3 Newtonian orbital dynamics

In this chapter we apply the tools developed in the previous two chapters to an exploration of the orbital dynamics of bodies subjected to their mutual gravitational attractions. Many aspects of what we learned in Chapters 1 and 2 will be put to good use, and the end result will be considerable insight into the behavior of our own solar system. To be sure, the field of *celestial mechanics* has a rich literature that goes back centuries, and this relatively short chapter will only ~~scrape~~ the surface. We believe, however, that we have sampled the literature well, and selected a good collection of interesting topics. Some of the themes introduced here will be featured in later chapters, when we turn to relativistic aspects of celestial mechanics.

We begin in Sec. 3.1 with a very brief survey of celestial mechanics, from Newton to Einstein. In Sec. 3.2 we give a complete description of Kepler's problem, the specification of the motion of two spherical bodies subjected to their mutual gravity. In Sec. 3.3 we introduce a powerful formalism to treat Keplerian orbits perturbed by external bodies or deformations of the two primary bodies; in this framework of *osculating Keplerian orbits*, the motion is at all times described by a sequence of Keplerian orbits, with constants of the motion that evolve as a result of the perturbation. We shall apply this formalism to a number of different situations, and highlight a number of important processes that take place in the solar system and beyond. In Sec. 3.5 we examine the three-body problem and briefly touch upon the general case of N bodies. We conclude in Sec. 3.6 with a review of the Lagrangian formulation of Newtonian mechanics.

3.1 Celestial mechanics from Newton to Einstein

The triumph of Newton's mechanics and universal gravitation is largely contained in the confrontation with the observed motion of celestial bodies in our solar system, which was initiated by Kepler even before Newton's laws became available. The two-body dynamics of Newton's theory immediately accounted for Kepler's laws, which state that the orbits of bodies in the solar system are ellipses that trace out equal areas in equal times, with periods inversely proportional to the 3/2 power of their diameters. Edmund Halley used Newton's equations to point out that the comets that had been observed in 1531, 1607, and 1682 were actually a single object that orbits the sun with a period ranging from 75 to 76 years. Refined orbital calculations carried out by Alexis Clairaut led to the successful

prediction of the comet's return in 1759. The difficult problem of describing the detailed motion of the Moon, in particular the advance of its perigee, was successfully tackled by Clairaut using Newtonian theory; a first attempt was actually made by Newton himself, but with somewhat incorrect results that discouraged him and motivated him to start a new career at the Royal Mint. Finally, in a triumph of theoretical prediction, Urbain Jean Joseph Le Verrier in France, and independently John Couch Adams in England, pointed out that certain anomalies in the orbit of Uranus could best be explained by the existence of an additional planet, and each astronomer made a rough prediction of where such a planet might be found. In 1846, a day after receiving Le Verrier's prediction for its position, German astronomers discovered the new planet, which is now called Neptune.

This accumulation of successes built such confidence in Newton's theory that when the crisis occurred, the shock was almost palpable. The crisis was caused by Mercury. By the middle of the nineteenth century, astronomers had established that the perihelion of Mercury (the point of closest approach to the Sun) was advancing at a rate of 575 arcseconds per century relative to the fixed stars. Although the two-body solution of Newton's theory requires the perihelion to be fixed in direction, it seemed clear that the advance should be caused by the gravitational influences of the other planets (mostly Venus because of its proximity, and Jupiter because of its large mass) on Mercury's orbit. Fresh from his success with the prediction of Neptune, Le Verrier applied his methods to the problem of Mercury. He calculated the amount that each planet would contribute to Mercury's perihelion advance (see Table 3.1), but the total fell short of the measured value, by an amount comparable to 40 arcseconds per century. The modern value of the discrepancy is 42.98 ± 0.04 arcseconds per century, based upon improved measurements of Mercury's orbit using radar ranging, combined with improved data on the masses and orbits of the other planets, and accurate numerical ephemeris codes for calculating orbits.

The discrepancy could not be attributed to calculational errors or faulty observations, and no viable explanation could be found for the next 50 years. In the spirit of Neptune, Le Verrier and others supported the existence of another planet between Mercury and the Sun, which was given the provisional name Vulcan. But despite systematic astronomical searches, no credible evidence for such a planet was ever discovered. If changing the solar system would not do, perhaps a change of theory might fare better? Simon Newcomb proposed that all would be well with Mercury if the inverse square law of Newtonian gravity were changed to the inverse power of 2.0000001574. Such a change, however, would also contribute to the advance of the lunar perigee, and once improved data on the lunar orbit became available, Newcomb's proposal was shown not to be viable.

The resolution of the problem of Mercury is by now legendary. The place is Berlin, the time November 1915. Albert Einstein, using the new equations of general relativity, calculates the motion of Mercury and shows that the relativistic laws of orbital motion account for the notorious discrepancy. Einstein was overjoyed, and later wrote to a friend that this discovery gave him palpitations of the heart. The modern value for the shift predicted by general relativity, using the best data for all relevant quantities, is 42.98 arcseconds per century, in perfect agreement with the measured value.

Table 3.1 Planetary contributions to Mercury's perihelion advance (in arcseconds per century).	
Planet	Advance
Venus	277.8
Earth	90.0
Mars	2.5
Jupiter	153.6
Saturn	7.3
Total	531.2
Discrepancy	42.9
Modern measured value	42.98 ± 0.04
General relativity prediction	42.98

3.2 Two bodies: Kepler's problem

Kepler's problem is to determine the motion of two bodies subjected to their mutual gravitational attraction, under the assumption that each body can be taken to be spherically symmetric. This is the simplest problem of celestial mechanics, but also one of the most relevant, because to a good first approximation, the motion of any planet around the Sun can be calculated while ignoring the effects of the other planets. It is also a problem that can be solved exactly and completely, in terms of simple functions.

3.2.1 Effective one-body description

The foundations for Kepler's problem were provided back in Sec. 1.6.7. We have a first body of mass m_1, position r_1, velocity $v_1 = dr_1/dt$, and acceleration $a_1 = dv_1/dt$, and a second body of mass m_2, position r_2, velocity $v_2 = dr_2/dt$, and acceleration $a_2 = dv_2/dt$. We place the origin of the coordinate system at the system's barycenter R, so that $m_1 r_1 + m_2 r_2 = 0$. The position of each body is then given by

$$r_1 = \frac{m_2}{m} r, \qquad r_2 = -\frac{m_1}{m} r, \tag{3.1}$$

in which $m := m_1 + m_2$ is the total mass and $r := r_1 - r_2$ the separation between bodies. Similar relations hold between v_1, v_2 and the relative velocity $v := v_1 - v_2 = dr/dt$. The relative acceleration $a := a_1 - a_2 = dv/dt$ is obtained from Eq. (1.220) by removing all terms that involve the multipole moments of each body (which vanish when the bodies are spherical); we have

$$a = -\frac{Gm}{r^2} n, \tag{3.2}$$

where $r := |r|$ is the distance between bodies, and $n := r/r$ a unit vector that points from body 2 to body 1. This is the equation of motion for the relative orbit, and thanks to Eq. (3.1),

its solution is sufficient to determine the individual motion of each body. Equation (3.2) can be interpreted as describing the motion of a fictitious particle at position r in the field of a gravitating center of mass m at $r = 0$; as such our two-body problem has been reformulated as an effective one-body problem. This radical simplification of the original problem, which involved six independent degrees of freedom (the three components of each position vector) instead of the current three (the three components of the separation vector), is a consequence of the conservation of total momentum, which implies that the motion of the barycenter position R is uniform and therefore trivial.

According to Eq. (1.221), the total energy of the two-body system is given by

$$E = \frac{1}{2}\mu v^2 - \frac{G\mu m}{r}, \tag{3.3}$$

in which $\mu := m_1 m_2 / m$ is the reduced mass of the system; to obtain this expression we once more dropped all terms involving multipole moments, and set $V = dR/dt = 0$ in the barycentric frame. The system's total angular momentum is given by Eq. (1.204), which reduces to

$$L = m_1 r_1 \times v_1 + m_2 r_2 \times v_2 = \mu r \times v \tag{3.4}$$

for a two-body system. Because the bodies are assumed to have no spin, and because the angular momentum is contained entirely in the orbital motion, we use the standard notation L instead of J for the angular-momentum vector. It is a simple matter to verify that $dE/dt = 0$ by virtue of Eq. (3.2), and that $dL/dt = 0$ by virtue of the sole fact that the acceleration a is directed along r.

The constancy of the angular-momentum vector has far-reaching consequences on the motion of the bodies. Because L is a constant vector orthogonal to both r and v, the motion must take place in a fixed plane that is at all times perpendicular to L. We therefore have achieved another radical simplification of the problem: by confining the motion to a plane we have eliminated one degree of freedom from the original three associated with the effective one-body problem.

3.2.2 Orbital plane

We take the orbital plane to coincide with the x–y plane of the coordinate system, and we align L with the z-direction. To simplify the notation in subsequent developments, we write

$$L = \mu h, \qquad h := r \times v = h e_z, \tag{3.5}$$

with $h := |h|$ denoting the magnitude of the constant vector $r \times v$.

It is helpful to describe the motion with the polar coordinates r and ϕ, defined such that the components of the separation vector are given by

$$r = [r \cos\phi, r \sin\phi, 0]; \tag{3.6}$$

both r and ϕ depend on time. A vectorial basis in the orbital plane can be built from the constant unit vectors e_x and e_y, but it is useful to introduce also the time-dependent unit

vectors

$$\boldsymbol{n} := [\cos\phi, \sin\phi, 0], \qquad \boldsymbol{\lambda} := [-\sin\phi, \cos\phi, 0], \tag{3.7}$$

which are closely tied to the description of the orbital motion. The vector $\boldsymbol{n} := \boldsymbol{r}/r$ points from body 2 to body 1, while $\boldsymbol{\lambda}$ is orthogonal to it. The vectors satisfy

$$\frac{d\boldsymbol{n}}{d\phi} = \boldsymbol{\lambda}, \qquad \frac{d\boldsymbol{\lambda}}{d\phi} = -\boldsymbol{n}. \tag{3.8}$$

The basis is completed with $\boldsymbol{e}_z$, which is normal to the orbital plane and aligned with the angular-momentum vector.

The vectors $\boldsymbol{r}$, $\boldsymbol{v}$, and $\boldsymbol{a}$ can each be decomposed in the orbital basis. Simple computations involving Eqs. (3.8) produce

$$\boldsymbol{r} = r\,\boldsymbol{n}, \tag{3.9a}$$

$$\boldsymbol{v} = \dot{r}\,\boldsymbol{n} + r\dot{\phi}\,\boldsymbol{\lambda}, \tag{3.9b}$$

$$\boldsymbol{a} = \left(\ddot{r} - r\dot{\phi}^2\right)\boldsymbol{n} + \frac{1}{r}\frac{d}{dt}\left(r^2\dot{\phi}\right)\boldsymbol{\lambda}, \tag{3.9c}$$

in which an overdot indicates differentiation with respect to t.

3.2.3 First integrals

The acceleration of Eq. (3.9) may now be inserted within Eq. (3.2), and the absence of a component along $\boldsymbol{\lambda}$ immediately implies that $r^2\dot{\phi}$ is a constant of the motion. Because $r^2\dot{\phi}$ is equal to the z (and only non-vanishing) component of the vector $\boldsymbol{r} \times \boldsymbol{v}$, we have rediscovered the statement of angular-momentum conservation. Taking Eq. (3.5) into account, we have that

$$r^2\dot{\phi} = h. \tag{3.10}$$

In this form we can see that conservation of angular momentum gives rise to Kepler's second law: $r(r\,d\phi)$ is twice the area swept by the orbit as it advances by an angle $d\phi$, and $r^2\dot{\phi}$ is twice the area swept per unit time; conservation of angular momentum implies equal areas for equal times.

The radial component of the equation of motion yields $\ddot{r} - r\dot{\phi}^2 = -Gm/r^2$, or

$$\ddot{r} - \frac{h^2}{r^3} = -\frac{Gm}{r^2} \tag{3.11}$$

after involving Eq. (3.10). This second-order differential equation for $r(t)$ can be integrated once by applying the $\dot{r}$-trick: multiply the equation by $\dot{r}$ and recognize that each term is a total derivative with respect to time. Integration produces

$$\frac{1}{2}\dot{r}^2 + \frac{h^2}{2r^2} - \frac{Gm}{r} = \varepsilon, \tag{3.12}$$

in which ε is another constant of the motion. From Eq. (3.3) it is easy to see that $E = \mu\varepsilon$, and we have rediscovered the statement of energy conservation.
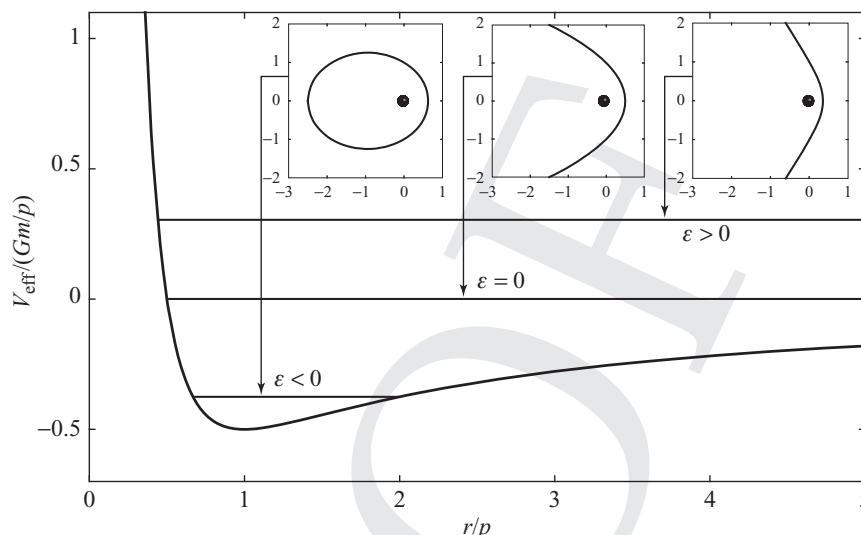
Effective potential for Kepler's problem, together with lines of constant ε. The radial variable r is presented in units of $p := h^2/(Gm)$, and V_{eff} is presented in units of Gm/p. The regions of allowed motion correspond to $\varepsilon \geq V_{\text{eff}}(r)$, and turning points occur when $\varepsilon = V_{\text{eff}}(r)$. The different types of Keplerian motion (hyperbolic, parabolic, and elliptical) are shown.

It is instructive to rewrite Eq. (3.12) in the form

$$\frac{1}{2}\dot{r}^2 = \varepsilon - V_{\text{eff}}(r), \tag{3.13}$$

in which the effective radial potential is defined as

$$V_{\text{eff}}(r) := \frac{h^2}{2r^2} - \frac{Gm}{r}. \tag{3.14}$$

This new form allows us to explore the qualitative features of Keplerian motion without having to perform additional calculations. In Fig. 3.1 we display a plot of $V_{\text{eff}}(r)$ for $h \neq 0$, with zero on the vertical axis denoting the limiting value of the effective potential as $r \to \infty$. The potential consists of an attractive (negative) gravitational well, and a repulsive (positive) centrifugal barrier rising to infinity as $r \to 0$. Motion is allowed when $\dot{r}^2 \geq 0$, that is, when $\varepsilon \geq V_{\text{eff}}(r)$, and regions where this condition is met are easily identified in the figure. A *turning point* occurs when $\dot{r} = 0$ and $\varepsilon = V_{\text{eff}}(r)$; at such points the radial velocity changes sign, and the motion changes from incoming to outgoing, or from outgoing to incoming.

We can easily imagine how a particle would move in this effective potential. If the particle has an energy $\varepsilon > 0$, then there is a single turning point at some innermost radius $r_{\min}$, and the motion takes place for $r \geq r_{\min}$. The particle starts at infinity with a negative radial velocity $\dot{r} = -\sqrt{2\varepsilon}$ and a vanishing angular velocity $\dot{\phi}$. As r decreases, $\dot{r}$ becomes increasingly negative, and $\dot{\phi}$ increases to obey conservation of angular momentum, until the particle reaches the turning point at $r = r_{\min}$. At this point $\dot{r}$ turns positive, and the particle begins its way back to infinity. As we shall see below, the particle follows a hyperbola in the

orbital plane, and motion with $\varepsilon > 0$ is known as *hyperbolic motion*. Such an orbit is not bound to the gravitating center, as revealed by the fact that the total energy is dominated by (positive) kinetic energy instead of (negative) gravitational potential energy. The limiting case of an unbound orbit corresponds to $\varepsilon = 0$. Here the particle begins from rest at infinity, proceeds to a single turning point at $r = r_{\min}$, and returns to a state of rest at infinity. In this case the path is a parabola in the orbital plane, and motion with $\varepsilon = 0$ is known as *parabolic motion*.

When $\varepsilon < 0$, that is, when gravitational potential energy dominates over kinetic energy, the diagram reveals that there are now two turning points at $r = r_{\min}$ and $r = r_{\max}$. In this case the orbital motion is bound to the gravitating center, and takes place between the innermost and outermost radii. We shall see that the bound orbit describes an ellipse, and motion with $\varepsilon < 0$ is known as *elliptical motion*. A special case of elliptical motion occurs when ε is made equal to the minimum value of the effective potential. In this case the turning points merge to a single radius r_0, and motion proceeds at this fixed radius; this is a *circular orbit* with $h^2 = Gmr_0$, $\dot{\phi} = \sqrt{Gm/r_0^3}$, and $\varepsilon = -Gm/(2r_0)$.

