

$$\frac{dL}{dt} = N. \tag{3.12}$$

The idea of the torque being a bivector is also natural as torques act over a plane. The plane in question is defined by the vector $\boldsymbol{f}$ and the chosen origin, so both L and N depend on the origin. Recall also that bivectors are additive, much like vectors, so the result of applying two torques is found by adding the respective bivectors.

The angular momentum bivector can be written in an alternative way by first defining $r = |\boldsymbol{x}|$ and writing

$$\boldsymbol{x} = r\hat{\boldsymbol{x}}. \tag{3.13}$$

We therefore have

$$\dot{\boldsymbol{x}} = \frac{d}{dt}(r\hat{\boldsymbol{x}}) = \dot{r}\hat{\boldsymbol{x}} + r\dot{\hat{\boldsymbol{x}}}, \tag{3.14}$$

so that

$$L = m\boldsymbol{x} \wedge (\dot{r}\hat{\boldsymbol{x}} + r\dot{\hat{\boldsymbol{x}}}) = mr\hat{\boldsymbol{x}} \wedge (\dot{r}\hat{\boldsymbol{x}} + r\dot{\hat{\boldsymbol{x}}}) = mr^2\, \hat{\boldsymbol{x}} \wedge \dot{\hat{\boldsymbol{x}}}. \tag{3.15}$$

But since $\hat{\boldsymbol{x}}^2 = 1$ we must have

$$0 = \frac{d}{dt}\hat{\boldsymbol{x}}^2 = 2\hat{\boldsymbol{x}} \cdot \dot{\hat{\boldsymbol{x}}}. \tag{3.16}$$

We can therefore eliminate the outer product in equation (3.15) and write

$$L = mr^2 \hat{\boldsymbol{x}}\dot{\hat{\boldsymbol{x}}} = -mr^2 \dot{\hat{\boldsymbol{x}}}\hat{\boldsymbol{x}}, \tag{3.17}$$

which is useful in a number of problems.

3.1.2 Systems of particles

The preceding definitions generalise easily to systems of particles. For these it is convenient to distinguish between internal and external forces, so the force on the ith particles is

$$\sum_j \boldsymbol{f}_{ji} + \boldsymbol{f}_i^e = \dot{\boldsymbol{p}}_i. \tag{3.18}$$

Here $\boldsymbol{f}_i^e$ is the external force and $\boldsymbol{f}_{ij}$ is the force on the jth particle due to the ith particle. We assume that $\boldsymbol{f}_{ii} = 0$. Newton's third law (in its weak form) states that

$$\boldsymbol{f}_{ij} = -\boldsymbol{f}_{ji}. \tag{3.19}$$

This is not obeyed by all forces, but is assumed to hold for the forces considered in this chapter. Summing the force equation over all particles we find that

$$\sum_i m_i \boldsymbol{a}_i = \sum_i \boldsymbol{f}_i^e + \sum_{i,j} \boldsymbol{f}_{ij} = \sum_i \boldsymbol{f}_i^e. \tag{3.20}$$

All of the internal forces cancel as a consequence of the third law. We define the centre of mass $\boldsymbol{X}$ by

$$\boldsymbol{X} = \frac{1}{M} \sum_i m_i \boldsymbol{x}_i, \tag{3.21}$$

where M is the total mass

$$M = \sum_i m_i. \tag{3.22}$$

The position of the centre of mass is governed by the force law

$$M \frac{d^2 \boldsymbol{X}}{dt^2} = \sum_i \boldsymbol{f}_i^e = \boldsymbol{f}^e \tag{3.23}$$

and so only responds to the total external force on the system. The total momentum of the system is defined by

$$\boldsymbol{P} = \sum_i \boldsymbol{p}_i = M \frac{d\boldsymbol{X}}{dt} \tag{3.24}$$

and is conserved if the total external force is zero.

The total angular momentum about the chosen origin is found by summing the individual bivector contributions,

$$L = \sum_i \boldsymbol{x}_i \wedge \boldsymbol{p}_i. \tag{3.25}$$

The rate of change of L is governed by

$$\dot{L} = \sum_i \boldsymbol{x}_i \wedge \dot{\boldsymbol{p}}_i = \sum_i \boldsymbol{x}_i \wedge \boldsymbol{f}_i^e + \sum_{i,j} \boldsymbol{x}_i \wedge \boldsymbol{f}_{ji}. \tag{3.26}$$

The final term is a double sum containing pairs of terms going as

$$\boldsymbol{x}_i \wedge \boldsymbol{f}_{ji} + \boldsymbol{x}_j \wedge \boldsymbol{f}_{ij} = (\boldsymbol{x}_i - \boldsymbol{x}_j) \wedge \boldsymbol{f}_{ji}. \tag{3.27}$$

The *strong* form of Newton's third law states that the interparticle force $\boldsymbol{f}_{ij}$ is directed along the vector $\boldsymbol{x}_i - \boldsymbol{x}_j$ between the two particles. This law is obeyed by a sufficiently wide range of forces to make it a useful restriction. (The most notable exception to this law is electromagnetism.) Under this restriction the total angular momentum satisfies

$$\frac{dL}{dt} = N^e, \tag{3.28}$$

where N^e is the total external torque. If the applied external torque is zero, and the strong law of action and reaction is obeyed, then the total angular momentum is conserved.

A useful expression for the angular momentum is obtained by introducing a set of position vectors relative to the centre of mass. We write

$$\boldsymbol{x}_i = \boldsymbol{x}_i' + \boldsymbol{X}, \tag{3.29}$$

so that

$$\sum_i m_i \boldsymbol{x}_i' = 0. \tag{3.30}$$

The velocity of the ith particle is now

$$\boldsymbol{v}_i = \boldsymbol{v}_i' + \boldsymbol{v}, \tag{3.31}$$

where $\boldsymbol{v} = \dot{\boldsymbol{X}}$ is the velocity of the centre of mass. The total angular momentum contains four terms:

$$L = \sum_i \big(\boldsymbol{X}\wedge m_i \boldsymbol{v} + \boldsymbol{x}_i'\wedge m_i \boldsymbol{v}_i' + m_i \boldsymbol{x}_i'\wedge \boldsymbol{v} + \boldsymbol{X}\wedge m_i \boldsymbol{v}_i'\big). \tag{3.32}$$

The final two terms both contain factors of $\sum m_i \boldsymbol{x}_i'$ and so vanish, leaving

$$L = \boldsymbol{X}\wedge \boldsymbol{P} + \sum_i \boldsymbol{x}_i'\wedge \boldsymbol{p}_i'. \tag{3.33}$$

The total angular momentum is therefore the sum of the angular momentum of the centre of mass about the origin, plus the angular momentum of the system about the centre of mass. In many cases it is possible to chose the origin so that the centre of mass is at rest, in which case L is simply the total angular momentum about the centre of mass. Similar considerations hold for the kinetic energy, and it is straightforward to show that

$$T = \sum_i \tfrac{1}{2} m_i \boldsymbol{v}_i^2 = \tfrac{1}{2} M \boldsymbol{v}^2 + \tfrac{1}{2} \sum_i m_i \boldsymbol{v}_i'^2. \tag{3.34}$$

3.2 Two-body central force interactions

One of the most significant applications of the preceding ideas is to a system of two point masses moving under the influence of each other. The force acting between the particles is directed along the vector between them, and all external forces are assumed to vanish. It follows that both the total momentum $\boldsymbol{P}$ and angular momentum L are conserved.

We suppose that the particles have positions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, and masses m_1 and m_2. Newton's second law for the central force problem takes the form

$$m_1 \ddot{\boldsymbol{x}}_1 = \boldsymbol{f}, \tag{3.35}$$

$$m_2 \ddot{\boldsymbol{x}}_2 = -\boldsymbol{f}, \tag{3.36}$$

where $\boldsymbol{f}$ is the interparticle force. We define the relative separation vector $\boldsymbol{x}$ by

$$\boldsymbol{x} = \boldsymbol{x}_1 - \boldsymbol{x}_2. \tag{3.37}$$

This vector satisfies

$$m_1 m_2 \ddot{\boldsymbol{x}} = (m_1 + m_2)\boldsymbol{f}. \tag{3.38}$$

We accordingly define the *reduced mass* μ by

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}, \tag{3.39}$$

so that the final force equation can be written as

$$\mu \ddot{\boldsymbol{x}} = \boldsymbol{f}. \tag{3.40}$$

The two-body problem has now been reduced to an equivalent single-body equation. The strong form of the third law assumed here means that the force $\boldsymbol{f}$ is directed along $\boldsymbol{x}$, so we can write $\boldsymbol{f}$ as $f\hat{\boldsymbol{x}}$.