The preceding discussion was couched in terms of a particle moving in the effective potential $V_{\text{eff}}(r)$. It is important to understand, however, that this "particle" at position $\boldsymbol{r}(t)$ is in fact a fictitious representation of the relative orbit, which, as we saw back in Eq. (3.2), is subjected to a fictitious gravitating center of mass m at $\boldsymbol{r} = \boldsymbol{0}$. The motion described previously is therefore a description of the relative orbital motion. But thanks to Eqs. (3.1), the motion of each body is merely a scaled version of the relative motion, and it can be described in the same language; the bodies move about each other on opposite sides of the barycenter. In the limit of small mass ratios, $m_1/m_2 \ll 1$, it becomes increasingly true that $\boldsymbol{r}_1 \to \boldsymbol{r}$ and $\boldsymbol{r}_2 \to \boldsymbol{0}$; in this limiting case m_1 becomes a test mass in the field of $m_2 \to m$, its orbit coincides with the relative orbit, and m_2 stays put at the barycentric position.

3.2.4 Solution to Kepler's problem

Formal solution; integrable systems

Formally, a solution to Kepler's problem can be obtained by integrating Eq. (3.13) to get

$$t - t_i = \pm \int_{r_i}^{r} \frac{dr'}{\sqrt{2[\varepsilon - V_{\text{eff}}(r')]}}, \tag{3.15}$$

inverting the result for $r(t)$, and integrating Eq. (3.10) to get

$$\phi - \phi_i = h \int_{t_i}^{t} \frac{dt'}{r(t')^2}, \tag{3.16}$$

which gives $\phi(t)$. In these equations, t_i is the time at which $r = r_i$, and ϕ_i is the orbital angle at time t_i. By following this procedure we convert the task of solving the second-order differential equations of (3.2) to the task of performing two integrations, or, to use an older but still popular terminology, evaluating two quadratures. When a dynamical problem such as this one can be reduced to doing quadratures, the problem is said to be *completely integrable*. A full discussion of integrable systems is beyond the scope of this book, but

roughly speaking, a dynamical system is completely integrable when it possesses a sufficient number of conserved quantities. In the case of Kepler's problem, it is the constancy of total momentum (which allowed us to eliminate the motion of the barycenter), of total angular momentum (which allowed us to restrict the motion to a plane and to reduce the angular equation to $\dot\phi = h/r^2$), and of energy (which resulted in an equation for $\dot r^2$) that make the problem completely integrable.

Completely integrable systems are mathematically elegant, and are extremely convenient when they come along, but unfortunately they are rather rare in physics. Fortunately for us, however, the Kepler problem is one of them. More generally, because Eq. (3.15) is valid for any effective potential $V_{\text{eff}}(r)$, motion in any spherically symmetric, static potential is always integrable.

The formal solutions of Eqs. (3.15) and (3.16) are not very useful from a practical point of view. We could try to carry out the integrals explicitly, but the results would not be illuminating. Alternatively, we could evaluate the integrals numerically, but the resulting tables would be of limited utility, and they would provide very little insight. We shall therefore proceed differently.

Spatial solution; conic sections

As a first step toward integrating the equations of motion, we eliminate t from the system of equations (3.10), (3.11), and adopt the orbital angle ϕ as independent variable. This strategy allows us to unravel the spatial aspects of the orbit – its shape in the orbital plane – and we shall return to the problem of describing the motion in time at a later stage. We also adopt $u := 1/r$ as a convenient substitute for r, and derive a differential equation for it by applying the chain rule of differential calculus. For example, we have that $\dot r = -u^{-2}\dot u = -u^{-2}\dot\phi u' = -hu'$ and $\ddot r = -h^2 u^2 u''$, in which a prime indicates differentiation with respect to ϕ. Making the substitutions within Eq. (3.11), we quickly arrive at

$$u'' + u = \frac{Gm}{h^2}. \tag{3.17}$$

The general solution to this simple equation is

$$u = \frac{Gm}{h^2}\big[1 + e\cos(\phi - \omega)\big], \tag{3.18}$$

in which e and ω are arbitrary constants of integration.

Returning to the original radial variable, the spatial solution to Kepler's problem is

$$r = \frac{p}{1 + e\cos(\phi - \omega)}, \tag{3.19}$$

in which

$$p := \frac{h^2}{Gm} \tag{3.20}$$

is a quantity known as the orbit's *semi-latus rectum*. We note that a solution with $e < 0$ is equivalent to one with $e > 0$, provided that ω is changed to $\omega + \pi$; we will adopt the convention that e is never negative.

The curve described by Eq. (3.19) is a *conic section*, an ellipse when $e < 1$, a hyperbola when $e > 1$, and a parabola when $e = 1$, with $r = 0$ situated at one of the curve's foci. The parameter e is called the *eccentricity* of the orbit. Note that r achieves a minimum when $\phi = \omega$. This is the point of closest approach in the orbit, called the *periapsis* or *pericenter*, and ω is known as the *longitude of pericenter*; its role is to fix the orientation of the orbit in the orbital plane. The term pericenter is usually adapted to reflect the identity of specific astronomical bodies. For example, we have perihelion for the Sun, perigee for the Earth, perijove for Jupiter, and periastron for binary star systems. No consensus has emerged to date for the closest approach to a black hole, but the word peribothron is gaining popularity; it was crafted by our colleague Sterl Phinney from the Greek root "bothros," which means hole or pit.

We examine first the elliptical orbits with $e < 1$. In this case the motion described by Eq. (3.19) is periodic, with period $\Delta\phi = 2\pi$, and we have recovered Kepler's first law, that planets move on elliptical paths around the Sun. We have already seen that $\phi = \omega$ describes the periapsis or pericenter, and we now see that $\phi = \omega + \pi$ describes the point of greatest separation, called *apoapsis* or *apocenter*. The pericenter and apocenter distances from the focus are given by

$$r_{\text{peri}} = \frac{p}{1+e}, \quad r_{\text{apo}} = \frac{p}{1-e}. \tag{3.21}$$

The sum of these is the major axis of the ellipse, and we define the *semi-major axis a* to be

$$a := \frac{1}{2}(r_{\text{peri}} + r_{\text{apo}}) = \frac{p}{1-e^2}. \tag{3.22}$$

The semi-latus rectum can also be expressed in terms of these quantities:

$$p = \frac{2r_{\text{peri}}r_{\text{apo}}}{r_{\text{peri}} + r_{\text{apo}}}. \tag{3.23}$$

A special case of elliptical motion occurs when $e = 0$. In this case $r = p$ is constant, and the orbit is circular.

We examine next the hyperbolic and parabolic orbits with $e \geq 1$. In these cases the motion is no longer periodic, and the notion of apocenter ceases to be meaningful. Similarly, while we can still define a quantity a related to $p := h^2/(Gm)$ by $a = p/(1 - e^2)$, this quantity also loses its usefulness, being negative for hyperbolic orbits and formally infinite for parabolic orbits. The pericenter still occurs at $\phi = \omega$, with $r_{\text{peri}} = p/(1 + e)$, and r approaches infinity as $\cos(\phi - \omega) \to -e^{-1}$. It is easy to show that the net change between the asymptotic angle ϕ_∞^{in} of the incoming part of the orbit and the asymptotic angle ϕ_∞^{out} of the outgoing orbit is given by

$$\Delta\phi := \phi_\infty^{\text{out}} - \phi_\infty^{\text{in}} = 2\arccos(-e^{-1}). \tag{3.24}$$

This reduces to $\Delta\phi = 2\pi$ when $e = 1$.

Returning to the general case with any e, we invoke Eqs. (3.10), (3.19), and (3.20) to derive the following useful expressions for the radial and angular velocities of a Keplerian

orbit:

$$\dot{r} = \sqrt{\frac{Gm}{p}} \, e \sin(\phi - \omega), \tag{3.25a}$$

$$\dot{\phi} = \sqrt{\frac{Gm}{p^3}} \big[1 + e \cos(\phi - \omega)\big]^2. \tag{3.25b}$$

From these we obtain

$$v^2 = \frac{Gm}{p}\big[1 + 2e\cos(\phi - \omega) + e^2\big] = Gm\left(\frac{2}{r} - \frac{1 - e^2}{p}\right) \tag{3.26}$$

for the squared orbital velocity. From Eq. (3.12) we get an expression for the total energy $E := \mu\varepsilon$,

$$E = -G\mu m \frac{1 - e^2}{2p} = -\frac{G\mu m}{2a}. \tag{3.27}$$

We recall that $\mu := m_1 m_2/m$ is the system's reduced mass, and that a is related to p by Eq. (3.22). And finally, from Eq. (3.5) we have that

$$\boldsymbol{L} = \mu\sqrt{Gmp}\,\boldsymbol{e}_z \tag{3.28}$$

is the system's total angular momentum.

Motion in time; eccentric anomaly

So far we have determined the orbit as a function of ϕ, and the description involves three arbitrary constants, p, e, and ω, known as *orbital elements*. As we have seen, the semi-latus rectum p is a substitute for angular momentum, while the associated semi-major axis $a := p/(1 - e^2)$ is a substitute for orbital energy. The true nature of the longitude of pericenter ω will be revealed below.

The description of the orbit is completed by giving ϕ as a function of time. This can be accomplished by integrating Eq. (3.25b), in the form of

$$t - T = \sqrt{\frac{p^3}{Gm}} \int_\omega^\phi \frac{d\phi'}{\big[1 + e\cos(\phi' - \omega)\big]^2}, \tag{3.29}$$

where the constant T, called the *time of pericenter passage*, is a fourth orbital element. This equation can be inverted to give $\phi(t)$, and Kepler's problem is now solved.

There remains, however, the practical problem of evaluating the integral of Eq. (3.29). We henceforth specialize to elliptical motion, and introduce another running parameter on the orbit, known as the *eccentric anomaly u*. This is defined by the relations

$$\cos f = \frac{\cos u - e}{1 - e\cos u}, \qquad \sin f = \frac{\sqrt{1 - e^2}\,\sin u}{1 - e\cos u}, \tag{3.30}$$

in which $f := \phi - \omega$ is the *true anomaly*, the orbital angle as measured from the pericenter. The inverted relations are

$$\cos u = \frac{\cos f + e}{1 + e \cos f}, \qquad \sin u = \frac{\sqrt{1 - e^2}\,\sin f}{1 + e \cos f}, \qquad (3.31)$$

and these relationships are neatly summarized by the half-angle formula

$$\tan \frac{1}{2} f = \sqrt{\frac{1 + e}{1 - e}}\, \tan \frac{1}{2} u. \qquad (3.32)$$

These equations reveal that $u = 0$ at pericenter (where $\phi = \omega$ and $f = 0$), that $u = \pi$ at apocenter (where $f = \pi$), and that $u = 2\pi$ at the end of a complete orbital cycle. Other useful relations are

$$\frac{df}{du} = \frac{\sqrt{1 - e^2}}{1 - e \cos u}, \qquad \frac{du}{df} = \frac{\sqrt{1 - e^2}}{1 + e \cos f}. \qquad (3.33)$$

The eccentric anomaly gives rise to a useful alternative description of the orbit. Making the substitutions reveals that

$$r = a(1 - e \cos u), \qquad (3.34a)$$

$$\dot{r} = \sqrt{\frac{Gm}{a}} \frac{e \sin u}{1 - e \cos u}, \qquad (3.34b)$$

$$\dot{u} = \sqrt{\frac{Gm}{a^3}} \frac{1}{1 - e \cos u}, \qquad (3.34c)$$

$$v^2 = \frac{Gm}{a} \frac{1 + e \cos u}{1 - e \cos u} = Gm\left(\frac{2}{r} - \frac{1}{a}\right). \qquad (3.34d)$$

The main advantage of this description resides in Eq. (3.34c), which can be immediately integrated to give

$$t - T = \sqrt{\frac{a^3}{Gm}}\,(u - e \sin u). \qquad (3.35)$$

This is known as *Kepler's equation*, which gives t as a simple function of the eccentric anomaly u. The equation can be inverted to yield $u(t)$, see Box 3.1, and this can finally be inserted within Eqs. (3.30) or (3.32) to express f as a function of time.

A description in terms of the eccentric anomaly is often a judicious choice when it is required to integrate over time. As a simple example, we can calculate the orbital period P of an elliptical orbit using Eq. (3.35). When u increases by 2π, t increases by P, and with very little effort we find that

$$P = 2\pi \sqrt{\frac{a^3}{Gm}}. \qquad (3.36)$$

While the period can also be obtained directly from Eq. (3.29), the integration is far more laborious. That P is proportional to $a^{3/2}$ is Kepler's third law of planetary motion.

Box 3.1	Solving Kepler's equation

We wish to find the eccentric anomaly u that corresponds to a given time t. For this purpose it is useful to rewrite Kepler's equation as

$$u - e \sin u = M := \sqrt{\frac{Gm}{a^3}}(t - T) = 2\pi \frac{t - T}{P},$$

and to map the *mean anomaly* M to the interval $0 \leq M < 2\pi$ by subtracting an appropriate multiple of P from $t - T$.

A time-honored method to solve Kepler's equation is based on Newton's root-finding method. It is an implementation of the iterative scheme

$$u_{n+1} = u_n + \frac{M - (u_n - e \sin u_n)}{1 - e \cos u_n}$$

until u converges to the desired accuracy. The iterations are seeded with $u_0 = M$, and the scheme typically requires a small number of iterations (fewer than six or so) for machine-precision accuracy.

Constancy of the pericenter; Runge–Lenz vector

We have seen that Keplerian motion within a fixed orbital plane is characterized by four constants of the motion, the orbital elements p, e, ω, and T. We have seen that constancy of p is tied to conservation of angular momentum, and that constancy of $a := p/(1 - e^2)$ is tied to conservation of energy (so that constancy of e is also assured). In addition, the appearance of T as an integration constant was expected from the fact that the gravitational potential Gm/r does not depend explicitly on time. Constancy of ω, however, is not related to the spherical symmetry of the potential nor its time independence; we must seek a deeper cause. It is worth emphasizing that constancy of ω is a very important property of a Keplerian orbit: It ensures that the orientation of the orbit stays fixed, that the position of the pericenter does not move, and when the orbit is bound, that the orbit retraces itself after each orbital cycle.

The constancy of ω is the result of a hidden symmetry of Kepler's problem, associated with the specific $1/r$ nature of the gravitational potential; the symmetry does not exist for other potentials. The symmetry gives rise to a conservation statement for the *Runge–Lenz vector*, defined by

$$A := \frac{v \times h}{Gm} - n, \tag{3.37}$$

in which $h := r \times v$ and $n := r/r$. A short computation using $a = -Gmn/r^2$ and $\dot{r} = n \cdot v$ shows that

$$\frac{dA}{dt} = 0, \tag{3.38}$$

and the manipulations do indeed reveal that constancy of A relies on the specific form of the gravitational acceleration. The Runge–Lenz vector can be evaluated explicitly by making use of the Keplerian results displayed previously, including Eqs. (3.5), (3.7), (3.9),

and (3.25). The result is

$$A = e(\cos \omega \, e_x + \sin \omega \, e_y), \tag{3.39}$$

and it reveals that the vector points in the fixed direction of the pericenter. The vector has a length e, and constancy of A as a vector implies that both e and ω are constants of the motion.

3.2.5 Keplerian orbits in space

The description of Keplerian orbits given previously achieved a remarkable degree of simplicity, thanks in part to conservation of angular momentum, which guarantees that the motion takes place in a fixed orbital plane. The description was entirely "orbit-centric," in that the reference frame was selected specifically to give a simple description of the orbital plane; and the simplicity was achieved largely by adopting the polar coordinates (r, ϕ) attached to the orbital plane. In many applications this description is entirely adequate, but in many others it is necessary to provide a fuller description that is less orbit-centric. For example, one might wish to describe the motion of two or more planets around the Sun (under the assumption that the Sun's motion and inter-planet interactions are negligible), with each planet moving in a different orbital plane; in such a case one would like to adopt the same reference frame for all the planets. Another example, which will be explored in some detail below, involves a two-body system perturbed by external bodies, or by a deformation of each body from a spherical configuration; these perturbations cause the orbital plane to move, and these motions must be described relative to a fixed reference frame. For such applications we require a fuller description of the orbital motion in space, relative to a reference frame that is not attached to the orbital plane.

We were already given an *orbital frame* with coordinates (x, y, z), such that the fixed orbital plane coincides with the x–y plane, and such that the angular-momentum vector is aligned with the z-direction. The orbital frame comes with the constant basis vectors e_x, e_y, and e_z, as well as the time-dependent basis n, λ, and e_z. We now introduce a *fundamental frame* with coordinates (X, Y, Z), and seek to describe the orbital motion in this new frame. We adopt the X–Y plane as a reference plane in the new frame, and the Z-axis as a reference direction. The fundamental frame comes with a constant vectorial basis e_X, e_Y, and e_Z. We assume that the orbital and fundamental frames share the same origin, so that $X = Y = Z = 0$ when $x = y = z = 0$. The choice of fundamental frame is arbitrary, and is often dictated by convention or convenience. For example, in the description of planetary motion in the solar system, the reference plane is chosen to coincide with Earth's own orbital plane (called the *ecliptic*). As another example, in the description of satellites orbiting the Earth, the fundamental plane is chosen to coincide with Earth's equatorial plane. In each case the direction of the X-axis is selected by convention.

The description of the orbital motion relative to the fundamental frame requires the introduction of additional orbital elements. The situation is represented in Fig. 3.2, which shows the orbital plane crossing the fundamental plane at an angle ι called the *inclination*; this is the angle between the z-direction of the orbital frame and the Z-direction of the fundamental frame. The line of intersection between the two planes is known as the *line of nodes*, and the point at which the orbit cuts the fundamental plane from below is the
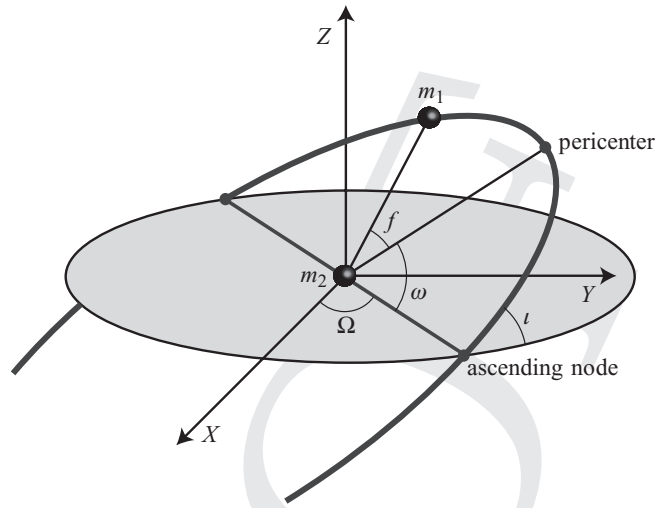
Fig. 3.2   Orbital motion viewed in the fundamental reference frame.

ascending node; the *descending node* is the point at which the orbit cuts the plane from above. The angle Ω between the X-direction and the line of nodes is the *longitude of the ascending node*. The diagram also shows ω, the *longitude of pericenter* introduced previously, which is now defined specifically as the angle between the line of nodes and the direction to the pericenter, as measured in the orbital plane. In this new description of the orbital motion, it is conventional to align the x-axis of the orbital plane with the direction to the pericenter, thereby deviating from our previous practice. And finally, the diagram shows the *true anomaly* f, the angle between the separation vector $\boldsymbol{r}$ and the direction to the pericenter, as measured in the orbital plane. The complete listing of orbital elements therefore consists of the principal elements p and e, the positional elements ι, Ω, and ω, and the time element T; the total number of elements is six, and it is no accident that this number corresponds to the number of initial conditions required to select a unique solution to Kepler's problem.

With these definitions and conventions, we show in Box 3.2 that the components of the separation vector $\boldsymbol{r}$ in the fundamental (X, Y, Z) frame are given by

$$r^X = r\big[\cos\Omega\cos(\omega + f) - \cos\iota\sin\Omega\sin(\omega + f)\big], \tag{3.40a}$$

$$r^Y = r\big[\sin\Omega\cos(\omega + f) + \cos\iota\cos\Omega\sin(\omega + f)\big], \tag{3.40b}$$

$$r^Z = r\sin\iota\sin(\omega + f), \tag{3.40c}$$

in which $r = p/(1 + e\cos f)$. Differentiation with respect to time and involvement of Eqs. (3.25) give us the components of the velocity vector,

$$v^X = -\sqrt{\frac{Gm}{p}}\Big\{\cos\Omega\big[\sin(\omega + f) + e\sin\omega\big] + \cos\iota\sin\Omega\big[\cos(\omega + f) + e\cos\omega\big]\Big\},$$

$$\tag{3.41a}$$

$$v^Y = -\sqrt{\frac{Gm}{p}} \left\{ \sin \Omega \left[\sin(\omega + f) + e \sin \omega \right] - \cos \iota \cos \Omega \left[\cos(\omega + f) + e \cos \omega \right] \right\},$$

$$(3.41b)$$

$$\cancel{s}v^Z = \sqrt{\frac{Gm}{p}} \sin \iota \left[\cos(\omega + f) + e \cos \omega \right].$$
$$(3.41c)$$

The results of Box 3.2 also allow us to express the orbital basis vectors $\boldsymbol{n}$ and $\boldsymbol{\lambda}$ as

$$\boldsymbol{n} = \left[\cos \Omega \cos(\omega + f) - \cos \iota \sin \Omega \sin(\omega + f) \right] \boldsymbol{e}_X$$
$$+ \left[\sin \Omega \cos(\omega + f) + \cos \iota \cos \Omega \sin(\omega + f) \right] \boldsymbol{e}_Y$$
$$+ \sin \iota \sin(\omega + f) \, \boldsymbol{e}_Z \tag{3.42}$$

and

$$\boldsymbol{\lambda} = \left[-\cos \Omega \sin(\omega + f) - \cos \iota \sin \Omega \cos(\omega + f) \right] \boldsymbol{e}_X$$
$$+ \left[-\sin \Omega \sin(\omega + f) + \cos \iota \cos \Omega \cos(\omega + f) \right] \boldsymbol{e}_Y$$
$$+ \sin \iota \cos(\omega + f) \, \boldsymbol{e}_Z. \tag{3.43}$$

In addition, the direction to the pericenter is given by

$$\boldsymbol{e}_x = \left(\cos \Omega \cos \omega - \cos \iota \sin \Omega \sin \omega \right) \boldsymbol{e}_X$$
$$+ \left(\sin \Omega \cos \omega + \cos \iota \cos \Omega \sin \omega \right) \boldsymbol{e}_Y$$
$$+ \sin \iota \sin \omega \, \boldsymbol{e}_Z, \tag{3.44}$$

while the direction of the angular-momentum vector is

$$\boldsymbol{e}_z = \sin \iota \sin \Omega \, \boldsymbol{e}_X - \sin \iota \cos \Omega \, \boldsymbol{e}_Y + \cos \iota \, \boldsymbol{e}_Z. \tag{3.45}$$

In terms of these we have that $\boldsymbol{A} = e \boldsymbol{e}_x$ and $\boldsymbol{h} = h \boldsymbol{e}_z$, with $h := \sqrt{Gmp}$.

The description of the orbital motion in the fundamental frame allows us to provide simple definitions for the orbital elements. We have that

$$p := \frac{h^2}{Gm}, \tag{3.46a}$$

$$e := |\boldsymbol{A}|, \tag{3.46b}$$

$$\cos \iota := \boldsymbol{e}_z \cdot \boldsymbol{e}_Z = \frac{\boldsymbol{h} \cdot \boldsymbol{e}_Z}{h}, \tag{3.46c}$$

$$\sin \iota \sin \Omega := \boldsymbol{e}_z \cdot \boldsymbol{e}_X = \frac{\boldsymbol{h} \cdot \boldsymbol{e}_X}{h}, \tag{3.46d}$$

$$\sin \iota \sin \omega := \boldsymbol{e}_x \cdot \boldsymbol{e}_Z = \frac{\boldsymbol{A} \cdot \boldsymbol{e}_Z}{e}. \tag{3.46e}$$

To this we may add $a := p/(1 - e^2)$ and the time element T, which is defined by Eq. (3.29). These definitions are elementary and fundamental, because they are formulated directly in terms of the fundamental vectors of the problem, the reduced angular-momentum vector $\boldsymbol{h} = \boldsymbol{r} \times \boldsymbol{v}$ and the Runge–Lenz vector $\boldsymbol{A}$ defined by Eq. (3.37). They will play a large role in the next section.

| Box 3.2 | Orbital and fundamental frames |

A close examination of Fig. 3.2 and some thought reveal that the fundamental (X, Y, Z) frame can be obtained from the orbital (x, y, z) frame by a sequence of three rotations. The first is a rotation R_1 by an angle $-\omega$ around the z-axis, to align the rotated x-axis with the line of nodes. The second is a rotation R_2 by an angle $-\iota$ around the new x-axis, to align the rotated z-axis with the final Z-axis. The third is a rotation R_3 by an angle $-\Omega$ around the new z-axis, to align the rotated x-axis with the final X-axis. The rotation matrices are given by

$$\mathsf{R}_1 = \begin{pmatrix} \cos\omega & -\sin\omega & 0 \\ \sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \mathsf{R}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\iota & -\sin\iota \\ 0 & \sin\iota & \cos\iota \end{pmatrix},$$

$$\mathsf{R}_3 = \begin{pmatrix} \cos\Omega & -\sin\Omega & 0 \\ \sin\Omega & \cos\Omega & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and the overall transformation between the orbital and fundamental frames is $\mathsf{X} = \mathsf{R}_3\mathsf{R}_2\mathsf{R}_1\,\mathsf{x}$, with x representing a column vector with components (x, y, z), and X a column vector with components (X, Y, Z).

The end result of the transformation is

$$\begin{aligned} X &= \big(\cos\Omega\cos\omega - \cos\iota\sin\Omega\sin\omega\big)x \\ &\quad - \big(\cos\Omega\sin\omega + \cos\iota\sin\Omega\cos\omega\big)y \\ &\quad + \big(\sin\iota\sin\Omega\big)z, \\ Y &= \big(\sin\Omega\cos\omega + \cos\iota\cos\Omega\sin\omega\big)x \\ &\quad - \big(\sin\Omega\sin\omega - \cos\iota\cos\Omega\cos\omega\big)y \\ &\quad - \big(\sin\iota\cos\Omega\big)z, \\ Z &= \big(\sin\iota\sin\omega\big)x + \big(\sin\iota\cos\omega\big)y + \big(\cos\iota\big)z. \end{aligned}$$

The components of $\boldsymbol{r}$ displayed in Eqs. (3.40) follow from these results by inserting $x = r\cos f$, $y = r\sin f$, and $z = 0$.

From the coordinate transformation we can easily obtain the old basis vectors in terms of the new:

$$\begin{aligned} \boldsymbol{e}_x = \frac{\partial \boldsymbol{x}}{\partial x} &= \big(\cos\Omega\cos\omega - \cos\iota\sin\Omega\sin\omega\big)\boldsymbol{e}_X \\ &\quad + \big(\sin\Omega\cos\omega + \cos\iota\cos\Omega\sin\omega\big)\boldsymbol{e}_Y \\ &\quad + \big(\sin\iota\sin\omega\big)\boldsymbol{e}_Z, \\ \boldsymbol{e}_y = \frac{\partial \boldsymbol{x}}{\partial y} &= -\big(\cos\Omega\sin\omega + \cos\iota\sin\Omega\cos\omega\big)\boldsymbol{e}_X \\ &\quad - \big(\sin\Omega\sin\omega - \cos\iota\cos\Omega\cos\omega\big)\boldsymbol{e}_Y \\ &\quad + \big(\sin\iota\cos\omega\big)\boldsymbol{e}_Z, \\ \boldsymbol{e}_z = \frac{\partial \boldsymbol{x}}{\partial z} &= \big(\sin\iota\sin\Omega\big)\boldsymbol{e}_X - \big(\sin\iota\cos\Omega\big)\boldsymbol{e}_Y + \big(\cos\iota\big)\boldsymbol{e}_Z, \end{aligned}$$

in which the position vector $\boldsymbol{x}$ is expressed as $\boldsymbol{x} = X\boldsymbol{e}_X + Y\boldsymbol{e}_Y + Z\boldsymbol{e}_Z$.

> The orbital vectors of Eqs. (3.42) and (3.43) are obtained by inserting these results within the relations $n = \cos f \, e_x + \sin f \, e_y$ and $\lambda = -\sin f \, e_x + \cos f \, e_y$, which are inherited from Eqs. (3.7); the new relations reflect the change of convention regarding the choice of x-direction.

3.3 Perturbed Kepler problem

We saw in the preceding section that the motion of two spherical bodies under their mutual gravitational attraction can be solved exactly and completely. As we shall see in Sec. 3.5, the same cannot be said of the three-body problem, and in general, the N-body problem does not admit an exact solution. Many situations of interest, however, involve more than two bodies. The Sun–Earth–Moon system is an extremely pertinent example, and the motion of any planet around the Sun is affected by the massive presence of Jupiter. While these systems cannot be given an exact description, we can nevertheless make progress by appealing to the fact that in many applications, the additional bodies have only a small effect on the orbital motion of a two-body system. In the Sun–Earth–Moon system, the dominant interaction is between the Sun and the Earth, and the gravitational effects of the Moon on Earth's orbital motion are small. Similarly, while Jupiter is indeed a massive body, it is sufficiently far from the other planets that its gravity has a small effect on them. The approximate analysis of small external influences on a system dominated by an internal interaction is the realm of *perturbation theory*, and in this section we formulate a perturbation theory for Kepler's problem.

3.3.1 Perturbing force

We return to the two-body problem of Sec. 3.2, but now suppose that the relative acceleration $a := a_1 - a_2$ between bodies is given by

$$a = -\frac{Gm}{r^2} n + f, \tag{3.47}$$

in which $m := m_1 + m_2$, $r := r_1 - r_2$, $r := |r|$, $n := r/r$, and where f is a perturbing force per unit mass, which may depend on r, $v := v_1 - v_2$, and time. The presence of the perturbing force implies that the Keplerian motion reviewed in Sec. 3.2 is no a longer a solution to the equations of motion. As a consequence we can no longer expect $h := r \times v$ to be a constant vector, the orbit to be elliptical, the pericenter to stay in a fixed position, and so on. But with f assumed to be small, we would expect the Keplerian description to remain approximately true. Our main goal is to find a useful way of describing the small deviations from Keplerian motion produced by the perturbing force. Below we shall follow the convention from celestial mechanics, and refer to f as a perturbing force, in spite of the fact that it is really a force per unit mass.

It is helpful to introduce, as we did previously, a vectorial basis adapted to the orbital motion of the two bodies. The first member of the basis will be n, which is already singled out in Eq. (3.47). The third member will be $e_z := h/h$, which is necessarily orthogonal to n; this vector can no longer be assumed to be a constant vector. The second member of the basis is then λ, which is defined to be orthogonal to both n and e_z. We take the triad of vectors (n, λ, e_z) to form a right-handed basis. When $f = 0$ the vector basis reduces to the one introduced back in Sec. 3.2, but the bases are quite distinct in the presence of a perturbing force; for reasons that will become clear later, we perversely adopt the same notation in spite of this important distinction.

The vectorial basis can be used to decompose any vector relevant to the description of the perturbed orbital motion. Examples are

$$r = r\,n, \tag{3.48a}$$

$$v = \dot{r}\,n + v_\perp\,\lambda, \tag{3.48b}$$

$$h = h\,e_z, \tag{3.48c}$$

in which $\dot{r} := v \cdot n$ and $v_\perp := v \cdot \lambda$; it is easy to verify that $h = rv_\perp$. Another example is the perturbing force, which is decomposed as

$$f = \mathcal{R}\,n + \mathcal{S}\,\lambda + \mathcal{W}\,e_z, \tag{3.49}$$

in terms of components $\mathcal{R}$, $\mathcal{S}$, and $\mathcal{W}$. Yet another example is the Runge–Lenz vector

$$A := \frac{v \times h}{Gm} - n = A\,e_x, \qquad e_x = \cos\alpha\,n - \sin\alpha\,\lambda, \tag{3.50}$$

which is given a component $A\cos\alpha$ along n, and a component $-A\sin\alpha$ along λ; there is no component along e_z because A is orthogonal to h. Here we have defined the Runge–Lenz vector exactly as in Sec. 3.2, but we can no longer expect A to be a constant vector. We have given the vector a length A, and Eq. (3.50) serves as a definition for the unit vector e_x, which is also not a constant vector. In an unperturbed situation A would be equal to e, α would be equal to the true anomaly f, and e_x would indeed be a constant vector; but none of these statements can be expected to hold when $f \neq 0$.

The effect of the perturbing force f on the vectors h and A can be calculated by appealing directly to their definitions. We find that

$$\frac{dh}{dt} = r \times f = -r\mathcal{W}\,\lambda + r\mathcal{S}\,e_z \tag{3.51}$$

and

$$Gm\frac{dA}{dt} = f \times h + v \times (r \times f) = 2h\mathcal{S}\,n - (h\mathcal{R} + r\dot{r}\mathcal{S})\,\lambda - r\dot{r}\mathcal{W}\,e_z. \tag{3.52}$$

These equations imply that

$$\frac{dh}{dt} = r\mathcal{S} \tag{3.53}$$

and

$$Gm\frac{dA}{dt} = h\sin\alpha\,\mathcal{R} + (2h\cos\alpha + r\dot{r}\sin\alpha)\mathcal{S}, \tag{3.54}$$

as well as

$$h\frac{d\boldsymbol{e}_z}{dt} = -r\mathcal{W}\boldsymbol{\lambda} \tag{3.55}$$

and

$$GmA\frac{d\boldsymbol{e}_x}{dt} = -\left[h\cos\alpha\,\mathcal{R} + (-2h\sin\alpha + r\dot{r}\cos\alpha)\mathcal{S}\right](\sin\alpha\,\boldsymbol{n} + \cos\alpha\,\boldsymbol{\lambda}) - r\dot{r}\mathcal{W}\,\boldsymbol{e}_z. \tag{3.56}$$

These equations tell us, for example, that $\mathcal{S}$ produces a change in the magnitude of the angular-momentum vector, while $\mathcal{W}$ produces a change in its direction. Similarly, both $\mathcal{R}$ and $\mathcal{S}$ produce a change in A's magnitude, as well as a change of direction orthogonal to $\boldsymbol{e}_z$.

3.3.2 Osculating orbits

How are we to solve Eq. (3.47)? A direct approach, of course, is always possible: one inserts the given expression for the external force and integrates the second-order differential equations directly, either by analytical methods when the problem is sufficiently simple (a rare occasion), or by numerical methods. The approach provides an answer, but very little insight into the effects of the perturbation on various key aspects of the orbit. We shall favor instead a clever reformulation of the perturbed problem that was initially devised by Euler and Lagrange. This reformulation, known as the *method of osculating orbital elements*, is an application to orbital dynamics of the well-known method of variation of arbitrary constants to integrate differential equations (reviewed in Box 3.3).