We next re-express the total angular momentum in terms of the centre of mass $\boldsymbol{X}$ and the relative vector $\boldsymbol{x}$. We start by writing

$$m_1 \boldsymbol{x}_1 = m_1 \boldsymbol{X} + \mu \boldsymbol{x}, \quad m_2 \boldsymbol{x}_2 = m_2 \boldsymbol{X} - \mu \boldsymbol{x}. \tag{3.41}$$

It follows that the total angular momentum L_t is given by

$$\begin{aligned} L_t &= m_1 \boldsymbol{x}_1 \wedge \dot{\boldsymbol{x}}_1 + m_2 \boldsymbol{x}_2 \wedge \dot{\boldsymbol{x}}_2 \\ &= M\boldsymbol{X} \wedge \dot{\boldsymbol{X}} + \mu \boldsymbol{x} \wedge \dot{\boldsymbol{x}}. \end{aligned} \tag{3.42}$$

We have assumed that there are no external forces acting, so both L_t and $\boldsymbol{P}$ are conserved. It follows that the internal angular momentum is also conserved and we write this as

$$L = \mu \boldsymbol{x} \wedge \dot{\boldsymbol{x}}. \tag{3.43}$$

Since L is constant, the motion of the particles is confined to the L plane. The trajectory of $\boldsymbol{x}$ must also sweep out area at a constant rate, since this is how L is defined. For planetary motion this is Kepler's second law, though he did not state it in quite this form. Kepler treated the sun as the origin, whereas L should be defined relative to the centre of mass.

The internal kinetic energy is

$$T = \tfrac{1}{2}\mu \dot{\boldsymbol{x}}^2 = \tfrac{1}{2}\mu(\dot{r}\hat{\boldsymbol{x}} + r\dot{\hat{\boldsymbol{x}}})^2 = \tfrac{1}{2}\mu \dot{r}^2 + \tfrac{1}{2}\mu r^2 \dot{\hat{\boldsymbol{x}}}^2. \tag{3.44}$$

From equation (3.17) we see that

$$L^2 = -\mu^2 r^4 \hat{\boldsymbol{x}}\dot{\hat{\boldsymbol{x}}}\dot{\hat{\boldsymbol{x}}}\hat{\boldsymbol{x}} = -\mu^2 r^4 \dot{\hat{\boldsymbol{x}}}^2. \tag{3.45}$$

We therefore define the constant l as the magnitude of L, so

$$l = \mu r^2 |\dot{\hat{\boldsymbol{x}}}|. \tag{3.46}$$

The kinetic energy can now be written as a function of r and $\dot{r}$ only:

$$T = \frac{\mu \dot{r}^2}{2} + \frac{l^2}{2\mu r^2}. \tag{3.47}$$

The force $\boldsymbol{f}$ is conservative and can be written in terms of a potential $V(r)$ as

$$\boldsymbol{f} = f\hat{\boldsymbol{x}} = -\boldsymbol{\nabla} V(r), \tag{3.48}$$

where

$$f = -\frac{dV}{dr}. \tag{3.49}$$

Since the force is conservative the total energy is conserved, so

$$E = \frac{\mu \dot{r}^2}{2} + \frac{l^2}{2\mu r^2} + V(r) \tag{3.50}$$

is a constant. For a given potential $V(r)$ this equation can be integrated to find the evolution of r. The full motion can then be recovered from L.

3.2.1 Inverse-square forces

The most important example of a two-body central force interaction is that described by an inverse-square force law. This case is encountered in gravitation and electrostatics and has been analysed in considerable detail by many authors (see the end of this chapter for suggested additional reading). In this section we review some of the key features of this system, highlighting the places where geometric algebra offers something new. An alternative approach to this problem is discussed in section 3.3.

Writing $f = -k/r^2$ the basic equation to solve is

$$\mu \ddot{\boldsymbol{x}} = -\frac{k}{r^2}\hat{\boldsymbol{x}} = -\frac{k}{r^3}\boldsymbol{x}. \tag{3.51}$$

The sign of k determines whether the force is attractive or repulsive (positive for attractive). This is a second-order vector differential equation, so we expect there to be two constant vectors in the solution — one for the initial position and one for the velocity. We already know that the angular momentum L is a constant of motion, and we can write this as

$$L = \mu r^2 \hat{\boldsymbol{x}}\dot{\hat{\boldsymbol{x}}} = -\mu r^2 \dot{\hat{\boldsymbol{x}}}\hat{\boldsymbol{x}}. \tag{3.52}$$

It follows that

$$L\dot{\boldsymbol{v}} = -\frac{k}{\mu r^2} L\hat{\boldsymbol{x}} = k\dot{\hat{\boldsymbol{x}}}, \tag{3.53}$$

which we can write in the form

$$\frac{d}{dt}(L\boldsymbol{v} - k\hat{\boldsymbol{x}}) = 0. \tag{3.54}$$

Eccentricity	Energy	Orbit
$e > 1$	$E > 0$	Hyperbola
$e = 1$	$E = 0$	Parabola
$e < 1$	$E < 0$	Ellipse
$e = 0$	$E = -\mu k^2/(2l^2)$	Circle

Table 3.1 *Classification of orbits for an inverse-square force law.*

The motion is therefore described by the simple equation

$$L\boldsymbol{v} = k(\hat{\boldsymbol{x}} + \boldsymbol{e}), \tag{3.55}$$

where the *eccentricity vector* $\boldsymbol{e}$ is a second vector constant of motion. This vector is also known in various contexts as the Laplace vector and as the Runge–Lenz vector. From its definition we can see that $\boldsymbol{e}$ must lie in the L plane.

To find a direct equation for the trajectory we first write

$$L\boldsymbol{v}\boldsymbol{x} = L(\boldsymbol{v}\cdot\boldsymbol{x} + \boldsymbol{v}\wedge\boldsymbol{x}) = \frac{1}{\mu}L\tilde{L} + \boldsymbol{v}\cdot\boldsymbol{x}\, L = k(r + \boldsymbol{e}\boldsymbol{x}). \tag{3.56}$$

The scalar part of this equation gives

$$r = \frac{l^2}{k\mu(1 + \boldsymbol{e}\cdot\hat{\boldsymbol{x}})}. \tag{3.57}$$

This equation specifies a conic surface in three dimensions with symmetry axis $\boldsymbol{e}$. The surface is formed by rotating a two-dimensional conic about this axis. Since the motion takes place entirely within the L plane the motion is described by a conic. That is, the trajectory $\boldsymbol{x}(t)$ is one of a hyperbola, parabola, ellipse or circle. The generic cases are ellipses for bound orbits and hyperbolae for free states. The cases of parabolic and circular orbits are exceptional as they require precise values of $|\boldsymbol{e}|$ (table 3.1).

In L and $\boldsymbol{e}$ we have found five of the six constants of motion (we only have two arbitrary constants in $\boldsymbol{e}$ as it is constrained to lie in the L plane). The final constant specifies where on the conic we start at time $t = 0$. We know that the energy is also a constant of motion, so it should be possible to express the energy directly in terms of L and $\boldsymbol{e}$. From equation (3.51) we see that the potential energy must go as k/r, provided we set the arbitrary constant so that $V = 0$ at infinity. The full energy is therefore given by

$$E = \frac{\mu}{2}\boldsymbol{v}^2 - \frac{k}{r}. \tag{3.58}$$

To simplify this we first form

$$L\boldsymbol{v}\boldsymbol{v}\tilde{L} = l^2 v^2 = k^2(\hat{\boldsymbol{x}} + \boldsymbol{e})^2. \tag{3.59}$$

It follows that

$$E = \frac{\mu k^2}{2l^2}(e^2 + 1 + 2\hat{\boldsymbol{x}}\cdot\boldsymbol{e}) - \frac{k}{r} = \frac{\mu k^2}{2l^2}(e^2 - 1), \tag{3.60}$$

where $e = |\boldsymbol{e}|$ is the eccentricity. The sign of the energy is governed entirely by e. Since the potential is set to zero at infinity, all bound states must have negative energy and hence an eccentricity $e < 1$. The limiting case of $e = 1$ describes a parabola (table 3.1).

3.2.2 Motion in time for elliptic orbits

Many methods can be used to find the trajectory as a function of time and these are discussed widely in the literature. Here we describe one of the simplest, which serves to highlight the essential difficulty of this problem. An alternative solution, which more fully exploits the techniques of geometric algebra, is described in section 3.3. From the energy equation we see that

$$\mu^2\dot{r}^2 = 2\mu E - \frac{l^2}{r^2} + \frac{2\mu k}{r}, \tag{3.61}$$

so t is given by

$$t = \mu\int_{r_0}^{r_1} \frac{r\,dr}{(2\mu k r + 2\mu E r^2 - l^2)^{1/2}}. \tag{3.62}$$

Evaluating this integral results in a rather complicated function of r, the general form of which is hard to invert and not very helpful. More useful formulae are obtained by specialising to one form of orbit. For bound problems we are interested in elliptic orbits for which E is negative. For these orbits it is useful to introduce the *semi-major axis* a defined by

$$a = \tfrac{1}{2}(r_1 + r_2) = -\frac{k}{2E}, \tag{3.63}$$

where r_1 and r_2 are the maximum and minimum values of r respectively. In terms of this we can write

$$2\mu k r + 2\mu E r^2 - l^2 = -\frac{\mu k}{a}(r^2 - 2ar) - l^2 = \frac{\mu k}{a}\big(a^2e^2 - (r-a)^2\big). \tag{3.64}$$

We now introduce a new variable Ψ, the *eccentric anomaly*, defined by

$$r = a\big(1 - e\cos(\Psi)\big). \tag{3.65}$$

In terms of this we find

$$t = \left(\frac{\mu a^3}{k}\right)^{1/2}\int_{\Psi_0}^{\Psi_1}\big(1 - e\cos(\Psi)\big)\,d\Psi, \tag{3.66}$$

so if we choose $t = 0$ to correspond to closest approach we have

$$\omega t = \Psi - e\sin(\Psi), \tag{3.67}$$

where

$$\omega^2 = \frac{k}{\mu a^3}. \tag{3.68}$$

Equations (3.65) and (3.67) provide a parametric solution relating r and t. This solution highlights the fact that the equation relating t and r is transcendental and does not have a simple closed form. The time taken for one orbit is $2\pi/\omega$, so the orbital period τ is related to the major axis a by

$$\tau^2 = \frac{4\pi^2\mu}{k} a^3. \tag{3.69}$$

This gives us the third of Kepler's three laws of planetary motion, that the square of the period is proportional to the cube of the major axis.

3.3 Celestial mechanics and perturbations

By far the most important application of the Newtonian theory of gravitation is to the motion of the planets in the solar system. This is a complicated subject of considerable historical and current importance, and we will only touch on a few applications. Detailed calculation of the motions of all of the planets in the solar system still represents a major computational challenge. Aside from the obvious problem of having to calculate the gravitational effects of every planet on every other planet, further effects must also be incorporated. These can include deviations of the shapes of the planets from spherical, the effects of tidal forces and ultimately general relativistic corrections.

A significant number of problems in celestial mechanics are best treated using perturbation theory. In this technique orbits are calculated as a series of ever smaller deviations from Kepler orbits. Since the Kepler orbit is specified entirely by L and $\boldsymbol{e}$, we should first form equations for these in the presence of a perturbing force. We modify the force law to read

$$\mu\ddot{\boldsymbol{x}} = -\frac{k}{r^3}\boldsymbol{x} + \boldsymbol{f}, \tag{3.70}$$

and assume that $\boldsymbol{f}$ is always small compared with the inverse-square term. The angular momentum L now satisfies

$$\dot{L} = \boldsymbol{x}\wedge\boldsymbol{f}, \tag{3.71}$$

so L is now only conserved if $\boldsymbol{f}$ is also a central force. With the eccentricity vector still defined by equation (3.55), we find that

$$k\dot{\boldsymbol{e}} = \dot{L}\cdot\boldsymbol{v} + \frac{1}{\mu}L\cdot\boldsymbol{f}. \tag{3.72}$$

Only five of the six equations for L and $\boldsymbol{e}$ are independent, as we always have $L\wedge\boldsymbol{e} = 0$.

For many problems the variation in L and $\boldsymbol{e}$ is slow compared to the orbital period. For these a useful approximation is obtained by finding the orbital average of $\boldsymbol{f}$ over one cycle, with L and $\boldsymbol{e}$ held constant. The quantities L and $\boldsymbol{e}$ are then assumed to vary slowly under the influence of the time-averaged force. Results for the orbital averages of numerous quantities can be found tabulated in many textbooks and are discussed in the exercises at the end of this chapter.

3.3.1 Example — general relativistic perturbations

Later in this book we will study how general relativity modifies the Newtonian view of gravity. For particles moving in a central potential, the modification is quite simple and can be handled efficiently using perturbation theory. The modified force law is

$$\ddot{\boldsymbol{x}} = -\frac{GM}{r^2}\left(1 + \frac{3l^2}{\mu^2 c^2 r^2}\right)\hat{\boldsymbol{x}}, \tag{3.73}$$

where c is the speed of light and we have replaced k by the gravitational expression $GM\mu$. (A small subtlety is that the derivatives here are with respect to proper time, but this does not affect our reasoning.) The force is still central, so the angular momentum L is still conserved. The eccentricity vector satisfies the simple equation

$$\dot{\boldsymbol{e}} = \frac{3l^2}{\mu^3 c^2 r^4} L\cdot\hat{\boldsymbol{x}}. \tag{3.74}$$

For bound orbits this gives rise to a precession of the major axis (see figure 3.2). The quantity of most interest is the amount $\boldsymbol{e}$ changes in one orbit. To get an approximate result for this we use the time-averaging idea and assume that the orbit is precisely elliptical. We therefore have

$$\Delta\boldsymbol{e} = -\frac{3l^2}{\mu^3 c^2} L \int_0^T dt \frac{\hat{\boldsymbol{x}}}{r^4}, \tag{3.75}$$

where T is the orbital period. Evaluating this integral is left as an exercise, and the final result is

$$\Delta\boldsymbol{e} = \frac{6\pi GM}{a(1-e^2)c^2}\boldsymbol{e}\cdot\hat{L}, \tag{3.76}$$

where $\hat{L} = L/l$. This gives a precession of $\boldsymbol{e}$ with the orientation of L, which corresponds to an advance (figure 3.2). For Mercury this gives rise to the famous advance in the perihelion of 43 arcseconds per century, which was finally explained by general relativity.

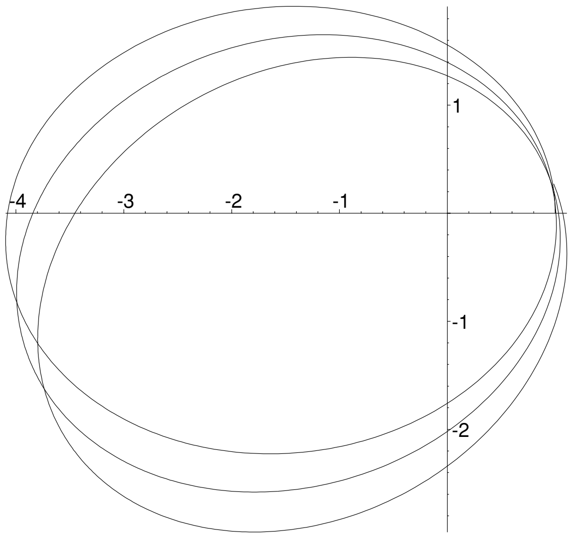

Figure 3.2 *Orbital precession.* The plot shows a modified orbit as predicted by general relativity. The ellipse precesses round in the same direction as the orbital motion. The parameters have been chosen to exaggerate the precession effect.

3.3.2 Spinor equations

An alternative method for analysing the Kepler problem is through the use of 'spinors'. These will be defined more carefully in later chapters, but in two and three dimensions they can be viewed as elements of the subalgebra of $\mathcal{G}_2$ and $\mathcal{G}_3$ consisting entirely of even elements. In two dimensions a spinor can therefore be identified with a complex number. The position vector $\boldsymbol{x}$ in two dimensions can be formed through a rotation and dilation via the polar decomposition

$$\boldsymbol{x} = \mathsf{e}_1 r \exp(\theta \mathsf{e}_1\mathsf{e}_2) = r\exp(-\theta\mathsf{e}_1\mathsf{e}_2)\mathsf{e}_1, \tag{3.77}$$

where $\{\mathsf{e}_1, \mathsf{e}_2\}$ denote a right-handed orthonormal frame and we assume that the vector lies in the $\mathsf{e}_1\mathsf{e}_2$ plane. We know from chapter 2 that the rotation formula only extends to higher dimensions if a double-sided prescription is adopted, so we write the vector $\boldsymbol{x}$ as

$$\boldsymbol{x} = U\mathsf{e}_1U^\dagger = U^2\mathsf{e}_1 = \mathsf{e}_1{U^\dagger}^2. \tag{3.78}$$

In writing this we have placed all of the dynamics in the complex number U.

For the Kepler problem it turns out that the equation for U is considerably easier than that for $\boldsymbol{x}$. We assume that the plane of L is given by $\mathsf{e}_1\mathsf{e}_2$, and start by forming

$$r = |\boldsymbol{x}| = UU^\dagger. \tag{3.79}$$

(Recall that, for a scalar + bivector combination in two dimensions, the reverse operator is the same as complex conjugation.) On differentiating we find that

$$\dot{\boldsymbol{x}} = 2\dot{U}U\mathsf{e}_1, \tag{3.80}$$

hence

$$2r\dot{U} = \dot{\boldsymbol{x}}\mathsf{e}_1 U^\dagger = \dot{\boldsymbol{x}}U\mathsf{e}_1. \tag{3.81}$$

We now introduce the new variable s defined by

$$\frac{d}{ds} = r\frac{d}{dt}, \qquad \frac{dt}{ds} = r. \tag{3.82}$$

In terms of this

$$2\frac{dU}{ds} = \dot{\boldsymbol{x}}U\mathsf{e}_1 \tag{3.83}$$

and

$$2\frac{d^2U}{ds^2} = r\ddot{\boldsymbol{x}}U\mathsf{e}_1 + \dot{\boldsymbol{x}}\frac{dU}{ds}\mathsf{e}_1 = U\big(\ddot{\boldsymbol{x}}\boldsymbol{x} + \tfrac{1}{2}\dot{\boldsymbol{x}}^2\big). \tag{3.84}$$

Now suppose we have motion in a central inverse-square force:

$$\mu\ddot{\boldsymbol{x}} = -k\frac{\boldsymbol{x}}{r^3}. \tag{3.85}$$

The equation for U becomes

$$\frac{d^2U}{ds^2} = \frac{1}{2\mu}U\left(\tfrac{1}{2}\mu\dot{\boldsymbol{x}}^2 - \frac{k}{r}\right) = \frac{E}{2\mu}U, \tag{3.86}$$

which is simply the equation for harmonic motion! This has a number of advantages. First of all, the equation is easy to solve. If we set

$$\omega^2 = -\frac{E}{2\mu} \tag{3.87}$$

then the general solution is

$$U = A\exp(\hat{L}\omega s) + B\exp(-\hat{L}\omega s), \tag{3.88}$$

where A and B are constants and $\hat{L}$ is the unit bivector for the plane of motion. The motion is illustrated in figure 3.3. The particle trajectory maps out an ellipse with the origin at one focus, whereas U defines an ellipse with the origin at the centre. The particle completes two orbits for each full cycle of U.