Box 3.3	Variation of arbitrary constants

The method of variation of arbitrary constants to integrate differential equations is best introduced by examining a simple example. Consider a simple harmonic oscillator of unit frequency driven by an external force $f(t)$. The system is described by the differential equation

$$\ddot{x} + x = f, \tag{1}$$

in which x is the displacement from equilibrium, and an overdot indicates differentiation with respect to t. To integrate this equation we begin with the general solution to the homogeneous problem, $\ddot{x} + x = 0$. This solution can be expressed as

$$x(t) = x_0\cos t + v_0\sin t, \tag{2}$$

in which $x_0 := x(t = 0)$ and $v_0 := \dot{x}(t = 0)$ are the arbitrary constants of the solution. This solution also gives us

$$\dot{x}(t) = v_0\cos t - x_0\sin t. \tag{3}$$

Returning to the original differential equation with a driving force $f(t)$, we declare that Eqs. (2) and (3) shall also be a solution to Eq. (1), and evade the obvious contradiction by allowing x_0 and v_0 to become functions of

time. Equation (3) will be compatible with Eq. (2) provided that $\dot{x}_0 \cos t + \dot{v}_0 \sin t = 0$, and substitution of Eqs. (2) and (3) within Eq. (1) yields $-\dot{x}_0 \sin t + \dot{v}_0 \cos t = f$. Solving for $\dot{x}_0$ and $\dot{v}_0$, we find that

$$\dot{x}_0 = -f \sin t, \qquad \dot{v}_0 = f \cos t, \tag{4}$$

and these equations now replace Eq. (1). The solutions are to be inserted within Eqs. (2) and (3) for a complete solution of the original problem.

These steps define the method of variation of arbitrary constants. One begins with the general solution to a homogeneous differential equation, with the appropriate number of arbitrary constants. One next solves the inhomogeneous problem by formally adopting the functional form of the homogeneous solution, and promoting the arbitrary constants to new variables. After some manipulations, the original problem acquires a new formulation as a set of first-order differential equations for the new variables.

The method

The starting point of the method of osculating orbital elements is to recall that the general solution to the unperturbed Kepler problem can be expressed as

$$\boldsymbol{r}(t) = \boldsymbol{r}_{\text{Kepler}}(t, \mu^a), \qquad \boldsymbol{v}(t) = \boldsymbol{v}_{\text{Kepler}}(t, \mu^a), \tag{3.57}$$

where $\boldsymbol{r}_{\text{Kepler}}$ and $\boldsymbol{v}_{\text{Kepler}}$ are the functional forms displayed in Eqs. (3.40) and (3.41), with μ^a denoting a collection of six constants of the motion, which we take to be the orbital elements $(p, e, \iota, \Omega, \omega, T)$. These vectors satisfy the differential equations

$$\frac{d\boldsymbol{r}_{\text{Kepler}}}{dt} = \boldsymbol{v}_{\text{Kepler}}, \qquad \frac{d\boldsymbol{v}_{\text{Kepler}}}{dt} = -Gm \frac{\boldsymbol{r}_{\text{Kepler}}}{r_{\text{Kepler}}^3}, \tag{3.58}$$

which define the unperturbed Kepler problem.

The vectors $\boldsymbol{r}_{\text{Kepler}}$ and $\boldsymbol{v}_{\text{Kepler}}$ can also provide a solution to the perturbed problem if we let the orbital elements μ^a become functions of time: $\mu^a \to \mu^a(t)$. The original mathematical problem expressed by Eq. (3.47) can then be formulated as a set of first-order differential equations for μ^a, and the solutions to these are to be inserted within Eqs. (3.57) to obtain a complete solution to the perturbed problem. The meaning of the method is that at any time t_1, the orbit is taken to be a Keplerian orbit with orbital elements $\mu^a(t_1)$ that refer to this time only; at another time t_2 the orbit will still be Keplerian, but the orbital elements will have evolved to new values $\mu^a(t_2)$. Another way of describing this is to say that at any time t_1, there exists a Keplerian orbit with elements $\mu^a(t_1)$ that is *tangent* to the perturbed orbit at that time; this is the *osculating orbit*, the term originating in the Latin word *osculatio*, which means "to kiss."

We shall therefore express the solution to Eq. (3.47) as

$$\boldsymbol{r}(t) = \boldsymbol{r}_{\text{Kepler}}\big(t, \mu^a(t)\big), \qquad \boldsymbol{v}(t) = \boldsymbol{v}_{\text{Kepler}}\big(t, \mu^a(t)\big), \tag{3.59}$$

with $\boldsymbol{r}_{\text{Kepler}}$ and $\boldsymbol{v}_{\text{Kepler}}$ still standing for the functional forms of Eqs. (3.40) and (3.41), but with the orbital elements μ^a now allowed to be functions of time. Differentiating $\boldsymbol{r}$ with

respect to time gives

$$\frac{d\mathbf{r}}{dt} = \frac{\partial \mathbf{r}_{\text{Kepler}}}{\partial t} + \sum_a \frac{\partial \mathbf{r}_{\text{Kepler}}}{\partial \mu^a} \frac{d\mu^a}{dt}. \tag{3.60}$$

The left-hand side is $\mathbf{v}$, and the first term on the right-hand side is $\mathbf{v}_{\text{Kepler}}$, because $\mathbf{r}_{\text{Kepler}}$ is differentiated with respect to t while keeping the orbital elements constant, thereby defining Keplerian motion. The equation is compatible with Eq. (3.59) provided that

$$\sum_a \frac{\partial \mathbf{r}_{\text{Kepler}}}{\partial \mu^a} \frac{d\mu^a}{dt} = \mathbf{0}, \tag{3.61}$$

and we have obtained our first osculating condition. Similarly, differentiating $\mathbf{v}$ with respect to t gives

$$\frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{v}_{\text{Kepler}}}{\partial t} + \sum_a \frac{\partial \mathbf{v}_{\text{Kepler}}}{\partial \mu^a} \frac{d\mu^a}{dt}. \tag{3.62}$$

The left-hand side is $\mathbf{a} = \mathbf{a}_{\text{Kepler}} + \mathbf{f}$, and the first term on the right-hand side is the Keplerian acceleration $\mathbf{a}_{\text{Kepler}}$. The second osculating condition is

$$\sum_a \frac{\partial \mathbf{v}_{\text{Kepler}}}{\partial \mu^a} \frac{d\mu^a}{dt} = \mathbf{f}, \tag{3.63}$$

and these equations can be solved for $d\mu^a/dt$.

Osculating equations

The purpose of the foregoing discussion was primarily to describe the conceptual aspects of the method of osculating orbital elements. While Eqs. (3.61) and (3.63) could in principle be solved for $d\mu^a/dt$ in terms of the perturbing force, it is easier in practice to obtain $d\mu^a/dt$ directly from the definitions of the orbital elements displayed in Eqs. (3.46). The orbital elements given explicitly in terms of the fixed vectorial basis ($\mathbf{e}_X$, $\mathbf{e}_Y$, $\mathbf{e}_Z$), as well as the angular-momentum vector $\mathbf{h}$ and the Runge–Lenz vector $\mathbf{A}$, which change according to Eqs. (3.51)–(3.56). The method authorizes us to substitute Keplerian relations on the right-hand side of these equations, so that h is at all times equal to $\sqrt{Gmp(t)}$, A is at all times equal to $e(t)$, α is at all times equal to the true anomaly f, r is at all times equal to $p/(1 + e\cos f)$, and $\dot{r}$ is at all times equal to $(Gm/p)^{1/2}e\sin f$.

By following this approach we quickly arrive at

$$\frac{dp}{dt} = 2\sqrt{\frac{p^3}{Gm}} \frac{1}{1 + e\cos f} \mathcal{S}, \tag{3.64a}$$

$$\frac{de}{dt} = \sqrt{\frac{p}{Gm}} \left[\sin f \, \mathcal{R} + \frac{2\cos f + e(1 + \cos^2 f)}{1 + e\cos f} \mathcal{S} \right], \tag{3.64b}$$

$$\frac{d\iota}{dt} = \sqrt{\frac{p}{Gm}} \frac{\cos(\omega + f)}{1 + e\cos f} \mathcal{W}, \tag{3.64c}$$

$$\sin\iota \frac{d\Omega}{dt} = \sqrt{\frac{p}{Gm}} \frac{\sin(\omega + f)}{1 + e\cos f} \mathcal{W}, \tag{3.64d}$$

$$\frac{d\omega}{dt} = \frac{1}{e}\sqrt{\frac{p}{Gm}}\left[-\cos f\,\mathcal{R} + \frac{2 + e\cos f}{1 + e\cos f}\sin f\,\mathcal{S} - e\cot\iota\,\frac{\sin(\omega + f)}{1 + e\cos f}\,\mathcal{W}\right],$$

(3.64e)

a listing of equations that govern the behavior of the osculating orbital elements. To these we can add

$$\frac{da}{dt} = 2\sqrt{\frac{a^3}{Gm}}(1 - e^2)^{-1/2}\left[e\sin f\,\mathcal{R} + (1 + e\cos f)\,\mathcal{S}\right],$$

(3.65)

in case $a := p/(1 - e^2)$ is preferred over p in the selection of orbital elements. Note that p and e (or a and e) are affected only by components of $\boldsymbol{f}$ in the orbital plane, while Ω and ι are affected only by the component out of the plane. All components of the perturbing force affect ω.

We are missing an equation for the variation of T, the time of pericenter passage, which determines the true anomaly f via Eq. (3.29). It is more practical, however, to close the system of equations by providing an expression for df/dt, from which the true anomaly can be obtained directly. Because f is the angle between the (varying) pericenter and the position vector $\boldsymbol{r}$, we have that $\cos f = \boldsymbol{r}\cdot\boldsymbol{e}_x/r$, and this can immediately be differentiated with respect to time. In this we insert Eq. (3.56) for $d\boldsymbol{e}_x/dt$, the usual Keplerian relations for $d\boldsymbol{r}/dt$, and we obtain

$$\frac{df}{dt} = \sqrt{\frac{Gm}{p^3}}(1 + e\cos f)^2 + \frac{1}{e}\sqrt{\frac{p}{Gm}}\left[\cos f\,\mathcal{R} - \frac{2 + e\cos f}{1 + e\cos f}\sin f\,\mathcal{S}\right]$$

(3.66)

after some simplification. The result can also be expressed as

$$\frac{df}{dt} = \left(\frac{df}{dt}\right)_{\text{Kepler}} - \left(\frac{d\omega}{dt} + \cos\iota\,\frac{d\Omega}{dt}\right),$$

(3.67)

which shows that df/dt differs from the usual Keplerian expression by a term $d\omega/dt + \cos\iota\,d\Omega/dt$ which possesses a direct geometrical meaning. We recall that ω is the angle from the (varying) pericenter to the (varying) line of nodes, as measured in the orbital plane, while Ω is the angle from the line of nodes to the (fixed) X-direction, as measured in the reference X–Y plane. The combination $d\omega + \cos\iota\,d\Omega$ can then be seen to describe the change in the direction to the pericenter relative to the X-direction, as measured entirely in the orbital plane. The non-Keplerian terms in Eqs. (3.66) and (3.67) therefore appear because the true anomaly f is measured relative to a varying set of directions.

The relevance of the combination $d\omega + \cos\iota\,d\Omega$ can also be inferred from Eq. (3.56), which gives the rate of change of the direction to the pericenter. After inserting the appropriate Keplerian relations in this expression, we find that

$$\frac{d\boldsymbol{e}_x}{dt} = \left(\frac{d\omega}{dt} + \cos\iota\,\frac{d\Omega}{dt}\right)\left(\sin f\,\boldsymbol{n} + \cos f\,\boldsymbol{\lambda}\right) + \left(\sin\omega\,\frac{d\iota}{dt} - \sin\iota\,\cos\Omega\,\frac{d\Omega}{dt}\right)\boldsymbol{e}_z,$$

(3.68)

which reveals that $d\omega + \cos\iota\,d\Omega$ does indeed describe the change of $\boldsymbol{e}_x$ within the orbital plane, while $\sin\omega\,d\iota - \sin\iota\,\cos\Omega\,d\Omega$ describes the change in the orthogonal direction.

First-order approximation

The formalism of osculating orbital elements, in the formulation of Eqs. (3.64) and (3.66), is *exactly equivalent* to the original formulation of the equations of motion in Eq. (3.47); no approximations have been introduced in the transcription. The usefulness of the formalism, however, is most immediate when the perturbing force is small, so that the changes in the orbital elements are small. In such a context one can achieve a very good approximation of the orbital dynamics by inserting the constant, zeroth-order values on the right-hand side of the equations, and integrating with respect to t to get the first-order changes. In such applications it can be convenient to use f as independent variable instead of t, and in this approximate context one can neglect the non-Keplerian terms on the right-hand side of Eq. (3.66). The system of osculating equations becomes

$$\frac{dp}{df} \simeq 2\frac{p^3}{Gm}\frac{1}{(1+e\cos f)^3}\,\mathcal{S}, \tag{3.69a}$$

$$\frac{de}{df} \simeq \frac{p^2}{Gm}\left[\frac{\sin f}{(1+e\cos f)^2}\,\mathcal{R} + \frac{2\cos f + e(1+\cos^2 f)}{(1+e\cos f)^3}\,\mathcal{S}\right], \tag{3.69b}$$

$$\frac{d\iota}{df} \simeq \frac{p^2}{Gm}\frac{\cos(\omega+f)}{(1+e\cos f)^3}\,\mathcal{W}, \tag{3.69c}$$

$$\sin\iota\frac{d\Omega}{df} \simeq \frac{p^2}{Gm}\frac{\sin(\omega+f)}{(1+e\cos f)^3}\,\mathcal{W}, \tag{3.69d}$$

$$\frac{d\omega}{df} \simeq \frac{1}{e}\frac{p^2}{Gm}\left[-\frac{\cos f}{(1+e\cos f)^2}\,\mathcal{R} + \frac{2+e\cos f}{(1+e\cos f)^3}\sin f\,\mathcal{S}\right.$$
$$\left. - e\cot\iota\frac{\sin(\omega+f)}{(1+e\cos f)^3}\,\mathcal{W}\right], \tag{3.69e}$$

with

$$\frac{dt}{df} \simeq \sqrt{\frac{p^3}{Gm}}\frac{1}{(1+e\cos f)^2}$$
$$\times\left\{1 - \frac{1}{e}\frac{p^2}{Gm}\left[\frac{\cos f}{(1+e\cos f)^2}\,\mathcal{R} - \frac{2+e\cos f}{(1+e\cos f)^3}\sin f\,\mathcal{S}\right]\right\} \tag{3.70}$$

providing the temporal information.

In most applications of the formalism it is found that the orbital elements undergo two types of changes. The first is an oscillation with a period equal to the orbital period P (or multiples of P), as given by Eq. (3.36); such changes are typically uninteresting. The second type is a steady drift that does not average out after a few orbital cycles; such changes, which are known as *secular changes*, are typically much more interesting, because they accumulate over time and eventually lead to large departures from the initial Keplerian orbit. The secular change of an orbital element μ^a over a complete orbit is given by

$$\Delta\mu^a = \int_0^P \frac{d\mu^a}{dt}\,dt = \int_0^{2\pi} \frac{d\mu^a}{df}\,df, \tag{3.71}$$

and division by the orbital period P gives a notion of average, or secular, time derivative:

$$\left(\frac{d\mu^a}{dt}\right)_{\text{sec}} := \frac{\Delta\mu^a}{P}. \tag{3.72}$$

The integrations are best carried out in the form involving the true anomaly f, but the factors of $1 + e\cos f$ in the denominator can sometimes produce integrals that are difficult to evaluate. When such a situation is encountered, it is usually advantageous to make a change of variables from f to the eccentric anomaly u, as defined by Eqs. (3.30)–(3.33).

3.4 Case studies of perturbed Keplerian motion

Our purpose in this section is to explore some of the non-Keplerian aspects of celestial mechanics by working through a number of examples that are of real relevance to the solar system. We shall investigate these effects by exploiting the method of osculating orbital elements introduced in the preceding section.

3.4.1 Perturbations by a third body

We first examine a two-body system, such as the Sun–Mercury system, perturbed by a remote third body such as Jupiter. This is an example of a three-body problem, and we shall have occasion in Sec. 3.5 to give this problem a more complete treatment. Here we provide an approximate discussion, relying on the fact that the third body is remote, so that its gravitational influence on the two-body system is weak. This influence can be described as a perturbation, and its effects can be investigated by means of the method of osculating orbits.

To set up the problem we recall that the two-body system involves a mass m_1 at position r_1 and a mass m_2 at position r_2. The third body has a mass m_3 and a position r_3. As usual, we let $m := m_1 + m_2$, $r := r_1 - r_2$, and we also introduce $R := r_3 - r_2$; ~~note that R n~~ ~~longer stands for the position of the two-body pericenter.~~ We also have $r := |r|$, $R := |R|$, $n := r/r$, and $N := R/R$. We assume that $R \gg r$.

From Eq. (1.202) we find that the acceleration of the first body is given by

$$a_1 = -Gm_2\frac{n}{r^2} - Gm_3\frac{r - R}{|r - R|^3}, \tag{3.73}$$

while

$$a_2 = Gm_1\frac{n}{r^2} + Gm_3\frac{N}{R^2} \tag{3.74}$$

is the acceleration of the second body. The relative acceleration is $a := a_1 - a_2$, and removal of the Keplerian term gives us the perturbing force

$$f = -Gm_3\left(\frac{r - R}{|r - R|^3} + \frac{N}{R^2}\right). \tag{3.75}$$

Taking into account the assumption that R is large compared with r, we simplify this expression by expanding in powers of r/R, and obtain, to leading order,

$$\boldsymbol{f} = -\frac{Gm_3 r}{R^3}\Big[\boldsymbol{n} - 3(\boldsymbol{n}\cdot\boldsymbol{N})\boldsymbol{N} + O(r/R)\Big]. \tag{3.76}$$

We see that the ratio of f to the Keplerian acceleration is of order $(m_3/m)(r/R)^3$, and therefore small by virtue of our assumption that $R \gg r$. The components of the perturbing force in the Keplerian basis $(\boldsymbol{n}, \boldsymbol{\lambda}, \boldsymbol{e}_z)$ are

$$\mathcal{R} := \boldsymbol{f}\cdot\boldsymbol{n} = -\frac{Gm_3 r}{R^3}\big[1 - 3(\boldsymbol{n}\cdot\boldsymbol{N})^2\big], \tag{3.77a}$$

$$[-2pt]\mathcal{S} := \boldsymbol{f}\cdot\boldsymbol{\lambda} = 3\frac{Gm_3 r}{R^3}(\boldsymbol{n}\cdot\boldsymbol{N})(\boldsymbol{\lambda}\cdot\boldsymbol{N}), \tag{3.77b}$$

$$\mathcal{W} := \boldsymbol{f}\cdot\boldsymbol{e}_z = 3\frac{Gm_3 r}{R^3}(\boldsymbol{n}\cdot\boldsymbol{N})(\boldsymbol{e}_z\cdot\boldsymbol{N}), \tag{3.77c}$$

and these can immediately be inserted within the equations that govern the evolution of the osculating orbital elements.

The expressions for $\mathcal{R}$, $\mathcal{S}$, and $\mathcal{W}$ refer to any two-body system perturbed by a remote third body, but to proceed we consider the specific situation mentioned previously: We examine the orbit of a planet like Mercury around the Sun, perturbed by an outer planet such as Jupiter. For simplicity we assume that Mercury and Jupiter move in the same orbital plane (when in reality they have a relative inclination of approximately 6 degrees), and we align the reference X–Y plane with this common orbital plane; this implies that $\iota = 0$, and since Ω is not defined, it may be set equal to zero. Again for simplicity, we assume that Jupiter moves on a circular orbit (a good approximation), with a constant radius R and angular frequency $\Omega_3 = \sqrt{G(m_2 + m_3)/R^3}$. Its true anomaly is given by $F = \Omega_3 t$, and it changes very little in the course of a complete Mercury orbit; we have that

$$\Delta F := \Omega_3 P = 2\pi\left(\frac{m_2 + m_3}{m_1 + m_2}\right)^{1/2}\left(\frac{a}{R}\right)^{3/2} \ll 1, \tag{3.78}$$

with P denoting Mercury's orbital period, and $a := p/(1 - e^2)$ its semi-major axis. The direction to Jupiter is

$$\boldsymbol{N} = \cos F\,\boldsymbol{e}_X + \sin F\,\boldsymbol{e}_Y, \tag{3.79}$$

and Eqs. (3.42), (3.43), and (3.45) reveal that

$$\boldsymbol{n} = \cos(f + \omega)\,\boldsymbol{e}_X + \sin(f + \omega)\,\boldsymbol{e}_Y, \tag{3.80a}$$

$$\boldsymbol{\lambda} = -\sin(f + \omega)\,\boldsymbol{e}_X + \cos(f + \omega)\,\boldsymbol{e}_Y, \tag{3.80b}$$

and $\boldsymbol{e}_z = \boldsymbol{e}_Z$. From these expressions we find that $\boldsymbol{n}\cdot\boldsymbol{N} = \cos(f + \omega - F)$, $\boldsymbol{\lambda}\cdot\boldsymbol{N} = -\sin(f + \omega - F)$, $\boldsymbol{e}_z\cdot\boldsymbol{N} = 0$, and the components of the perturbing force become

$$\mathcal{R} = -\frac{Gm_3 r}{R^3}\big[1 - 3\cos^2(f + \omega - F)\big], \tag{3.81a}$$

$$\mathcal{S} = -3\frac{Gm_3 r}{R^3}\sin(f + \omega - F)\cos(f + \omega - F), \tag{3.81b}$$

$$\mathcal{W} = 0, \tag{3.81c}$$

in which we may substitute $r = p/(1 + e\cos f)$.

Mercury's orbital evolution in response to Jupiter's perturbation is governed by Eqs. (3.64). As we explained near the end of Sec. 3.3.2, we are primarily interested in the *secular variation* of the orbital elements p, e, and ω, as calculated according to Eq. (3.71) using the equations listed in Eq. (3.69). The integrals over f are easily evaluated by adopting the eccentric anomaly u as an integration variable, and by assuming that F is constant over Mercury's orbital cycle – refer to Eq. (3.78). After some manipulations we arrive at

$$\Delta p := -15\pi \frac{m_3 p^4}{m R^3} e^2 (1 - e^2)^{-7/2} \sin 2(\omega - F), \tag{3.82a}$$

$$\Delta e := \frac{15\pi}{2} \frac{m_3 p^3}{m R^3} e (1 - e^2)^{-5/2} \sin 2(\omega - F), \tag{3.82b}$$

$$\Delta \omega := \frac{3\pi}{2} \frac{m_3 p^3}{m R^3} (1 - e^2)^{-5/2} \big[1 + 5 \cos 2(\omega - F) \big] \tag{3.82c}$$

for the net change in the orbital elements after a complete Mercury orbit. From these results it can be inferred that $\Delta a = 0$: Mercury does not undergo a secular variation of its semi-major axis, and the change in $p = a(1 - e^2)$ simply reflects a change in eccentricity.

These expressions reveal that depending on Jupiter's position on its orbit, Mercury's eccentricity e and longitude of perihelion ω either increase or decrease in the course of a complete orbit. After an extremely long time, much longer than Jupiter's own orbital period, the change in e will average out, but Mercury's perihelion will advance at an average rate of

$$\langle \Delta \omega \rangle = \frac{3\pi}{2} \frac{m_3}{m} \left(\frac{p}{R} \right)^3 (1 - e^2)^{-5/2} = \frac{3\pi}{2} \frac{m_3}{m} \left(\frac{a}{R} \right)^3 (1 - e^2)^{1/2} \tag{3.83}$$

per orbit.

Inserting the numerical values for Mercury and Jupiter from Table 3.2, we find that $\langle \Delta \omega \rangle = 1.79 \times 10^{-6}$ radians per orbit; the smallness of this effect provides ample justification for a calculation limited to first-order perturbation theory. It is customary to express the perihelion advance as a rate in units of arcseconds (as) per century. To achieve this conversion we use the fact that 1 rad $= 2.063 \times 10^5$ as, and take into account the fact that Mercury's orbital period is $P = 0.2408$ yr; we obtain $\langle (d\omega/dt)_{\text{sec}} \rangle = 154$ as/century. This is very close to the accurately computed value of 153.6 as/century displayed in Table 3.1. The accuracy, however, would not be as good if we applied Eq. (3.83) to the perturbation produced by Earth's gravitational attraction; here we would get 69.2 as/century, quite a difference from the accurate value of 90 as/century. The reason for this lack of success is that $r/R \simeq 0.38$ for the Earth, so that an accurate computation of the perihelion advance requires higher-order terms in the expansion of the perturbing force in powers of r/R. In addition, the approximation formulated in Eq. (3.78) is not so good in the case of the Earth, because Earth's period is only four times longer than Mercury's (instead of forty times longer for Jupiter). Nevertheless, the method of osculating orbital elements would allow us to incorporate such details systematically in a straightforward manner, leading to the accurately calculated planetary perturbations listed in Table 3.1.

As we have seen in Sec. 3.1, the sum total of the contributions from all the planets to Mercury's perihelion advance does not account for the observed effect. The discrepancy,

Table 3.2 Orbital elements of selected planets. The astronomical unit (AU) is the Earth–Sun distance, equal to 149.60 × 10^6 km. The inclination is measured in degrees, minutes of arc, and seconds of arc. The inverse mass is measured in units of the inverse solar mass, with $M_\odot = 1.9889 \times 10^{30}$ kg.

Planet	Semi-major axis (AU)	Orbital period (yr)	Eccentricity	Inclination to ecliptic ° . ′.″	Inverse mass $1/M_\odot = 1$
Mercury	0.387099	0.24085	0.205628	7.0.15	6010000
Venus	0.723332	0.61521	0.006787	3.23.40	408400
Earth	1.000000	1.00004	0.016722	0.0.0	328910
Mars	1.523691	1.88089	0.093377	1.51.0	3098500
Jupiter	5.202803	11.86223	0.04845	1.18.17	1047.39
Saturn	9.53884	29.4577	0.05565	2.29.22	3498.5

the famous 43 as/century, is accounted for by relativistic corrections to the equations of motion, which can be incorporated as additional contributions to the perturbing force. We will return to this question in Chapter 10.

3.4.2 The Kozai mechanism

In 1962, the Soviet dynamicist Michael Lidov and the Japanese astronomer Yoshihide Kozai independently discovered a remarkable phenomenon that occurs when an object orbits a massive body under the perturbing influence of a distant body, when the two orbits are inclined relative to each other. As a result of the perturbed dynamics, there is a periodic exchange between the orbital eccentricity and inclination, such that when one increases, the other decreases. This effect, known as the *Kozai mechanism*, has important implications for the behavior of asteroids, satellites of planets, extrasolar planets, and multiple-star systems.

To study this effect we re-examine the situation of the previous section, but now allow the orbiting body (m_1) to have an inclination ι relative to the orbit of the perturbing body (m_3), which is placed in the reference X–Y plane. For simplicity we take the perturbing body to move on a circular orbit, and we choose the line of nodes of the inclined orbit to be situated at $\Omega = 0$. The components of the perturbing force are still given by Eqs. (3.77), but the orbital basis vectors are now

$$\boldsymbol{n} = \cos(\omega + f)\,\boldsymbol{e}_X + \cos\iota \sin(\omega + f)\,\boldsymbol{e}_Y + \sin\iota \sin(\omega + f)\,\boldsymbol{e}_Z, \tag{3.84a}$$

$$\boldsymbol{\lambda} = -\sin(\omega + f)\,\boldsymbol{e}_X + \cos\iota \cos(\omega + f)\,\boldsymbol{e}_Y + \sin\iota \cos(\omega + f)\,\boldsymbol{e}_Z, \tag{3.84b}$$

$$\boldsymbol{e}_z = -\sin\iota\,\boldsymbol{e}_Y + \cos\iota\,\boldsymbol{e}_Z; \tag{3.84c}$$

the direction to the perturbed body is still given by Eq. (3.79). We calculate the secular variation of the orbital elements a, e, ω, and ι by inserting these expressions within Eqs. (3.69),

integrating over a complete orbital cycle, and averaging over the position of the perturbing body. The final results are

$$\langle \Delta a \rangle = 0, \tag{3.85a}$$

$$\langle \Delta e \rangle = \frac{15\pi}{2} \frac{m_3}{m} \left(\frac{a}{R} \right)^3 e(1 - e^2)^{1/2} \sin^2 \iota \sin \omega \cos \omega, \tag{3.85b}$$

$$\langle \Delta \omega \rangle = \frac{3\pi}{2} \frac{m_3}{m} \left(\frac{a}{R} \right)^3 (1 - e^2)^{-1/2} \Big[5 \cos^2 \iota \sin^2 \omega + (1 - e^2)(5 \cos^2 \omega - 3) \Big], \tag{3.85c}$$

$$\langle \Delta \iota \rangle = -\frac{15\pi}{2} \frac{m_3}{m} \left(\frac{a}{R} \right)^3 e^2(1 - e^2)^{-1/2} \sin \iota \cos \iota \sin \omega \cos \omega. \tag{3.85d}$$

From Eqs. (3.85b) and (3.85d) we notice that $e(1 - e^2)^{-1} \langle \Delta e \rangle + \tan \iota \langle \Delta \iota \rangle = 0$, which leads to the remarkable conservation law,

$$\sqrt{1 - e^2} \, \cos \iota = \text{constant (over long time scales).} \tag{3.86}$$

This is actually the consequence of angular-momentum conservation. The component of the total angular-momentum vector normal to the orbital plane of the perturbing body is $L_Z = L_Z^{1+2} + L_Z^3$, and if we assume that L_Z^3 is separately conserved (thereby ignoring the gravitational effects of m_1 and m_2 on the perturbing body), we find that $L_Z^{1+2} = \mu \boldsymbol{h} \cdot \boldsymbol{e}_Z = \mu \sqrt{Gmp} \cos \iota$ must be conserved. Because $a := p/(1 - e^2)$ does not vary over long time scales, conservation of angular momentum implies that $(1 - e^2)^{1/2} \cos \iota$ cannot vary, in agreement with our previous statement.

When the longitude of pericenter ω lies in the first or third quadrants, we see from Eqs. (3.85) that the eccentricity increases while the inclination decreases; when ω lies instead in the second or fourth quadrants, the eccentricity decreases while the inclination increases. The equations also reveal that as long as $\cos^2 \iota / (1 - e^2) > 3/5$, ω increases monotonically, on the same time scale as the variations in e and ι. For orbits with sufficiently low inclinations, therefore, there is a periodic exchange between eccentricity and inclination as the pericenter keeps on advancing.

For high-inclination orbits, such that $\cos^2 \iota / (1 - e^2) < 3/5$, the orbital evolution is very different. In this case, ω approaches a critical angle ω_c such that $\langle \Delta \omega \rangle = 0$, and the motion of the pericenter ceases; the critical angle is determined by

$$\cos^2 \omega_c = \frac{3(1 - e^2) - 5 \cos^2 \iota}{5(1 - e^2 - \cos^2 \iota)}. \tag{3.87}$$

Meanwhile, if e is initially in an increasing phase and ι in a decreasing phase, the ratio $\cos^2 \iota / (1 - e^2)$ will be driven toward the limiting value of $3/5$, and ω will settle to a final value of $\omega_c = \frac{\pi}{2}$ or $\omega_c = \frac{3\pi}{2}$. The final outcome is a stationary solution of the perturbation equations with $\langle \Delta e \rangle = \langle \Delta \iota \rangle = \langle \Delta \omega \rangle = 0$, occurring at critical values e_c, ι_c, and ω_c constrained by

$$\omega_c = \frac{\pi}{2} \text{ or } \frac{3\pi}{2}, \qquad \frac{\cos^2 \iota_c}{1 - e_c^2} = \frac{3}{5}. \tag{3.88}$$

This equilibrium is stable; it can be shown that a perturbation away from equilibrium leads to oscillations of e, ι, and ω about their critical values, with a period

$$P_{\text{Kozai}} = \frac{2}{3\sqrt{15}} \frac{m}{m_3} \left(\frac{R}{a}\right)^3 \frac{1}{e_c \sin \iota_c} P,$$ (3.89)

where P is the period of the orbiting body. This stationary solution is called the *Kozai resonance*. You will be asked to verify these statements in Exercise 3.7.

The Kozai mechanism can have important consequences: as the orbital eccentricity increases, the orbiting body can pass sufficiently close to the inner regions of the system at pericenter to interact with other bodies there, or even be tidally captured or disrupted by the central body. Conversely, the body could find itself sufficiently far away at apocenter to interact with more distant objects and be ejected from the system. This mechanism accounts for many observed features of the asteroid belts, of minor satellites of Jupiter, and of objects within the Kuiper belt. As an example, many high-eccentricity comets are found to be in Kozai resonances, with their longitude of pericenter in the vicinity of $\frac{\pi}{2}$ or $\frac{3\pi}{2}$.

3.4.3 Effects of oblateness

As our next case study we examine a two-body system perturbed by the oblateness of one of its members. As we saw back in Sec. 2.3, the oblateness is most often due to rotational flattening, and here the deformed body is taken to have an axisymmetric shape; the deformation from spherical symmetry is measured by its quadrupole moment, and we neglect the gravitational influence of higher-order multipole moments. Examples of such systems abound. We could be dealing with Mercury orbiting an oblate, rotating Sun, or an artificial satellite orbiting the Earth, or else a neutron star orbiting a rotating, normal star.

The equations of motion for such a system were worked out back in Sec. 1.6.7, and according to Eq. (1.227), the perturbing force is given by

$$\boldsymbol{f} = \frac{3}{2} J_2 \frac{GmR^2}{r^4} \left\{ \left[5(\boldsymbol{e} \cdot \boldsymbol{n})^2 - 1 \right] \boldsymbol{n} - 2(\boldsymbol{e} \cdot \boldsymbol{n}) \boldsymbol{e} \right\},$$ (3.90)

with J_2 denoting the dimensionless quadrupole moment of the deformed body, R its radius, and the unit vector $\boldsymbol{e}$ indicating the direction of the symmetry axis. As usual we have $m := m_1 + m_2$, $\boldsymbol{r} := \boldsymbol{r}_1 - \boldsymbol{r}_2$, $r := |\boldsymbol{r}|$, and $\boldsymbol{n} := \boldsymbol{r}/r$. In Eq. (3.90) we specifically identify the oblate body with m_2, and m_1 is assumed to be spherical; a swap of identities involves changing the sign of $\boldsymbol{n}$, and therefore the sign of the perturbing force.

We choose the reference X–Y plane to be orthogonal to the symmetry axis of m_2, so that $\boldsymbol{e}$ is aligned with the Z-direction. The orbital plane has an inclination ι relative to the reference plane, and the orbital basis is given by Eqs. (3.42)–(3.45). According to these equations, $\boldsymbol{e} \cdot \boldsymbol{n} = \sin \iota \sin(\omega + f)$, $\boldsymbol{e} \cdot \boldsymbol{\lambda} = \sin \iota \cos(\omega + f)$, $\boldsymbol{e} \cdot \boldsymbol{e}_z = \cos \iota$, and the components of the

perturbing force in the orbital basis are

$$\mathcal{R} = \frac{3}{2} J_2 \frac{GmR^2}{r^4} \left[3 \sin^2 \iota \sin^2(\omega + f) - 1 \right], \tag{3.91a}$$

$$\mathcal{S} = -3 J_2 \frac{GmR^2}{r^4} \sin^2 \iota \sin(\omega + f) \cos(\omega + f), \tag{3.91b}$$

$$\mathcal{W} = -3 J_2 \frac{GmR^2}{r^4} \sin \iota \cos \iota \sin(\omega + f), \tag{3.91c}$$

with $r = p/(1 + e \cos f)$.