Further advantages of formulating the dynamics in terms of U are that the equation for U is linear, so is better suited to perturbation theory, and that there is no singularity at $r = 0$, which provides better numerical stability. (Removing this singularity is called ‘regularization’.) In addition, equation (3.86) is universal

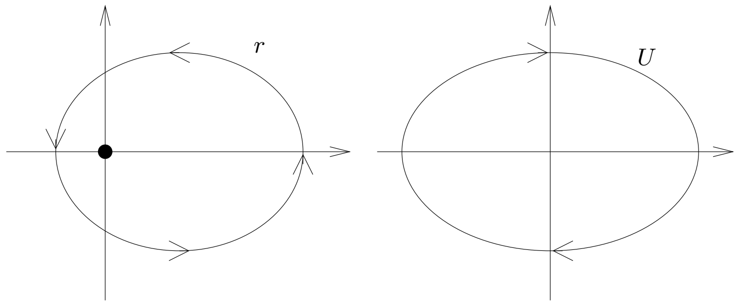

Figure 3.3 *Solution to the Kepler problem.* The particle orbit is shown on the left, and the corresponding spinor on the right. The particle completes two orbits every time U completes one cycle, since U and $-U$ describe the same position.

— it holds for $E > 0$ and $E < 0$. The solution when $E > 0$ simply has trigonometric functions replaced by exponentials. This universality is important, because perturbations can often send bound orbits into unbound ones.

For the method to be truly powerful, however, it must extend to three dimensions. The relevant formula in three dimensions is

$$\boldsymbol{x} = U\mathsf{e}_1 U^\dagger, \tag{3.89}$$

where U is a general even element. This means that U has four degrees of freedom now, whereas only three are required to specify $\boldsymbol{x}$. We are therefore free to impose a further additional constraint on U, which we will use to ensure the equations take on a convenient form. The quantity $UU^\dagger$ is still a scalar in three dimensions, so we have

$$r = UU^\dagger = U^\dagger U. \tag{3.90}$$

We next form $\dot{\boldsymbol{x}}$:

$$\dot{\boldsymbol{x}} = \dot{U}\mathsf{e}_1 U^\dagger + U\mathsf{e}_1 \dot{U}^\dagger. \tag{3.91}$$

We would like this to equal $2\dot{U}\mathsf{e}_1 U^\dagger$ for the preceding analysis to follow through. For this to hold we require

$$\dot{U}\mathsf{e}_1 U^\dagger - U\mathsf{e}_1 \dot{U}^\dagger = \dot{U}\mathsf{e}_1 U^\dagger - (\dot{U}\mathsf{e}_1 U^\dagger)^\dagger = 0. \tag{3.92}$$

The quantity $\dot{U}\mathsf{e}_1 U^\dagger$ only contains odd grade terms (grade-1 and grade-3). If we subtract its reverse, all that remains is the trivector (pseudoscalar) term. We therefore require that

$$\langle \dot{U}\mathsf{e}_1 U^\dagger \rangle_3 = 0, \tag{3.93}$$

which we adopt as our extra condition on U. With this condition satisfied we have

$$2\frac{dU}{ds} = \dot{\boldsymbol{x}}U\mathsf{e}_1 \tag{3.94}$$

and

$$2\frac{d^2U}{ds^2} = \big(\ddot{\boldsymbol{x}}\boldsymbol{x} + \tfrac{1}{2}\dot{\boldsymbol{x}}^2\big)U. \tag{3.95}$$

For an inverse-square force law we therefore recover the same harmonic oscillator equation. In the presence of a perturbing force we have

$$2\mu\frac{d^2U}{ds^2} - EU = \boldsymbol{f}\boldsymbol{x}U = r\boldsymbol{f}U\mathsf{e}_1. \tag{3.96}$$

This equation for U can be handled using standard techniques from perturbation theory. The equation was first found (in matrix form) by Kustaanheimo and Stiefel in 1964 . The analysis was refined and cast in its present form by Hestenes (1999).

3.4 Rotating systems and rigid-body motion

Rigid bodies can be viewed as another example of a system of particles, where now the effect of the internal forces is to keep all of the interparticle distances fixed. For such systems the internal forces can be ignored once one has found a set of dynamical variables that enforce the rigid-body constraint. The problem then reduces to solving for the motion of the centre of mass and for the angular momentum in the presence of any external forces or torques. Suitable variables are a vector $\boldsymbol{x}(t)$ for the centre of mass, and a set of variables to describe the attitude of the rigid body in space. Many forms exist for the latter variables, but here we will concentrate on parameterising the attitude of the rigid body with a *rotor*. Before applying this idea to rigid-body motion, we first look at the description of rotating frames with rotors.

3.4.1 Rotating frames

Suppose that the frame of vectors $\{\mathsf{f}_k\}$ is rotating in space. These can be related to a fixed orthonormal frame $\{\mathsf{e}_k\}$ by the time-dependent rotor $R(t)$:

$$\mathsf{f}_k(t) = R(t)\mathsf{e}_k R^\dagger(t). \tag{3.97}$$

The angular velocity vector $\boldsymbol{\omega}$ is traditionally defined by the formula

$$\dot{\mathsf{f}}_k = \boldsymbol{\omega}\times\mathsf{f}_k, \tag{3.98}$$

where the cross denotes the vector cross product. From section 2.4.3 we know that the cross product is related to the inner product with a bivector by

$$\boldsymbol{\omega}\times\mathsf{f}_k = (-I\boldsymbol{\omega})\cdot\mathsf{f}_k = \mathsf{f}_k\cdot(I\boldsymbol{\omega}). \tag{3.99}$$

We are now used to the idea that angular momentum is best viewed as a bivector, and we must expect the same to be true for angular velocity. We therefore define the angular velocity bivector Ω by

$$\Omega = I\boldsymbol{\omega}. \tag{3.100}$$

This choice ensures that the rotation has the orientation implied by Ω.

To see how Ω is related to the rotor R we start by differentiating equation (3.97):

$$\dot{\mathsf{f}}_k = \dot{R}\mathsf{e}_k R^\dagger + R\mathsf{e}_k\dot{R}^\dagger = \dot{R}R^\dagger\mathsf{f}_k + \mathsf{f}_k R\dot{R}^\dagger. \tag{3.101}$$

From the normalisation equation $RR^\dagger = 1$ we find that

$$0 = \frac{d}{dt}(RR^\dagger) = \dot{R}R^\dagger + R\dot{R}^\dagger. \tag{3.102}$$

Since differentiation and reversion are interchangeable operations we now have

$$\dot{R}R^\dagger = -R\dot{R}^\dagger = -(\dot{R}R^\dagger)^\dagger. \tag{3.103}$$

The quantity $\dot{R}R^\dagger$ is equal to minus its own reverse and has even grade, so must be a pure bivector. The equation for f_k now becomes

$$\dot{\mathsf{f}}_k = \dot{R}R^\dagger\mathsf{f}_k - \mathsf{f}_k\dot{R}R^\dagger = (2\dot{R}R^\dagger)\cdot\mathsf{f}_k. \tag{3.104}$$

Comparing this with equation (3.99) and equation (3.100) we see that $2\dot{R}R^\dagger$ must equal minus the angular velocity bivector Ω, so

$$2\dot{R}R^\dagger = -\Omega. \tag{3.105}$$

The dynamics is therefore contained in the single *rotor equation*

$$\dot{R} = -\tfrac{1}{2}\Omega R. \tag{3.106}$$

The reversed form of this is also useful:

$$\dot{R}^\dagger = \tfrac{1}{2}R^\dagger\Omega. \tag{3.107}$$

Equations of this type are surprisingly ubiquitous in physics. In the more general setting, rotors are viewed as elements of a *Lie group*, and the bivectors form their *Lie algebra*. We will have more to say about this in chapter 11.