We calculate the secular variation of the orbital elements p, e, ι, ω, and Ω by inserting these expressions within Eqs. (3.69), and integrating over a complete orbital cycle. (In this case there is no need to involve the eccentric anomaly u in the evaluation of the integrals; the factor of r^{-4} in the perturbing force ensures that there are no remaining factors of $1 + e \cos f$ in the denominators, and all integrations are elementary.) The end results are

$$\Delta p = 0, \tag{3.92a}$$

$$\Delta e = 0, \tag{3.92b}$$

$$\Delta \iota = 0, \tag{3.92c}$$

$$\Delta \omega = 6\pi J_2 \left(\frac{R}{p} \right)^2 \left(1 - \frac{5}{4} \sin^2 \iota \right), \tag{3.92d}$$

$$\Delta \Omega = -3\pi J_2 \left(\frac{R}{p} \right)^2 \cos \iota. \tag{3.92e}$$

The first two equations imply that $\Delta a = 0$; there is no secular change in the eccentricity, no secular change in the semi-major axis, and no secular change in the inclination. There are, however, secular changes in the line of nodes (measured by $\Delta \Omega$) and in the pericenter (measured relative to the line of nodes by $\Delta \omega$). The last two equations imply

$$\Delta \omega + \cos \iota \, \Delta \Omega = 3\pi J_2 \left(\frac{R}{p} \right)^2 \left(1 - \frac{3}{2} \sin^2 \iota \right); \tag{3.93}$$

as we saw back in Sec. 3.3.2, this is the pericenter advance relative to the reference X-direction, as measured in the orbital plane. At zero inclination, when $e \cdot n = 0$, the changes are produced by the r^{-4} modification to the gravitational acceleration, which is still directed along n. For inclined orbits, the angular dependence of the perturbing force produces an additional contribution. An associated effect is a precession of the angular-momentum vector h around the symmetry axis e, which produces a rotation of the line of nodes. The precession is predicted by Eq. (3.51), which becomes

$$\frac{dh}{dt} = 3 J_2 \frac{GmR^2}{r^3} (e \cdot n)(e \times n) \tag{3.94}$$

in the case of the perturbing force of Eq. (3.90). This describes a precession because h does not change on a secular time scale, and because $e \cdot h$ stays constant during the orbital evolution.

For Mercury, the perihelion advance induced by the Sun's oblateness is negligible. Inserting the relevant orbital elements and the value $(J_2)_\odot = 2.2 \times 10^{-7}$, we obtain $(d\omega/dt)_{\text{sec}} + \cos\iota(d\Omega/dt)_{\text{sec}} = 0.03$ as/century, which is just below the observational uncertainties assigned to the measurement of the advance (refer to Table 3.1). The conclusion relies on a reliable determination of the Sun's oblateness, something that was difficult to come by. In fact, it is extremely difficult to measure the Sun's J_2 directly, because at planetary distances, its effects are just too small to be measured. The best way to obtain a reliable estimate would be to send a spacecraft to a low solar orbit, well inside the orbit of Mercury, and to measure precisely how J_2 affects its orbit; but despite a number of proposals for such a mission over the years, none has yet come to pass.

During the 1960s, Robert Dicke and Mark Goldenberg attempted to determine J_2 by measuring the Sun's visual shape. Because the surface of the Sun is an equipotential (recall the discussion in Sec. 2.3.1), its shape is affected by J_2 in a way that can be directly related to the deformation of the external gravitational field. The shape was measured by inserting a circular, opaque disk in front of a telescopic image of the Sun, leaving only a thin visible ring at the edge of the Sun, and measuring the difference in brightness of the visible ring between the equator and pole of the Sun. If the Sun were oblate, the ring at the equator should extend further beyond the occulting disk, and should therefore be brighter. But many factors had to be corrected for, including the effects of atmospheric distortion on the observed shape of the Sun, and the effects of possible temperature differences between the polar and equatorial regions of the Sun, which would lead to brightness differences not associated with the shape. Dicke and Goldenberg claimed to have measured a J_2 of the order of 2.5×10^{-5}, over 100 times larger than the currently accepted value. Dicke postulated that such a large oblateness would occur if the core of the Sun were rotating much faster than its outer layers, thereby generating more centrifugal flattening than would be expected on the basis of the observed surface rotation alone.

A value of J_2 this large would mean that solar oblateness contributes as much as 4 as/century to Mercury's perihelion advance, which would destroy the agreement of the measured advance with the prediction of general relativity. But it would have supported the scalar-tensor theory of gravity (to be presented in Chapter 13) that Dicke himself had developed with Carl Brans, whose prediction for the relativistic part of the advance could be fit to 39 as/century by tuning a coupling constant. (The flexibility to tune the Brans–Dicke parameter to obtain this low value was short lived; accurate observations of the relativistic light deflection eventually constrained the coupling constant to such an extent that the Brans–Dicke prediction for the perihelion advance of Mercury became only marginally different from Einstein's.) Later observations of the visible shape of the Sun by Henry Hill and others, along with observations to try to better understand the temperature differences, did not fully resolve this controversy.

The resolution came with the advance of helioseismology. This was the discovery that the Sun vibrates in a superposition of thousands of normal modes with an array of frequencies, as could be observed by measuring the frequency spectrum of Doppler-shifted solar spectral lines. The specific pattern of frequencies depends on the Sun's angular-velocity profile. Through a systematic program of ground-based and space-based observations of the Sun, it became possible to determine the Sun's rotational profile over much of its interior. The

conclusion was that the core does not rotate much faster than the surface, and solar models consistent with this information produced the currently accepted value of $J_2 = 2.2 \times 10^{-7}$; this is approximately what one would infer from a Sun that rotates uniformly at its observed surface rate. The bottom line is that, as far as Mercury's motion and general relativity are concerned, the solar quadrupole moment does not play a significant role.

The quadrupole moment of the Earth does play an important role in the motion of artificial satellites. The effect on the line of nodes can be examined for a satellite placed on a circular orbit with radius a. Using $(J_2)_\oplus = 1.08 \times 10^{-3}$ and dividing $\Delta\Omega$ by the satellite's orbital period $P = 2\pi a^{3/2}/(Gm)^{1/2} = 83.91(a/R)^{3/2}$ minutes, we can show that

$$\left(\frac{d\Omega}{dt}\right)_{\text{sec}} = -3639 \left(\frac{R}{a}\right)^{7/2} \cos\iota \quad \text{degrees/year}, \tag{3.95}$$

where R is the Earth's radius. For example, the orbit of the Laser Geodynamics Satellite (LAGEOS) I, with $a = 1.93R$ and $\iota = 109°.8$, precesses at a rate of 120 degrees per year. A satellite with $a = 1.5R$ and $\iota = 65.9°$ would precess at a rate of 360 degrees per year. Such an orbit is called *Sun synchronous* because it always presents the same face to the Sun, varying only in its inclination relative to the ecliptic because of the 23.5° tilt of the Earth's spin axis.

3.4.4 Tidally interacting bodies

In our next case study we examine a system of two bodies that deform each other by tidal interactions. The physics of tides was reviewed in some depth in Sec. 2.5, and here we consider a situation in which the orbital period $P = 2\pi a^{3/2}/(Gm)^{1/2}$ is long compared with the hydrodynamical time scale $T_{\text{int}} \sim (G\rho)^{-1/2} \sim R^{3/2}/(Gm)^{1/2}$ that characterizes the internal dynamics of each body; this is the realm of *static tides*.

We first take the bodies to be non-rotating. According to the theory developed in Sec. 2.5.1, a body of radius R subjected to a tidal potential $U_{\text{tidal}} = -\frac{1}{2}\mathcal{E}_{jk}x^j x^k$ acquires a mass quadrupole moment given by $GI_{\langle jk\rangle} = -\frac{2}{3}k_2 R^5 \mathcal{E}_{jk}$, with k_2 denoting the body's gravitational Love number, which depends on the details of its internal structure. In the case of a two-body system, we have a first body of mass m_1, radius R_1, Love number $(k_2)_1$, and quadrupole moment $I_1^{\langle jk\rangle}$, and a second body of mass m_2, radius R_2, Love number $(k_2)_2$, and quadrupole moment $I_2^{\langle jk\rangle}$. According to Eq. (2.275), the tidal quadrupole moment created by m_2 and acting on m_1 is given by $\mathcal{E}_1^{jk} = -(Gm_2/r^3)(3n^j n^k - \delta^{jk})$, while the tidal moment created by m_1 and acting on m_2 is $\mathcal{E}_2^{jk} = -(Gm_1/r^3)(3n^j n^k - \delta^{jk})$; as usual, we have that $m := m_1 + m_2$, $\boldsymbol{r} := \boldsymbol{r}_1 - \boldsymbol{r}_2$, $r := |\boldsymbol{r}|$, and $\boldsymbol{n} := \boldsymbol{r}/r$. From all this we find that the tidally-induced mass quadrupole moment of the first body is given by

$$I_1^{\langle jk\rangle} = \frac{2}{3}(k_2)_1 \frac{m_2 R_1^5}{r^3}\left(3n^j n^k - \delta^{jk}\right), \tag{3.96}$$

while

$$I_2^{\langle jk\rangle} = \frac{2}{3}(k_2)_2 \frac{m_1 R_2^5}{r^3}\left(3n^j n^k - \delta^{jk}\right) \tag{3.97}$$

is the mass quadrupole moment of the second body.

The equations of motion for two deformed bodies were obtained back in Sec. 1.6.7, and Eq. (1.220) provides an expression for the relative acceleration $\boldsymbol{a} := \boldsymbol{a}_1 - \boldsymbol{a}_2$. If we remove the Keplerian term, truncate the multipole expansions to the leading, quadrupole order, and neglect the quadrupole–quadrupole interaction (which is much smaller than the monopole–quadrupole interactions that we keep), we obtain

$$ f^j = \frac{1}{2} Gm \left(\frac{I_1^{\langle kn \rangle}}{m_1} + \frac{I_2^{\langle kn \rangle}}{m_2} \right) \partial_{jkn} \left(\frac{1}{r} \right) \tag{3.98} $$

for the perturbing force. Inserting the preceding expressions for the mass quadrupole moments, as well as Eq. (1.152) for $\partial_{jkn} r^{-1}$, we find that this becomes

$$ \boldsymbol{f} = -6 \frac{Gm}{r^7} \left((k_2)_1 \frac{m_2}{m_1} R_1^5 + (k_2)_2 \frac{m_1}{m_2} R_2^5 \right) \boldsymbol{n}. \tag{3.99} $$

The perturbing force is directed along $\boldsymbol{n}$, and its only non-vanishing component is $\mathcal{R}$. With $\mathcal{S} = \mathcal{W} = 0$, we see that the tidal interaction has no effect on p, ι, and Ω; the only orbital elements that can undergo a change are e and ω.

We insert our expression for $\mathcal{R}$ within Eqs. (3.69), integrate over a complete orbital cycle, and obtain the secular changes $\Delta e = 0$ and

$$ \Delta \omega = 30\pi \left(1 + \frac{3}{2} e^2 + \frac{1}{8} e^4 \right) \left[(k_2)_1 \frac{m_2}{m_1} \left(\frac{R_1}{p} \right)^5 + (k_2)_2 \frac{m_1}{m_2} \left(\frac{R_2}{p} \right)^5 \right]. \tag{3.100} $$

These results imply that e (and therefore a) do not undergo a secular change, but the pericenter advances at a steady rate that depends on the orbital parameters, the body radii, and the gravitational Love numbers. Astronomers call this phenomenon the *apsidal advance*, and for a class of close binaries, this effect gives an important clue to the internal structure of each star, via the ability to infer k_2. These systems have masses of the order of a few solar masses, orbital periods of the order of 10 days, and modest eccentricities. The resulting apsidal advance can then be expressed as the rate

$$ \left(\frac{d\omega}{dt} \right)_{\text{sec}} = 0.06 \, f(e) \left(\frac{M_\odot}{m} \right)^{5/3} \left(\frac{10 \text{ days}}{P} \right)^{13/3} $$

$$ \times \left[\frac{(k_2)_1}{0.01} \frac{m_2}{m_1} \left(\frac{R_1}{R_\odot} \right)^5 + \frac{(k_2)_2}{0.01} \frac{m_1}{m_2} \left(\frac{R_2}{R_\odot} \right)^5 \right] \text{ degrees/century}, \tag{3.101} $$

where $f(e) = (1 + \frac{3}{2} e^2 + \frac{1}{8} e^4)/(1 - e^2)^5$. Many close binaries are eclipsing, and by combining light-curve data on the timing of eclipses with spectroscopic data on the orbital motion, it is possible to determine the masses and radii as well as the apsidal advance rate, and thereby estimate k_2 for each star. In most cases, there is good agreement between the values of k_2 obtained by implementing this method and those determined by stellar models for stars having the observed masses and spectra. As we shall see in Chapter 10, the general relativistic contribution to the apsidal advance is small, generally a few percent of the tidally induced advance.

Box 3.4	*DI Herculis*: A tidal troublemaker

There are a few systems for which the nice picture of a tidally induced apsidal advance does not work. The most famous is *DI Herculis*, a high eccentricity ($e \simeq 0.48$), high mass ($m \simeq 10\ M_\odot$) system with a 10.55 day period. In this case, the relativistic contribution to the apsidal advance is large, 2.34 degrees per century, while the tidal contribution is predicted to be 1.93 degrees per century, giving a total predicted advance of 4.27 degrees per century. Unfortunately, the *observed* apsidal advance is only $1.00 \pm .30$ degrees per century. While there were a few attempts to use this discrepancy as evidence against general relativity, more conventional explanations involve purely Newtonian mechanisms. First, there could be a large misalignment between the rotation axis of each star and the orbital plane, such that the contribution of the rotationally-induced quadrupole moments to $\Delta\omega$ would be negative, or retrograde. Second, there could be a third star in the vicinity of the system, whose perturbative effects on all the orbital elements would complicate how $\Delta\omega$ is inferred from the eclipse light-curve data, leading to a value in better agreement with the prediction. The book is not yet closed on *DI Herculis*.

Our discussion thus far has excluded the dissipative aspects of tidal dynamics that were reviewed in Sec. 2.5.2. To explore these effects we examine specifically the case of a moon orbiting a planet. We take the first body to be the planet, the second body to be the moon, and we place the bodies on a circular orbit of relative separation r. Both bodies are now rotating, and we take the moon to be tidally locked, so that it rotates with the same angular velocity as the orbit. The planet, on the other hand, rotates with an angular velocity ω_1 that differs, in general, from the orbital angular velocity Ω.

According to Eq. (2.277), the quadrupole moment of each body is given by

$$GI_{\langle jk \rangle} = -\frac{2}{3}k_2 R^5\left(\mathcal{E}_{jk} - \tau\dot{\mathcal{E}}_{jk}\right) \tag{3.102}$$

in the body's rotating frame. The effect of the first term was investigated previously, and we henceforth ignore it. We keep the second term, the one associated with the dissipative aspects of the tidal interaction; the parameter τ is the body's viscous delay. Because the moon is co-rotating with the orbit, the tidal field measured in its rotating frame is constant, and this implies that we can set $I_2^{\langle jk \rangle} = 0$; for our purposes here, the moon is not deformed by the tidal forces exerted by the planet. The planet, however, is deformed by the moon's tidal field, and we find that its quadrupole moment is given by

$$I_1^{\langle jk \rangle} = -2(k_2)_1 \frac{m_2 R_1^5}{r^3}(\Omega - \omega_1)\tau_1\left(n^j\lambda^k + \lambda^j n^k\right) \tag{3.103}$$

in the rotating frame. Inserting this within the perturbing force of Eq. (3.98), we find that its dissipative piece is given by

$$\boldsymbol{f}_{\rm diss} = -6(k_2)_1 Gm\frac{m_2}{m_1}\frac{R_1^5}{r^7}(\Omega - \omega_1)\tau_1\,\boldsymbol{\lambda}. \tag{3.104}$$

Its only non-vanishing component is $\mathcal{S}_{\rm diss} := \boldsymbol{f}_{\rm diss} \cdot \boldsymbol{\lambda}$.

The dissipative piece of the perturbing force can now be inserted within the osculating equations (3.69). We find that r is the only orbital element that is affected by the perturbation, and that it changes by

$$\Delta r = -24\pi (k_2)_1 \frac{m_2}{m_1} \frac{R_1^5}{r^4} (\Omega - \omega_1)\tau_1 \tag{3.105}$$

in the course of a complete orbital cycle. The sign of the change is related to the sign of $\Omega - \omega_1$. When $\omega_1 < \Omega$, that is, when the moon orbits faster than the planet rotates, we find that $\Delta r < 0$, so that the moon slowly approaches the planet in a shrinking orbit; an example of such a situation is found in the Mars–Phobos system. When $\omega_1 > \Omega$, that is, when the planet rotates faster than the moon orbits, we find instead that $\Delta r > 0$, so that the moon recedes from the planet in an expanding orbit; the Earth–Moon system is a familiar example of such a situation. These conclusions are in accord with our discussion of angular-momentum transfer in Sec. 2.5.2. We saw that the planet gains spin angular momentum when $\omega_1 < \Omega$, and the shrinking orbit confirms that this gain comes at the expense of the orbital angular momentum. Similarly, the planet loses its spin when $\omega_1 > \Omega$, and this loss allows the orbit to gain angular momentum and to expand.