3.4.2 Constant Ω

For the case of constant Ω equation (3.106) integrates immediately to give

$$R = \mathrm{e}^{-\Omega t/2}R_0, \tag{3.108}$$

which is the rotor for a constant frequency rotation in the positive sense in the Ω plane. The frame rotates according to

$$\mathsf{f}_k(t) = \mathrm{e}^{-\Omega t/2}R_0\mathsf{e}_k R_0^\dagger\mathrm{e}^{\Omega t/2}. \tag{3.109}$$

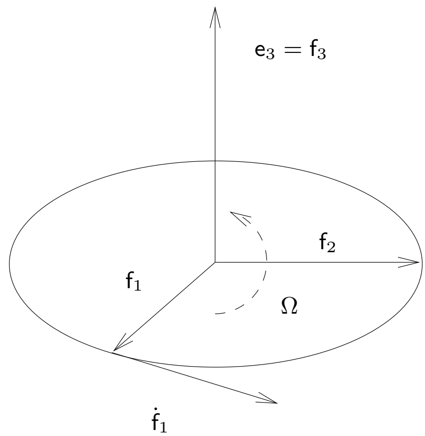

Figure 3.4 *Orientation of the angular velocity bivector.* Ω has the orientation of $\mathsf{f}_1 \wedge \dot{\mathsf{f}}_1$. It must therefore have orientation $+\mathsf{e}_1 \wedge \mathsf{e}_2$ when $\boldsymbol{\omega} = \mathsf{e}_3$.

The constant term R_0 describes the orientation of the frame at $t = 0$, relative to the $\{\mathsf{e}_k\}$ frame.

As an example, consider the case of motion about the e_3 axis (figure 3.4). We have

$$\Omega = \omega I \mathsf{e}_3 = \omega \mathsf{e}_1 \mathsf{e}_2, \tag{3.110}$$

and for convenience we set $R_0 = 1$. The motion is described by

$$\mathsf{f}_k(t) = \exp\Big(-\tfrac{1}{2}\mathsf{e}_1\mathsf{e}_2\omega t\Big)\mathsf{e}_k \exp\Big(\tfrac{1}{2}\mathsf{e}_1\mathsf{e}_2\omega t\Big), \tag{3.111}$$

so that the f_1 axis rotates as

$$\mathsf{f}_1 = \mathsf{e}_1 \exp(\mathsf{e}_1\mathsf{e}_2\omega t) = \cos(\omega t)\mathsf{e}_1 + \sin(\omega t)\mathsf{e}_2. \tag{3.112}$$

This defines a *right-handed* (anticlockwise) rotation in the $\mathsf{e}_1\mathsf{e}_2$ plane, as prescribed by the orientation of Ω.

3.4.3 Rigid-body motion

Suppose that a rigid body is moving through space. To describe the position in space of any part of the body, we need to specify the position of the centre of mass, and the vector to the point in the body from the centre of mass. The latter can be encoded in terms of a rotation from a fixed 'reference' body onto the body in space (figure 3.5). We let $\boldsymbol{x}_0$ denote the position of the centre of mass and $\boldsymbol{y}_i(t)$ denote the position (in space) of a point in the body. These are related by

$$\boldsymbol{y}_i(t) = R(t)\boldsymbol{x}_i R^\dagger(t) + \boldsymbol{x}_0(t), \tag{3.113}$$

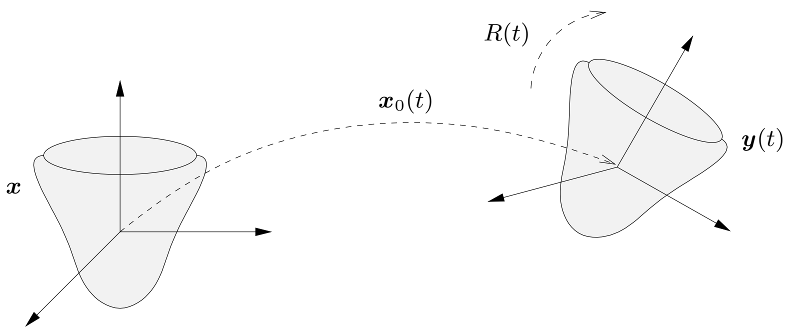

Figure 3.5 *Description of a rigid body.* The vector $\boldsymbol{x}_0(t)$ specifies the position of the centre of mass, relative to the origin. The rotor $R(t)$ defines the orientation of the body, relative to a fixed copy imagined to be placed at the origin. $\boldsymbol{x}$ is a vector in the reference body, and $\boldsymbol{y}$ is the vector in space of the equivalent point on the moving body.

where $\boldsymbol{x}_i$ is a fixed constant vector in the reference copy of the body. In this manner we have placed all of the rotational motion in the time-dependent rotor $R(t)$.

The velocity of the point $\boldsymbol{y} = R\boldsymbol{x}R^\dagger + \boldsymbol{x}_0$ is

$$\begin{aligned}\boldsymbol{v}(t) &= \dot{R}\boldsymbol{x}R^\dagger + R\boldsymbol{x}\dot{R}^\dagger + \dot{\boldsymbol{x}}_0 \\ &= -\tfrac{1}{2}\Omega R\boldsymbol{x}R^\dagger + \tfrac{1}{2}R\boldsymbol{x}R^\dagger\Omega + \boldsymbol{v}_0 \\ &= (R\boldsymbol{x}R^\dagger)\cdot\Omega + \boldsymbol{v}_0,\end{aligned} \tag{3.114}$$

where $\boldsymbol{v}_0$ is the velocity of the centre of mass. The bivector Ω defines the plane of rotation in space. This plane will lie at some orientation relative to the current position of the rigid body. For studying the motion it turns out to be extremely useful to transform the rotation plane back into the fixed, reference copy of the body. Since bivectors are subject to the same rotor transformation law as vectors we define the 'body' angular velocity Ω_B by

$$\Omega_B = R^\dagger \Omega R. \tag{3.115}$$

In terms of the body angular velocity the rotor equation becomes

$$\dot{R} = -\tfrac{1}{2}\Omega R = -\tfrac{1}{2}R\Omega_B, \qquad \dot{R}^\dagger = \tfrac{1}{2}\Omega_B R^\dagger. \tag{3.116}$$

The velocity of the body is now re-expressed as

$$\boldsymbol{v}(t) = R\,\boldsymbol{x}\cdot\Omega_B\,R^\dagger + \boldsymbol{v}_0, \tag{3.117}$$

which will turn out to be the more convenient form. (We have used the operator ordering conventions of section 2.5 to suppress unnecessary brackets in writing $R\,\boldsymbol{x}\cdot\Omega_B\,R^\dagger$ in place of $R(\boldsymbol{x}\cdot\Omega_B)R^\dagger$.)

To calculate the momentum of the rigid body we need the masses of each of the constituent particles. It is easier at this point to go to a continuum approximation and introduce a density $\rho = \rho(\boldsymbol{x})$. The position vector $\boldsymbol{x}$ is taken relative to the centre of mass, so we have

$$\int d^3x\,\rho = M \quad \text{and} \quad \int d^3x\,\rho\boldsymbol{x} = 0. \tag{3.118}$$

The momentum of the rigid body is simply

$$\int d^3x\,\rho\boldsymbol{v} = \int d^3x\,\rho(R\,\boldsymbol{x}\cdot\Omega_B\,R^\dagger + \boldsymbol{v}_0) = M\boldsymbol{v}_0, \tag{3.119}$$

so is specified entirely by the motion of the centre of mass. This is the continuum version of the result of section 3.1.2.

3.4.4 The inertia tensor

The next quantity we require is the angular momentum bivector L for the body about its centre of mass. We therefore form

$$\begin{aligned} L &= \int d^3x\,\rho(\boldsymbol{y} - \boldsymbol{x}_0)\wedge\boldsymbol{v} \\ &= \int d^3x\,\rho(R\boldsymbol{x}R^\dagger)\wedge(R\,\boldsymbol{x}\cdot\Omega_B\,R^\dagger + \boldsymbol{v}_0) \\ &= R\left(\int d^3x\,\rho\boldsymbol{x}\wedge(\boldsymbol{x}\cdot\Omega_B)\right)R^\dagger. \end{aligned} \tag{3.120}$$

The integral inside the brackets refers only to the fixed copy and so defines a time-independent function of Ω_B. This is the reason for working with Ω_B instead of the space angular velocity Ω. We define the *inertia tensor* $\mathcal{I}(B)$ by

$$\mathcal{I}(B) = \int d^3x\,\rho\boldsymbol{x}\wedge(\boldsymbol{x}\cdot B). \tag{3.121}$$

This is a linear function mapping bivectors to bivectors. This way of writing linear functions may be unfamiliar to those used to seeing tensors labelled with indices, but the notation is the natural extension to linear functions of the index-free approach advocated in this book. The linearity of the map is easy to check:

$$\begin{aligned} \mathcal{I}(\lambda A + \mu B) &= \int d^3x\,\rho\boldsymbol{x}\wedge\big(\boldsymbol{x}\cdot(\lambda A + \mu B)\big) \\ &= \int d^3x\,\rho\big(\lambda\boldsymbol{x}\wedge(\boldsymbol{x}\cdot A) + \mu\boldsymbol{x}\wedge(\boldsymbol{x}\cdot B)\big) \\ &= \lambda\mathcal{I}(A) + \mu\mathcal{I}(B). \end{aligned} \tag{3.122}$$

The fact that the inertia tensor maps bivectors to bivectors, rather than vectors to vectors, is also a break from tradition. This viewpoint is very natural given our earlier comments about the merits of bivectors over axial vectors, and provides a

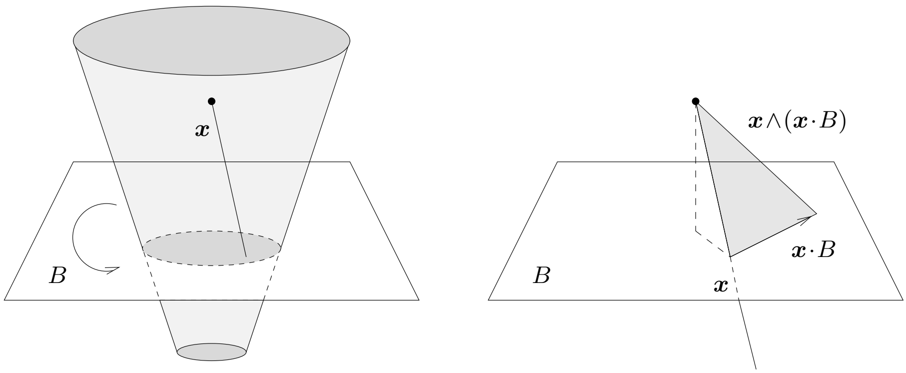

Figure 3.6 *The inertia tensor.* The inertia tensor $\mathcal{I}(B)$ is a linear function mapping its bivector argument B onto a bivector. It returns the total angular momentum about the centre of mass for rotation in the B plane.

clear geometric picture of the tensor (figure 3.6). Since both vectors and bivectors belong to a three-dimensional linear space, there is no additional complexity introduced in this new picture.

To understand the effect of the inertia tensor, suppose that the body rotates in the B plane at a fixed rate $|B|$, and we place the origin at the centre of mass (which is fixed). The velocity of the vector $\boldsymbol{x}$ is simply $\boldsymbol{x}\cdot B$, and the momentum density at this point is $\rho\boldsymbol{x}\cdot B$, as shown in figure 3.6. The angular momentum density bivector is therefore $\boldsymbol{x}\wedge(\rho\boldsymbol{x}\cdot B)$, and integrating this over the entire body returns the total angular momentum bivector for rotation in the B plane.

In general, the total angular momentum will not lie in the same plane as the angular velocity. This is one reason why rigid-body dynamics can often seem quite counterintuitive. When we see a body rotating, our eyes naturally pick out the angular *velocity* by focusing on the vector the body rotates around. Deciding the plane of the angular momentum is less easy, particularly if the internal mass distribution is hidden from us. But it is the angular momentum that responds directly to external torques, not the angular velocity, and this can have some unexpected consequences.

We have calculated the inertia tensor about the centre of mass, but bodies rotating around a fixed axis can be forced to rotate about any point. A useful theorem relates the inertia tensor about an arbitrary point to one about the centre of mass. Suppose that we want the inertia tensor relative to the point $\boldsymbol{a}$, where $\boldsymbol{a}$ is a vector taken from the centre of mass. Returning to the definition of equation (3.121) we see that we need to compute

$$\mathcal{I}_a(B) = \int d^3x\, \rho(\boldsymbol{x}-\boldsymbol{a})\wedge\big((\boldsymbol{x}-\boldsymbol{a})\cdot B\big). \tag{3.123}$$

This integral evaluates to give

$$\begin{aligned}\mathcal{I}_a(B) &= \int d^3x\,\rho\big(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot B) - \boldsymbol{x}\wedge(\boldsymbol{a}\cdot B) - \boldsymbol{a}\wedge(\boldsymbol{x}\cdot B) + \boldsymbol{a}\wedge(\boldsymbol{a}\cdot B)\big) \\ &= \mathcal{I}(B) + M\boldsymbol{a}\wedge(\boldsymbol{a}\cdot B).\end{aligned} \tag{3.124}$$

The inertia tensor relative to $\boldsymbol{a}$ is simply the inertia tensor about the centre of mass, plus the tensor for a point mass M at position $\boldsymbol{a}$.

3.4.5 Principal axes

So far we have only given an abstract specification of the inertia tensor. For most calculations it is necessary to introduce a set of basis vectors fixed in the body. As we are free to choose the directions of these vectors, we should ensure that this choice simplifies the equations of motion as much as possible. To see how to do this, consider the $\{\mathsf{e}_i\}$ frame and define the matrix $\mathcal{I}_{ij}$ by

$$\mathcal{I}_{ij} = -(I\mathsf{e}_i)\cdot\mathcal{I}(I\mathsf{e}_j). \tag{3.125}$$

This defines a *symmetric* matrix, as follows from the result

$$A\cdot(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot B)) = \langle A\boldsymbol{x}(\boldsymbol{x}\cdot B)\rangle = \langle (A\cdot\boldsymbol{x})\boldsymbol{x}B\rangle = B\cdot(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot A)). \tag{3.126}$$

(This sort of manipulation, where one uses the projection onto grade to replace inner and outer products by geometric products, is very common in geometric algebra.) This result ensures that

$$\begin{aligned}\mathcal{I}_{ij} &= -\int d^3x\,\rho(I\mathsf{e}_i)\cdot\big(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot(I\mathsf{e}_j))\big) \\ &= -\int d^3x\,\rho(I\mathsf{e}_j)\cdot\big(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot(I\mathsf{e}_i))\big) = \mathcal{I}_{ji}.\end{aligned} \tag{3.127}$$

It follows that the matrix $\mathcal{I}_{ij}$ will be diagonal if the $\{\mathsf{e}_i\}$ frame is chosen to coincide with the eigendirections of the inertia tensor. These directions are called the *principal axes*, and we always choose our frame along these directions.

The matrix $\mathcal{I}_{ij}$ is also positive-(semi)definite, as can be seen from

$$\begin{aligned}a_i a_j \mathcal{I}_{ij} &= -\int d^3x\,\rho(I\boldsymbol{a})\cdot\big(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot(I\boldsymbol{a}))\big) \\ &= \int d^3x\,\rho\big(\boldsymbol{x}\cdot(I\boldsymbol{a})\big)^2 \geq 0.\end{aligned} \tag{3.128}$$

It follows that all of the eigenvalues of $\mathcal{I}_{ij}$ must be positive (or possibly zero for the case of point or line masses). These eigenvalues are the principal moments of inertia and are crucial in specifying the properties of a rigid body. We denote these $\{i_1, i_2, i_3\}$, so that

$$\mathcal{I}_{jk} = \delta_{jk} i_k \quad \text{(no sum)}. \tag{3.129}$$

(It is more traditional to use a capital I for the moments of inertia, but this symbol is already employed for the pseudoscalar.) If two or three of the principal moments are the same the principal axes are not uniquely specified. In this case one simply chooses one orthonormal set of eigenvectors from the degenerate family of possibilities.

Returning to the index-free presentation, we see that the principal axes satisfy

$$\mathcal{I}(I\mathsf{e}_j) = I\mathsf{e}_k\mathcal{I}_{jk} = i_j I\mathsf{e}_j, \tag{3.130}$$

where again there is no sum implied between eigenvectors and their associated eigenvalue in the final expression. To calculate the effect of the inertia tensor on an arbitrary bivector B we decompose B in terms of the principal axes as

$$B = B_j I\mathsf{e}_j. \tag{3.131}$$

It follows that

$$\mathcal{I}(B) = \sum_{j=1}^{3} i_j B_j I\mathsf{e}_j = i_1 B_1 \mathsf{e}_2\mathsf{e}_3 + i_2 B_2 \mathsf{e}_3\mathsf{e}_1 + i_3 B_3 \mathsf{e}_1\mathsf{e}_2. \tag{3.132}$$

The fact that for most bodies the principal moments are not equal demonstrates that $\mathcal{I}(B)$ will not lie in the same plane as B, unless B is perpendicular to one of the principal axes.

A useful result for calculating the inertia tensor is that the principal axes of a body always coincide with symmetry axes, if any are present. This simplifies the calculation of the inertia tensor for a range of standard bodies, the results for which can be found in some of the books listed at the end of this chapter.