3.4.5 Luni-solar precession of the Earth

In this final case study we describe how the gravitational attraction of the Moon and the Sun on an oblate Earth causes a precession of its spin around the normal to the ecliptic plane. This is the origin of the famous *precession of the equinoxes*, with its period of approximately 26 000 years.

The motion of a body's spin vector $\boldsymbol{S}$ subjected to gravitational torques exerted by other bodies was determined back in Sec 1.6.8. According to Eq. (1.240), an axisymmetric body of mass m, radius R, and dimensionless quadrupole moment J_2, placed in the gravitational field of a first mass m_1 at a distance r_1, and of a second mass m_2 at a distance r_2, undergoes a spin precession described by

$$\frac{d\boldsymbol{S}}{dt} = -3GmJ_2R^2 \left[\frac{m_1}{r_1^3}(\boldsymbol{e}\cdot\boldsymbol{n}_1)(\boldsymbol{e}\times\boldsymbol{n}_1) + \frac{m_2}{r_2^3}(\boldsymbol{e}\cdot\boldsymbol{n}_2)(\boldsymbol{e}\times\boldsymbol{n}_2) \right]. \tag{3.106}$$

Here, $\boldsymbol{e}$ is the direction of the body's rotation axis, so that $\boldsymbol{S} = S\boldsymbol{e}$, $\boldsymbol{n}_1$ is the direction to m_1, and $\boldsymbol{n}_2$ the direction to m_2. Note that in this application, m stands for the body's mass instead of the total mass of the system. It is easy to see from Eq. (3.106) that S is a constant, so that $d\boldsymbol{S}/dt = Sd\boldsymbol{e}/dt$.

For simplicity, and as a rather good approximation, we take the Moon and the Sun to move on circular orbits in the ecliptic plane, which we identify with the reference X–Y plane. The direction to the Moon is then

$$\boldsymbol{n}_1 = \cos(\Omega_1 t)\,\boldsymbol{e}_X + \sin(\Omega_1 t)\,\boldsymbol{e}_Y, \tag{3.107}$$

with Ω_1 denoting the Moon's angular velocity, while the direction to the Sun is

$$\boldsymbol{n}_2 = \cos(\Omega_2 t + \psi)\,\boldsymbol{e}_X + \sin(\Omega_2 t + \psi)\,\boldsymbol{e}_Y, \tag{3.108}$$

with Ω_2 denoting the Sun's angular velocity, and ψ representing the Sun's initial phase on its orbit. The Earth's rotation axis is tilted by an angle $\alpha = 23.4°$ relative to the ecliptic, and we express its direction as

$$\boldsymbol{e} = \sin\alpha\cos\beta\,\boldsymbol{e}_X + \sin\alpha\sin\beta\,\boldsymbol{e}_Y + \cos\alpha\,\boldsymbol{e}_Z, \tag{3.109}$$

in terms of a time-dependent precessional angle β. As we shall see, Eq. (3.106) implies that α remains constant on a secular time scale.

We insert these expressions within Eq. (3.106) and average the result over a large number of orbital cycles. We obtain

$$\left(\frac{d\boldsymbol{e}}{dt}\right)_{\text{sec}} = -\frac{3GmJ_2R^2}{2S}\sin\alpha\cos\alpha\left(\frac{m_1}{r_1^3} + \frac{m_2}{r_2^3}\right)\left(-\sin\beta\,\boldsymbol{e}_X + \cos\beta\,\boldsymbol{e}_Y\right). \tag{3.110}$$

Substitution of Eq. (3.109) on the left-hand side confirms that $(d\alpha/dt)_{\text{sec}} = 0$, and gives rise to a rate of change for the precessional angle,

$$\left(\frac{d\beta}{dt}\right)_{\text{sec}} = -\frac{3GmJ_2R^2}{2S}\cos\alpha\left(\frac{m_1}{r_1^3} + \frac{m_2}{r_2^3}\right). \tag{3.111}$$

Inserting the relevant numbers, including $S = 5.86 \times 10^{33}$ kg m^2/s and $J_2 = 1.08 \times 10^{-3}$, we obtain $|(d\beta/dt)_{\text{sec}}| = 7.74 \times 10^{-12}$ rad/s, or 5040 as/century, which is very close to the accepted value of 5029 as/century. The precessional period $P_{\text{prec}} = 2\pi\,|(d\beta/dt)_{\text{sec}}|^{-1}$ amounts to the previously quoted value of 26 000 years.

3.5 More bodies

We saw back in Sec. 3.2 that the problem of two spherical bodies moving under their mutual gravitational attraction can be solved completely and exactly. The same cannot be said, however, when the bodies have significant deformations from spherical symmetry, or when we go beyond two bodies; in such cases there are no exact solutions, except in very special situations. In this section we touch briefly on the three-body problem, and say a few words about the general N-body problem; throughout this section the bodies are assumed to be spherical.

3.5.1 The 3-body problem

General problem

Adding just one body to a two-body system brings a remarkable complexity to the problem of motion. The equations of motion are simple enough to write down. In the familiar

notation we have

$$a_1 = -Gm_2\frac{r_{12}}{r_{12}^3} - Gm_3\frac{r_{13}}{r_{13}^3}, \tag{3.112a}$$

$$a_2 = +Gm_1\frac{r_{12}}{r_{12}^3} - Gm_3\frac{r_{23}}{r_{23}^3}, \tag{3.112b}$$

$$a_3 = +Gm_1\frac{r_{13}}{r_{13}^3} + Gm_2\frac{r_{23}}{r_{23}^3}, \tag{3.112c}$$

where, for example, $r_{12} := r_1 - r_2$. According to the discussion of Sec. 1.6.6, the equations of motion admit the conserved quantities

$$P = m_1v_1 + m_2v_2 + m_3v_3, \tag{3.113a}$$

$$L = m_1r_1 \times v_1 + m_2r_2 \times v_2 + m_3r_3 \times v_3, \tag{3.113b}$$

$$E = \frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 + \frac{1}{2}m_1v_2^2 - \frac{Gm_1m_2}{r_{12}} - \frac{Gm_1m_3}{r_{13}} - \frac{Gm_2m_3}{r_{23}}. \tag{3.113c}$$

Momentum conservation implies that the system's barycenter, situated at

$$R := \frac{1}{M}(m_1r_1 + m_2r_2 + m_3r_3), \tag{3.114}$$

with $M := m_1 + m_2 + m_3$, moves uniformly according to $R(t) := R(0) + Pt/M$. We shall describe the 3-body system in the *barycentric frame*, and set $R = 0$ at all times.

In the two-body problem it proved helpful to write the equations of motion in terms of the separation vector between the two bodies, which produces a reduction to an effective one-body problem. Here we follow Jacobi and accomplish a similar reduction to an effective two-body problem by introducing the alternative variables

$$r := r_1 - r_2, \tag{3.115a}$$

$$\rho := r_3 - \frac{m_1}{m}r_1 - \frac{m_2}{m}r_2, \tag{3.115b}$$

in which $m := m_1 + m_2$; the vector r is the separation between body 1 and body 2, while ρ is the separation between body 3 and the barycenter of the 1–2 subsystem. The barycenter condition $R = 0$ can be used to express each position vector in terms of the new variables; we find that

$$r_1 = \frac{m_2}{m}r - \frac{m_3}{M}\rho, \tag{3.116a}$$

$$r_2 = -\frac{m_1}{m}r - \frac{m_3}{M}\rho, \tag{3.116b}$$

$$r_3 = \frac{m}{M}\rho. \tag{3.116c}$$

Similar relations hold between the individual velocity vectors and the relative velocities $v := dr/dt$ and $V := d\rho/dt$. The separation vectors become $r_{12} = r$, $r_{13} = -(\rho - m_2r/m)$, and $r_{23} = -(\rho + m_1r/m)$ in the Jacobi variables.

Making the substitutions, we quickly find that the equations of motion of the effective two-body problem are

$$\frac{d^2\boldsymbol{r}}{dt^2} = -Gm\frac{\boldsymbol{r}}{r^3} + Gm_3\left(\frac{\boldsymbol{\rho} - m_2\boldsymbol{r}/m}{|\boldsymbol{\rho} - m_2\boldsymbol{r}/m|^3} - \frac{\boldsymbol{\rho} + m_1\boldsymbol{r}/m}{|\boldsymbol{\rho} + m_1\boldsymbol{r}/m|^3}\right), \tag{3.117a}$$

$$\frac{d^2\boldsymbol{\rho}}{dt^2} = -\frac{GM}{m}\left(m_1\frac{\boldsymbol{\rho} - m_2\boldsymbol{r}/m}{|\boldsymbol{\rho} - m_2\boldsymbol{r}/m|^3} + m_2\frac{\boldsymbol{\rho} + m_1\boldsymbol{r}/m}{|\boldsymbol{\rho} + m_1\boldsymbol{r}/m|^3}\right), \tag{3.117b}$$

and that the conserved quantities become

$$\boldsymbol{L} = \frac{m_1 m_2}{m}\boldsymbol{r} \times \boldsymbol{v} + \frac{mm_3}{M}\boldsymbol{\rho} \times \boldsymbol{V}, \tag{3.118a}$$

$$E = \frac{1}{2}\frac{m_1 m_2}{m}v^2 + \frac{1}{2}\frac{mm_3}{M}V^2 - \frac{Gm_1 m_2}{r} - \frac{Gm_1 m_3}{|\boldsymbol{\rho} - m_2\boldsymbol{r}/m|} - \frac{Gm_2 m_3}{|\boldsymbol{\rho} + m_1\boldsymbol{r}/m|} \tag{3.118b}$$

in the Jacobi variables.

The equations of motion (3.117) are extraordinarily complex, and in spite of massive efforts by the best physicists and mathematicians of past centuries, including Lagrange, Laplace, Jacobi, Hill, and especially Poincaré, the solution space is still poorly understood today. Special solutions are known. For example, when the three bodies are confined to a plane it is possible to find solutions with $r_{12} = r_{13} = r_{23}$, so that the bodies form an equilateral triangle. In this configuration, each body follows its own elliptical orbit around the barycenter, with the triangle expanding and shrinking as the bodies pass simultaneously through their apocenter and pericenter; in the special case of circular orbits, the three bodies rotate as a rigid equilateral triangle around the barycenter. Another example is the "figure-of-eight" solution, in which three equal-mass bodies chase each other on a closed orbit that takes the shape of the number eight; the solution has zero total angular momentum. Understanding the three-body problem is of considerable astrophysical interest, since over 500 stars in our galaxy have been identified as triple systems, including the north star Polaris and our nearest neighbor, Alpha Centauri.

Restricted problem

There is a special case of the three-body problem that offers sufficient simplicity to permit a detailed exploration. Known as the *restricted three-body problem*, this is a situation in which one of the three bodies has a negligible mass compared to the others, and the other two bodies are placed on a circular orbit, undisturbed by the third body. In this case the three-body problem reduces to an independent two-body problem with a specified solution (the circular orbit), and the decoupled problem of a test body moving in the gravitational field of the orbiting bodies.

We take the light body to be m_3, and let $m_3 \ll m_1$ and $m_3 \ll m_2$, so that $M \simeq m = m_1 + m_2$. In this limit the equations of motion become $d^2\bar{\boldsymbol{r}}/dt^2 = -Gm\bar{\boldsymbol{r}}/r^3$ for the two-body system, and

$$\frac{d^2\bar{\boldsymbol{\rho}}}{dt^2} = -Gm_1\frac{\bar{\boldsymbol{\rho}} - m_2\bar{\boldsymbol{r}}/m}{|\bar{\boldsymbol{\rho}} - m_2\bar{\boldsymbol{r}}/m|^3} - Gm_2\frac{\bar{\boldsymbol{\rho}} + m_1\bar{\boldsymbol{r}}/m}{|\bar{\boldsymbol{\rho}} + m_1\bar{\boldsymbol{r}}/m|^3} \tag{3.119}$$

for the third body; the reason for adorning each vector with an overbar will be revealed presently. In this limit we have $\bar{r}_1 = (m_2/m)\bar{r}$, $\bar{r}_2 = -(m_1/m)\bar{r}$, and $r_3 = \bar{\rho}$. The circular orbit of the two-body system is described by

$$\bar{r} = r \cos(\Omega t)\, \bar{e}_x + r \sin(\Omega t)\, \bar{e}_y, \tag{3.120}$$

with $r = $ constant and $\Omega := \sqrt{Gm/r^3}$ denoting the orbital angular velocity.

The equations are presented in a frame $(\bar{x}, \bar{y}, \bar{z})$ that is attached to the barycenter of the three-body system. This frame is non-rotating, and accordingly, the motion of the two-body system is described by the circular orbit of Eq. (3.120). To proceed it is helpful to switch reference frames, and to choose instead a frame (x, y, z) that is co-rotating with the orbiting bodies. The tools to achieve this were developed near the end of Sec. 2.3.1, and the notation adopted here, with bars placed on the non-rotating coordinates and no decoration placed on rotating coordinates, follows the conventions adopted back in Chapter 2. The coordinate transformation is given by $\bar{x} = x \cos \Omega t - y \sin \Omega t$, $\bar{y} = x \sin \Omega t + y \cos \Omega t$, $\bar{z} = z$, and the two-body system appears stationary in the rotating frame; the separation vector is now given by

$$r = r\, e_x, \tag{3.121}$$

and body 1 is now fixed at coordinates $x_1 = (m_2/m)r$ and $y_1 = 0$, while body 2 is at coordinates $x_2 = -(m_1/m)r$ and $y_2 = 0$.

Simple manipulations reveal that the equations of motion of the test body become

$$\frac{d^2\boldsymbol{\rho}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\boldsymbol{\rho}}{dt} = \boldsymbol{\nabla}\Phi \tag{3.122}$$

in the rotating frame, where $\boldsymbol{\Omega} := \Omega\, e_z$ is the angular-velocity vector, and

$$\Phi = \frac{1}{2}\Omega^2\big[\rho^2 - (\boldsymbol{\rho} \cdot e_z)^2\big] + \frac{Gm_1}{|\boldsymbol{\rho} - m_2 r/m|} + \frac{Gm_2}{|\boldsymbol{\rho} + m_1 r/m|}, \tag{3.123}$$

a generalized potential that includes both gravitational and centrifugal terms; in Eq. (3.122) the gradient operator $\boldsymbol{\nabla}$ refers to the three components of the vector $\boldsymbol{\rho}$ in the rotating frame. The equations admit a first integral, which is obtained by taking the scalar product of Eq. (3.122) with $d\boldsymbol{\rho}/dt$ and recognizing that each term is a total derivative with respect to t. Integration produces the *Jacobi integral*

$$\frac{1}{2}V^2 - \Phi = C, \tag{3.124}$$

where $V := |d\boldsymbol{\rho}/dt|$ is the speed of the test body in the rotating frame, and C is a constant of the motion, a generalized energy known as Jacobi's constant.

Equilibria of the restricted problem

The motions predicted by Eq. (3.122) are extremely diverse, and a complete exploration would require numerical integration of the equations. The question we shall investigate here is whether the equations admit stationary solutions with $d\boldsymbol{\rho}/dt = \mathbf{0} = d^2\boldsymbol{\rho}/dt^2$. If such solutions exist, they would occur at values of $\boldsymbol{\rho}$ such that $\boldsymbol{\nabla}\Phi = \mathbf{0}$, that is, at stationary

points of the generalized potential. A little thought reveals that a stationary solution requires the test body to be in the same plane as the orbiting bodies.

To simplify the analysis we adopt r as a unit of length, and divide all distances by r so that they become dimensionless. Similarly, we adopt $m := m_1 + m_2$ as a unit of mass, and divide all masses by m so that they too become dimensionless. In those units, the first body is situated at $x_1 = m_2$, the second body at $x_2 = -m_1$, and the orbital angular velocity is $\sqrt{G}$. We take the test body to have coordinates $\boldsymbol{\rho} = (x_3, y_3, 0)$, and let

$$r_1 := r_{13} = |\boldsymbol{\rho} - m_2 \boldsymbol{r}/m| = \sqrt{(x_3 - m_2)^2 + y_3^2}, \tag{3.125a}$$

$$r_2 := r_{23} = |\boldsymbol{\rho} + m_1 \boldsymbol{r}/m| = \sqrt{(x_3 + m_1)^2 + y_3^2}. \tag{3.125b}$$

With $m_1 + m_2$ now restricted to be unity, the individual masses are determined by the mass ratio $q := m_1/m_2 \leq 1$; we have that

$$m_1 = \frac{q}{1+q}, \qquad m_2 = \frac{1}{1+q}, \tag{3.126}$$

and we adopt the convention that the first body will be the less massive of the two main bodies.

The generalized potential is given by

$$G^{-1}\Phi = \frac{1}{2}\left(x_3^2 + y_3^2\right) + \frac{m_1}{r_1} + \frac{m_2}{r_2} \tag{3.127}$$

in the rescaled variables, and differentiation with respect to x_3 and y_3 produces the equilibrium conditions $x_3 - m_1(x_3 - m_2)/r_1^3 - m_2(x_3 + m_1)/r_2^3 = 0$ and $y_3 - m_1 y_3/r_1^3 - m_2 y_3/r_2^3 = 0$. These can be re-expressed as

$$0 = m_1(x_3 - m_2)(1 - r_1^{-3}) + m_2(x_3 + m_1)(1 - r_2^{-3}), \tag{3.128a}$$

$$0 = y_3\left[m_1(1 - r_1^{-3}) + m_2(1 - r_2^{-3})\right], \tag{3.128b}$$

by exploiting the identity $m_1 + m_2 = 1$. The existence of stationary solutions hinges on finding solutions to these algebraic equations. As we shall see below, there are five distinct solutions, which are known as the *Lagrange points* L_1, L_2, L_3, L_4, and L_5. The first three equilibria have $y_3 = 0$ and describe co-linear configurations, with all three bodies aligned on the x-axis; these equilibria are unstable. The last two equilibria have $r_1 = r_2 = 1$ and describe triangular configurations, with all three bodies equidistant from each other; these equilibria are stable for sufficiently small mass ratios. The five equilibria are represented in Fig. 3.3.