3.4.6 Kinetic energy and angular momentum

To calculate the kinetic energy of the body from the velocity of equation (3.114) we form the integral

$$\begin{aligned} T &= \tfrac{1}{2}\int d^3x\, \rho(R\,\boldsymbol{x}\cdot\Omega_B\, R^\dagger + \boldsymbol{v}_0)^2 \\ &= \tfrac{1}{2}\int d^3x\, \rho\big((\boldsymbol{x}\cdot\Omega_B)^2 + 2\boldsymbol{v}_0\cdot(R\,\boldsymbol{x}\cdot\Omega_B\, R^\dagger) + \boldsymbol{v}_0^2\big) \\ &= \tfrac{1}{2}\int d^3x\, \rho(\boldsymbol{x}\cdot\Omega_B)^2 + M\boldsymbol{v}_0^2. \end{aligned} \tag{3.133}$$

Again, there is a clean split into a rotational contribution and a term due to the motion of the centre of mass. Concentrating on the former, we use the manipulation

$$(\boldsymbol{x}\cdot\Omega_B)^2 = \langle \boldsymbol{x}\cdot\Omega_B \boldsymbol{x}\Omega_B\rangle = -\Omega_B\cdot\big(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot\Omega_B)\big) \tag{3.134}$$

to write the rotational contribution as

$$-\tfrac{1}{2}\Omega_B \cdot \Big(\int d^3x\, \rho \boldsymbol{x} \wedge (\boldsymbol{x} \cdot \Omega_B) \Big) = -\tfrac{1}{2}\Omega_B \cdot \mathcal{I}(\Omega_B). \tag{3.135}$$

The minus sign is to be expected because bivectors all have negative squares. The sign can be removed by reversing one of the bivectors to construct a positive-definite product. The total kinetic energy is therefore

$$T = \tfrac{1}{2}M\boldsymbol{v}_0^2 + \tfrac{1}{2}\Omega_B^\dagger \cdot \mathcal{I}(\Omega_B). \tag{3.136}$$

The inertia tensor is constructed from the point of view of the fixed body. From equation (3.120) we see that the angular momentum in space is obtained by rotating the body angular momentum $\mathcal{I}(\Omega_B)$ onto the space configuration, that is,

$$L = R\mathcal{I}(\Omega_B)\, R^\dagger. \tag{3.137}$$

We can understand this expression as follows. Suppose that a body rotates in space with angular velocity Ω. At a given instant we carry out a fixed rotation to align everything back with the fixed reference configuration. This reference copy then has angular velocity $\Omega_B = R^\dagger \Omega R$. The inertia tensor (fixed in the reference copy) returns the angular momentum, given an input angular velocity. The result of this is then rotated forwards onto the body in space, to return L.

The space and body angular velocities are related by $\Omega = R\Omega_B R^\dagger$, so the kinetic energy can be written in the form

$$T = \tfrac{1}{2}M\boldsymbol{v}_0^2 + \tfrac{1}{2}\Omega^\dagger \cdot L. \tag{3.138}$$

We now introduce components $\{\omega_k\}$ for both Ω and Ω_B by writing

$$\Omega = \sum_{k=1}^{3} \omega_k I \mathsf{f}_k, \quad \Omega_B = \sum_{k=1}^{3} \omega_k I \mathsf{e}_k. \tag{3.139}$$

In terms of these we recover the standard expression

$$T = \tfrac{1}{2}M\boldsymbol{v}_0^2 + \sum_{k=1}^{3} \tfrac{1}{2} i_k \omega_k^2. \tag{3.140}$$

3.4.7 Equations of motion

The equations of motion are $\dot{L} = N$, where N is the external torque. The inertia tensor is time-independent since it only refers to the static 'reference' copy of the rigid body, so we find that

$$\begin{aligned}\dot{L} &= \dot{R}\mathcal{I}(\Omega_B)R^\dagger + R\mathcal{I}(\Omega_B)\dot{R}^\dagger + R\mathcal{I}(\dot{\Omega}_B)R^\dagger \\ &= R\big(\mathcal{I}(\dot{\Omega}_B) - \tfrac{1}{2}\Omega_B\mathcal{I}(\Omega_B) + \tfrac{1}{2}\mathcal{I}(\Omega_B)\Omega_B\big)R^\dagger.\end{aligned} \tag{3.141}$$

At this point it is extremely useful to have a symbol to denote one-half of the commutator of two bivectors. The standard symbol for this is the cross, $\times$, so we define the *commutator product* by

$$A\times B = \tfrac{1}{2}(AB - BA). \tag{3.142}$$

This notation does raise the possibility of confusion with the vector cross product, but as the latter is not needed any more this should not pose a problem. The commutator product is so ubiquitous in applications that it needs its own symbol, and the cross is particularly convenient as it correctly conveys the antisymmetry of the product. In section 4.1.3 we prove that the commutator of any two bivectors results in a third bivector. This is easily confirmed in three dimensions by expressing both bivectors in terms of their dual vectors.

With the commutator product at our disposal the equations of motion are now written concisely as

$$\dot{L} = R\big(\mathcal{I}(\dot{\Omega}_B) - \Omega_B\times\mathcal{I}(\Omega_B)\big)R^\dagger. \tag{3.143}$$

The typical form of the rigid-body equations is recovered by expanding in terms of components. In terms of these we have

$$\begin{aligned}\dot{L} &= R\left(\sum_{k=1}^{3} i_k\dot{\omega}_k I\mathsf{e}_k - \sum_{j,k=1}^{3} i_k\omega_j\omega_k (I\mathsf{e}_j)\times(I\mathsf{e}_k)\right)R^\dagger \\ &= \sum_{k=1}^{3}\dot{\omega}_k I\mathsf{f}_k + \sum_{j,k,l=1}^{3}\epsilon_{jkl} i_k\omega_j\omega_k I\mathsf{f}_l.\end{aligned} \tag{3.144}$$

If we let N_k denote the components of the torque N in the rotating f_k frame,

$$N = \sum_{k=1}^{3} N_k I\mathsf{f}_k, \tag{3.145}$$

we recover the Euler equations of motion for a rigid body:

$$\begin{aligned} i_1\dot{\omega}_1 - \omega_2\omega_3(i_2 - i_3) &= N_1, \\ i_2\dot{\omega}_2 - \omega_3\omega_1(i_3 - i_1) &= N_2, \\ i_3\dot{\omega}_3 - \omega_1\omega_2(i_1 - i_2) &= N_3. \end{aligned} \tag{3.146}$$

Various methods can be used to solve these equations and are described in most mechanics textbooks. Here we will simply illustrate some features of the equations, and describe a solution method which does not resort to the explicit coordinate equations.

3.4.8 Torque-free motion

The torque-free equation $\dot{L} = 0$ reduces to

$$\mathcal{I}(\dot{\Omega}_B) - \Omega_B \times \mathcal{I}(\Omega_B) = 0. \tag{3.147}$$

This is a first-order constant coefficient differential equation for the bivector Ω_B. Closed form solutions exist, but before discussing some of these it is useful to consider the conserved quantities. Throughout this section we ignore any overall motion of the centre of mass of the rigid body. Since $\dot{L} = 0$ both the kinetic energy and the magnitude of L are constant. To exploit this we introduce the components

$$L_k = i_k \omega_k, \quad L = \sum_{k=1}^{3} L_k I \mathsf{f}_k. \tag{3.148}$$

These are the components of L in the rotating f_k frame. So, even though L is constant, the components L_k are time-dependent. In terms of these components the magnitude of L is

$$LL^\dagger = L_1^2 + L_2^2 + L_3^3 \tag{3.149}$$

and the kinetic energy is

$$T = \frac{L_1^2}{2i_1} + \frac{L_2^2}{2i_2} + \frac{L_3^2}{2i_3}. \tag{3.150}$$

Both $|L|$ and T are constants of motion, which imposes two constraints on the three components L_k. A useful way to visualise this is to think in terms of a vector $\boldsymbol{l}$ with components L_k:

$$\boldsymbol{l} = \sum_{k=1}^{3} L_k \mathsf{e}_k = -IR^\dagger L R. \tag{3.151}$$

This is the vector perpendicular to $R^\dagger L R$ — a rotating vector in the fixed reference body. Conservation of $|L|$ means that $\boldsymbol{l}$ is constrained to lie on a sphere, and conservation of T restricts $\boldsymbol{l}$ to the surface of an ellipsoid. Possible paths for $\boldsymbol{l}$ for a given rigid body are therefore defined by the intersections of a sphere with a family of ellipsoids (governed by T). For the case of unequal principal moments these orbits are non-degenerate. Examples of these orbits are shown in figure 3.7. This figure shows that orbits around the axes with the smallest and largest principal moments are stable, whereas around the middle axis the orbits are unstable. Any small change in the energy of the body will tend to throw it into a very different orbit if the orbit of $\boldsymbol{l}$ approaches close to e_2.

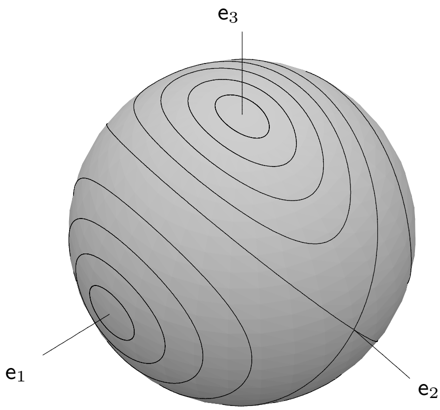

Figure 3.7 *Angular momentum orbits.* The point described by the vector $\boldsymbol{l}$ simultaneously lies on the surface of a sphere and an ellipse. The figure shows possible paths on the sphere for $\boldsymbol{l}$ in the case of $i_1 < i_2 < i_3$, with the 3 axis vertical.

3.4.9 The symmetric top

The full analytic solution for torque-free motion is complicated and requires elliptic functions. If the body has a single symmetry axis, however, the solution is quite straightforward. In this case the body has two equal moments of inertia, $i_1 = i_2$, and the third principal moment i_3 is assumed to be different. With this assignment e_3 is the symmetry axis of the body. The action of the inertia tensor on Ω_B is

$$\begin{aligned}\mathcal{I}(\Omega_B) &= i_1\omega_1\mathsf{e}_2\mathsf{e}_3 + i_1\omega_2\mathsf{e}_3\mathsf{e}_1 + i_3\omega_3\mathsf{e}_1\mathsf{e}_2 \\ &= i_1\Omega_B + (i_3 - i_1)\omega_3 I\mathsf{e}_3,\end{aligned} \tag{3.152}$$

so we can write $\mathcal{I}(\Omega_B)$ in the compact form

$$\mathcal{I}(\Omega_B) = i_1\Omega_B + (i_3 - i_1)(\Omega_B\wedge\mathsf{e}_3)\mathsf{e}_3. \tag{3.153}$$

(This type of expression offers many advantages over the alternative 'dyad' notation.) The torque-free equations of motion are now

$$\mathcal{I}(\dot{\Omega}_B) = \Omega_B\times\mathcal{I}(\Omega_B) = (i_3 - i_1)\Omega_B\times\big((\Omega_B\wedge\mathsf{e}_3)\mathsf{e}_3\big). \tag{3.154}$$

Since $\Omega_B\wedge\mathsf{e}_3$ is a trivector, we can dualise the final term and write

$$\mathcal{I}(\dot{\Omega}_B) = -(i_3 - i_1)\mathsf{e}_3\wedge\big((\Omega_B\wedge\mathsf{e}_3)\Omega_B\big). \tag{3.155}$$

It follows that

$$\mathsf{e}_3\wedge\mathcal{I}(\dot{\Omega}_B) = 0 = i_3\dot{\omega}_3 I, \tag{3.156}$$

which shows that ω_3 is a constant. This result can be read off directly from the Euler equations, but it is useful to see how it can be derived without dropping down to the individual component equations. The ability to do this becomes ever more valuable as the complexity of the equations increases.

Next we use the result that

$$\begin{aligned} i_1\Omega_B &= \mathcal{I}(\Omega_B) - (i_3 - i_1)(\Omega_B \wedge \mathsf{e}_3)\mathsf{e}_3 \\ &= \mathcal{I}(\Omega_B) + (i_1 - i_3)\omega_3 I\mathsf{e}_3 \end{aligned} \tag{3.157}$$

to write

$$\Omega = R\Omega_B R^\dagger = \frac{1}{i_1}L + \frac{i_1 - i_3}{i_1}\omega_3 RI\mathsf{e}_3 R^\dagger. \tag{3.158}$$

Our rotor equation now becomes

$$\dot{R} = -\tfrac{1}{2}\Omega R = -\frac{1}{2i_1}(LR + R(i_1 - i_3)\omega_3 I\mathsf{e}_3). \tag{3.159}$$

The right-hand side of this equation involves two constant bivectors, one multiplying R to the left and the other to the right. We therefore define the two bivectors

$$\Omega_l = \frac{1}{i_1}L, \qquad \Omega_r = \omega_3\frac{i_1 - i_3}{i_1}I\mathsf{e}_3, \tag{3.160}$$

so that the rotor equation becomes

$$\dot{R} = -\tfrac{1}{2}\Omega_l R - \tfrac{1}{2}R\Omega_r. \tag{3.161}$$

This equation integrates immediately to give

$$R(t) = \exp(-\tfrac{1}{2}\Omega_l t)R_0 \exp(-\tfrac{1}{2}\Omega_r t). \tag{3.162}$$

This fully describes the motion of a symmetric top. It shows that there is an 'internal' rotation in the $\mathsf{e}_1\mathsf{e}_2$ plane (the symmetry plane of the body). This is responsible for the precession of a symmetric top. The constant rotor R_0 defines the attitude of the rigid body at $t = 0$ and can be set to 1. The resultant body is then rotated in the plane of its angular momentum to obtain the final attitude in space.

3.5 Notes

Much of this chapter follows *New Foundations for Classical Mechanics* by David Hestenes (1999), which gives a comprehensive account of the applications to classical mechanics of geometric algebra in three dimensions. Readers are encouraged to compare the techniques used in this chapter with more traditional methods, a good description of which can be found in *Classical Mechanics* by Goldstein (1950), or *Analytical Mechanics* by Hand & Finch (1998). The standard reference for the Kustaanheimo–Stiefel equation is *Linear and Regular Celestial Mechanics* by Stiefel and Scheifele (1971). Many authors have explored this technique, particularly in the quaternionic framework. These include Hestenes' 'Celestial mechanics with geometric algebra' (1983) and the papers by Aramanovitch (1995) and Vrbik (1994, 1995).

3.6 Exercises

3.1 An elliptical orbit in an inverse-square force law is parameterised in terms of a scalar + pseudoscalar quantity U by $\boldsymbol{x} = U^2 \mathsf{e}_1$. Prove that U can be written

$$U = A_0 \mathrm{e}^{I\omega s} + B_0 \mathrm{e}^{-I\omega s},$$

where $dt/ds = r$, $r = |\boldsymbol{x}| = UU^\dagger$ and I is the unit bivector for the plane. What is the value of ω? Find the conditions on A_0 and B_0 such that at time $t = 0$, $s = 0$ and the particle lies on the positive e_1 axis with velocity in the positive e_2 direction. For which value of s does the velocity point in the $-\mathsf{e}_1$ direction? Find the values for the shortest and longest diameters of the ellipse, and verify that we can write

$$U = \sqrt{a(1+e)}\cos(\omega s) - \sqrt{a(1-e)} I \sin(\omega s),$$

where e is the eccentricity and a is the semi-major axis.

3.2 For elliptical orbits the semi-major axis a is defined by $a = \frac{1}{2}(r_1 + r_2)$, where r_1 and r_2 are the distances of closest and furthest approach. Prove that

$$\frac{l^2}{k\mu} = a(1 - e^2).$$

Hence show that we can write

$$r = \frac{a(1-e^2)}{1 + e\cos(\theta)},$$

where $e\cos(\theta) = \boldsymbol{e} \cdot \hat{\boldsymbol{x}}$. The eccentricity vector points to the point of closest approach. Why would we expect the orbital average of $\hat{\boldsymbol{x}}/r^4$ to also point in this direction? Prove that

$$\int_0^T dt \frac{\hat{\boldsymbol{x}}}{r^4} = \hat{\boldsymbol{e}} \frac{\mu}{la^2(1-e^2)^2} \int_0^{2\pi} \bigl(1 + e\cos(\theta)\bigr)^2 \cos(\theta)\, d\theta$$

and evaluate the integral.

3.3 A particle in three dimensions moves along a curve $\boldsymbol{x}(t)$ such that $|\boldsymbol{v}|$ is constant. Show that there exists a bivector Ω such that

$$\dot{\boldsymbol{v}} = \Omega \cdot \boldsymbol{v},$$

and give an explicit formula for Ω. Is this bivector unique?

3.4 Suppose that we measure components of the position vector $\boldsymbol{x}$ in a rotating frame $\{\mathsf{f}_i\}$. By referring this frame to a fixed frame, show that the components of $\boldsymbol{x}$ are given by

$$x_i = \mathsf{e}_i \cdot (R^\dagger \boldsymbol{x} R).$$

By differentiating this expression twice, prove that we can write

$$\mathsf{f}_i \ddot{x}_i = \ddot{\boldsymbol{x}} + \Omega \cdot (\Omega \cdot \boldsymbol{x}) + 2\Omega \cdot \dot{\boldsymbol{x}} + \dot{\Omega} \cdot \boldsymbol{x}.$$

Hence deduce expressions for the centrifugal, Coriolis and Euler forces in terms of the angular velocity bivector Ω.

3.5 Show that the inertia tensor satisfies the following properties:

$$\begin{array}{rl} \text{linearity:} & \mathcal{I}(\lambda A + \mu B) = \lambda \mathcal{I}(A) + \mu \mathcal{I}(B) \\ \text{symmetry:} & \langle A \mathcal{I}(B) \rangle = \langle \mathcal{I}(A) B \rangle. \end{array}$$

3.6 Prove that the inertia tensor $\mathcal{I}(B)$ for a solid cylinder of height h and radius a can be written

$$\mathcal{I}(B) = \frac{Mh^2}{12}(B - B \wedge \mathsf{e}_3\, \mathsf{e}_3) + \frac{Ma^2}{4}(B + B \wedge \mathsf{e}_3\, \mathsf{e}_3),$$

where e_3 is the symmetry axis.

3.7 For a torque-free symmetric top prove that the angular momentum, viewed back in the reference copy, rotates around the symmetry axis at an angular frequency ω, where

$$\omega = \omega_3 \frac{i_3 - i_1}{i_1}.$$

Show that the angle between the symmetry axis and the vector $\boldsymbol{l} = -IL$ is given by

$$\cos(\theta) = \frac{i_3 \omega}{l},$$

where $l^2 = \boldsymbol{l}^2 = LL^\dagger$. Hence show that the symmetry axis rotates in space in the L plane at an angular frequency ω', where

$$\omega' = \frac{i_3 \omega_3}{i_1 \cos(\theta)}.$$