We first consider the solutions to Eqs. (3.128) with $y_3 = 0$. In this case $r_1 = |x_3 - m_2|$, $r_2 = |x_3 + m_1|$, and we must distinguish three subcases. For L_1 we take the test body to be between the other two bodies, so that $-m_1 < x_3 < m_2$; in this case we have that $r_1 = m_2 - x_3$, $r_2 = x_3 + m_1$, and Eq. (3.128a) can be expressed as

$$q = \frac{r_2 - r_2^{-2}}{r_1 - r_1^{-2}}, \qquad r_1 + r_2 = 1 \qquad (L_1). \tag{3.129}$$
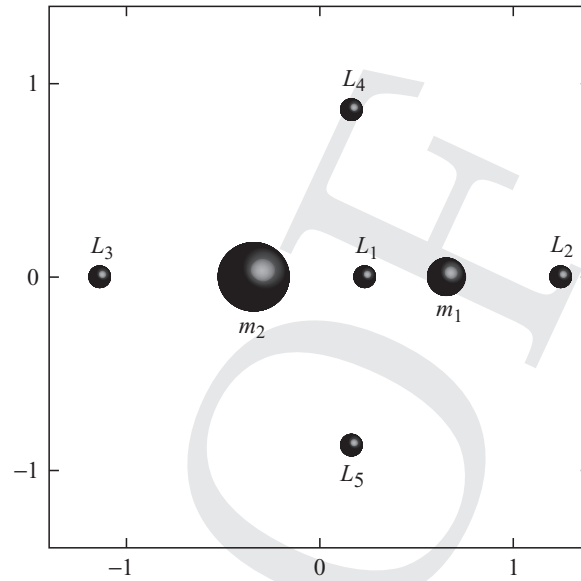
Fig. 3.3 The five Lagrange points of a restricted three-body system with mass ratio $q = \frac{1}{2}$.

It is easy to see that a solution exists for any value of $q \leq 1$, with $r_1 \leq \frac{1}{2}$. For example, with $q = \frac{1}{2}$ so that $m_1 = \frac{1}{3}$ and $m_2 = \frac{2}{3}$, the solution is $x_3 \simeq 0.23742$. For L_2 we take the test body to be beyond body 1, so that $x_3 > m_2$; in this case we have that $r_1 = x_3 - m_2$, $r_2 = x_3 + m_1$, and Eq. (3.128a) can now be expressed as

$$q = -\frac{r_2 - r_2^{-2}}{r_1 - r_1^{-2}}, \qquad r_2 - r_1 = 1 \qquad (L_2). \tag{3.130}$$

A solution exists for any value of $q \leq 1$ with $r_1 \leq 0.7$. For example, with $q = \frac{1}{2}$ the solution is $x_3 \simeq 1.24905$. Finally, for L_3 we take the test body to be beyond body 2, so that $x_3 < -m_1$; in this case we have that $r_1 = m_2 - x_3, r_2 = -(x_3 + m_1)$, and Eq. (3.128a) yields

$$q = -\frac{r_2 - r_2^{-2}}{r_1 - r_1^{-2}}, \qquad r_1 - r_2 = 1 \qquad (L_3). \tag{3.131}$$

A solution exists for any value of $q \leq 1$ with $1.7 \leq r_1 \leq 2$. For example, with $q = \frac{1}{2}$ the solution is $x_3 \simeq -1.13636$.

We next consider the solutions to Eqs. (3.128) with $y_3 \neq 0$. It is easy to see that a simultaneous solution to these equations requires $r_1 = r_2 = 1$, so that the configuration is an equilateral triangle. Inserting this constraint within Eq. (3.125) reveals that there are two distinct solutions described by $x_3 = \frac{1}{2} - m_1 = m_2 - \frac{1}{2}$ and $y_3 = \pm\frac{1}{2}\sqrt{3}$; the solution with $y_3 > 0$ is L_4, and the one with $y_3 < 0$ is L_5. For $q = \frac{1}{2}$, L_4 is situated at $x_3 = 0.16667$ and $y_3 = 0.86603$, while L_5 is situated at $x_3 = 0.16667$ and $y_3 = -0.86603$.

A stability analysis reveals that L_1, L_2, and L_3 are unstable equilibria: small displacements from these points produce motions that take the test body far away. The remaining Lagrange points, however, are stable provided that (refer to Exercise 3.12)

$$q < \frac{1 - \frac{1}{9}\sqrt{69}}{1 + \frac{1}{9}\sqrt{69}} \simeq 0.04; \tag{3.132}$$

in this case a small displacement produces a small oscillation around the equilibrium point.

In the solar system, the stable Lagrange points of the Sun–Jupiter system are occupied by large collections of asteroids. These orbit the Sun along Jupiter's orbital path, with one group, called the "Greek" asteroids, $60°$ ahead of it (at L_4), and the other group, called the "Trojan" asteroids, $60°$ behind it (at L_5). The asteroids remain in those positions because they are less disturbed by planetary perturbations, being protected by the weak potential well provided by the gravitational fields of the Sun and Jupiter combined with centrifugal forces. Similarly, there is a distribution of interplanetary dust at the Earth–Sun Lagrange points, and two small moons lie at the stable Lagrange points of the system formed by Saturn and Tethys, one of its moons.

It turns out that the L_1 and L_2 points of the Earth–Sun system are very useful places to park a spacecraft. Even though these equilibria are unstable, the instability is weak, and a relatively small amount of thrust is sufficient to keep the spacecraft in a small orbit around the Lagrange point. For example, the Solar and Heliospheric Observatory (SOHO) was placed at L_1, pointing continuously sunward. The Wilkinson Microwave Anisotropy Probe (WMAP) and the Planck satellite were both placed at L_2, so that they could point away from the Sun and Earth. The James Webb Space Telescope (the replacement for Hubble) is also destined for L_2.

3.5.2 The N-body problem

Very little can be said about the detailed motion of an N-body system, even when the equation of motion of each body is given by the simplest expression

$$\boldsymbol{a}_A = -\sum_{B \neq A} G m_B \frac{\boldsymbol{r}_{AB}}{r_{AB}^3}. \tag{3.133}$$

General statements can still be made, but they concern only the most global and coarse-grained aspects of the motion. For example, we know from Sec. 1.6.6 that the system's total momentum, energy, and angular momentum are conserved quantities. These, we recall, are given by $\boldsymbol{P} = \sum_A m_A \boldsymbol{v}_A$, $E = \mathcal{T} + \Omega$, and $\boldsymbol{L} = \sum_A m_A \boldsymbol{r}_A \times \boldsymbol{v}_A$, with

$$\mathcal{T} = \frac{1}{2} \sum_A m_A v_A^2 \tag{3.134}$$

denoting the total kinetic energy, and

$$\Omega = -\frac{1}{2} \sum_A \sum_{B \neq A} \frac{G m_A m_B}{r_{AB}} \tag{3.135}$$

the total gravitational potential energy. Another global statement is the virial theorem

$$\frac{1}{2}\frac{d^2 I}{dt^2} = 2\mathcal{T} + \Omega, \tag{3.136}$$

where $I = \sum_A m_A |\boldsymbol{r}_A|^2$ is the quadrupole-moment scalar of the N-body system. This version of the virial theorem is a specialization to isolated bodies of the general theorem of Sec. 1.4.3, which applies to any fluid configuration. Though they are limited in information content, the global statements are nevertheless useful because they rest upon the most fundamental aspects of Newtonian dynamics. They can, actually, be very powerful at probing the nature of complicated systems. For example, the virial theorem played an important role in deducing the existence of dark matter in the universe.

The virial theorem applies when an N-body system such as a star cluster can be regarded as being isolated and gravitationally bound, and it is useful when the system can also be regarded as being in a steady state, so that $\ddot{I} \approx 0$. In such situations the theorem implies that $\mathcal{T} \approx \frac{1}{2}|\Omega|$. The kinetic energy of the cluster can be expressed as $\frac{1}{2}M\langle v^2 \rangle$, where M is the total mass of the cluster, and $\langle v^2 \rangle$ is a suitable average of v^2 over all stars in the cluster. On the other hand, the gravitational potential energy is approximately equal to $-GM^2/(2R)$, where R is the cluster's radius. If the distribution of velocities is assumed to be isotropic, then $\langle v^2 \rangle \approx 3\langle v_r^2 \rangle$, with v_r denoting the component of the stellar velocity along the line of sight, which can be measured from the Doppler shift of the spectral lines (after removing the overall shift caused by the bulk motion of the cluster relative to the Earth). The radius R is determined from the cluster's angular size combined with an estimate of its distance. With these ingredients, the "virial mass" of the cluster is given by $GM_{\text{virial}} \simeq 6R\langle v_r^2 \rangle$, and within a factor of order unity, we should expect this to be a reliable estimate of the cluster's total mass. The problem is that the virial mass determined in this way turns out to be as much as 10 times larger than the "visible mass" determined by measuring the brightness of the cluster and multiplying by the mass-to-light ratio known from stellar-structure theory for the kinds of stars contained in the cluster. When the virial method was applied to clusters of galaxies (with each galaxy viewed as a point body), the discrepancy between the virial and visible masses was even larger.

One proposed explanation for the discrepancy held that star clusters are really not bound at all, so that the virial theorem does not apply to them. Another possible explanation is that the clusters are not in a steady state, so that the assumption $\ddot{I} \approx 0$ is not valid. But observations of many clusters and detailed analyses have suggested that these avenues are not promising. An alternative explanation is that there is additional mass in the system that contributes to the gravitational binding, but does not emit any light – this is the dark matter hypothesis. Eventually it was shown that the amount of dark matter implied by the virial-mass discrepancy in star clusters and galaxies is consistent with the amount required to explain the rotational properties of many spiral galaxies, the angular fluctuations in the cosmic microwave background radiation, and the growth of large-scale structure in the early universe. The dark matter hypothesis has now become a central feature of the standard cosmological model.

3.6 Lagrangian formulation of Newtonian dynamics

Our exploration of Newtonian gravity is coming to a close, but we would be remiss if we didn't touch upon the wonderful developments initiated by Maupertuis and Euler and completed by Lagrange and Hamilton. Their *principle of least action* has become a unifying principle for all of physics, being applicable to such disparate fields as mechanics, fluid dynamics, and field theory including general relativity. Our treatment here will be extremely brief; we expect the reader to have encountered this topic in greater depth elsewhere. We include it regardless because we shall have occasion to refer to Lagrangians and action functionals in later portions of the book, and wish to provide a quick reference manual to accompany these references.

3.6.1 Lagrangian and action principle

We consider a dynamical system with an arbitrary number of degrees of freedom. We describe the motion of the system with *generalized coordinates* $q^j(t)$, which may be chosen arbitrarily to give the simplest description; the generalized coordinates give rise to the *generalized velocities* $\dot{q}^j(t) := dq^j/dt$. The system possesses a kinetic energy $\mathcal{T}$ and a potential energy Ω which depend upon the generalized coordinates and velocities. The system's *Lagrangian function* is defined to be $L(\dot{q}^j, q^j) := \mathcal{T} - \Omega$. The system's *action functional* is

$$S[q] := \int_{t_1}^{t_2} L(\dot{q}^j, q^j)\, dt, \tag{3.137}$$

the integral of the Lagrangian function between two fixed times. The action can be evaluated for any path $q^j(t)$ that joins an initial configuration $q^j(t_1)$ to a final configuration $q^j(t_2)$, whether or not this path satisfies the equations of motion of the dynamical system.

The true path of the dynamical system – the one that satisfies the equations of motion – is identified as the one that produces an *extremum* of the action functional. This means the following. Suppose that the true path is $\bar{q}^j(t)$, and that we examine an arbitrary variation of this path described by the small displacements $\delta q^j(t)$, thereby constructing a trial path $q_{\text{trial}}^j(t) = \bar{q}^j(t) + \delta q^j(t)$. We insist that the trial path should begin from the same configuration as the true path, and that it should also end in the same configuration as the true path, so that the displacements must obey $\delta q^j(t_1) = 0 = \delta q^j(t_2)$; apart from this the functions $\delta q^j(t)$ are arbitrary. The action functional can be evaluated on the true path, and it can be evaluated on the trial path; it will be an extremum when the difference $\delta S := S[q_{\text{trial}}] - S[\bar{q}]$ vanishes to first order in the displacements. As we shall see, the condition $\delta S = 0$ gives us a means to identify $\bar{q}^j(t)$. It can be shown that when $t_2 - t_1$ is not too large, the extremum is in fact a minimum – the action is minimized when evaluated along the true path. When $t_2 - t_1$ increases beyond a certain threshold (called a kinetic focus), the extremum turns into a saddle point of the action functional.

To carry out the variation of the action we write

$$\delta S = S[q_{\text{trial}}] - S[\bar{q}]$$
$$= \int_{t_1}^{t_2} \left[L\left(\dot{q}_{\text{trial}}^j, q_{\text{trial}}^j\right) - L\left(\dot{\bar{q}}^j, \bar{q}^j\right) \right] dt \tag{3.138}$$

and express the trial Lagrangian as a Taylor expansion about the true path, obtaining

$$L\left(\dot{\bar{q}}^j, \bar{q}^j\right) + \frac{\partial L}{\partial \dot{q}^j} \delta \dot{q}^j + \frac{\partial L}{\partial q^j} \delta q^j$$

to first order in the displacements; summation over the repeated index j is understood. The variation of the action becomes

$$\delta S = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial \dot{q}^j} \frac{d}{dt} \delta q^j + \frac{\partial L}{\partial q^j} \delta q^j \right) dt, \tag{3.139}$$

and integration by parts of the first term produces

$$\delta S = \frac{\partial L}{\partial \dot{q}^j} \delta q^j \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(-\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^j} + \frac{\partial L}{\partial q^j} \right) \delta q^j \, dt. \tag{3.140}$$

The boundary terms vanish because of the conditions $\delta q^j(t_1) = 0 = \delta q^j(t_2)$, and since the variations $\delta q^j(t)$ are otherwise arbitrary, we find that $\delta S = 0$ requires

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^j} - \frac{\partial L}{\partial q^j} = 0. \tag{3.141}$$

These are the *Euler–Lagrange equations* for the dynamical system, which give rise to second-order differential equations for the generalized coordinates $q^j(t)$. The Euler–Lagrange equations, therefore, provide the equations of motion to be satisfied by the dynamical system, as dictated by the principle of extremum action.

The Euler–Lagrange equations can be written in the form

$$\dot{p}^j = \frac{\partial L}{\partial q^j}, \tag{3.142}$$

in which $p^j := \partial L / \partial \dot{q}^j$ is the *generalized momentum* associated with the coordinate q^j. An immediate consequence of this equation is that when L does not depend explicitly on one (or more) of its generalized coordinates, say q^*, then the associated momentum p^* is necessarily a constant of the motion. In this way, momentum conservation is seen to arise as a consequence of a symmetry of the dynamical system.

3.6.2 Lagrangian mechanics of a two-body system

We apply these ideas to a system of two spherical bodies moving under their mutual gravitational attraction. We follow the description of Sec. 1.6.7 and adopt the variables $\boldsymbol{R} := (m_1/m)\boldsymbol{r}_1 + (m_2/m)\boldsymbol{r}_2$ and $\boldsymbol{r} := \boldsymbol{r}_1 - \boldsymbol{r}_2$ as generalized coordinates; here $m := m_1 + m_2$ is the total mass of the system, and $\mu := m_1 m_2/m$ is the reduced mass. The corresponding generalized velocities are $\boldsymbol{V} := d\boldsymbol{R}/dt$ and $\boldsymbol{v} = d\boldsymbol{r}/dt$. The kinetic energy of

the system is $\mathcal{T} = \frac{1}{2}mV^2 + \frac{1}{2}\mu v^2$, and its gravitational potential energy is $\Omega = -G\mu m/r$. The Lagrangian of the two-body system is therefore

$$L = \frac{1}{2}mV^2 + \frac{1}{2}\mu v^2 + \frac{G\mu m}{r}. \tag{3.143}$$

We see immediately that L does not depend explicitly on the coordinates $\boldsymbol{R}$, and this implies that the corresponding momentum $\boldsymbol{P} := \partial L/\partial \boldsymbol{V} = m\boldsymbol{V}$ is constant; we have recovered the fact that the total momentum of a two-body system is a conserved quantity. Taking this into account, we recognize that the first term in the Lagrangian is an irrelevant constant which plays no role in the dynamics of the relative orbit. We may remove it, and take the effective Lagrangian to be

$$L = \frac{1}{2}\mu v^2 + \frac{G\mu m}{r}. \tag{3.144}$$

The reduction to an effective one-body problem has been achieved at the level of the Lagrangian, instead of at the level of the equations of motion.

The gravitational potential energy Ω is spherically symmetric, and this motivates the use of spherical polar coordinates (r, θ, ϕ) instead of the original Cartesian coordinates (x, y, z) associated with the separation vector $\boldsymbol{r}$. The transformation, given by $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, and $z = r\cos\theta$, can implemented directly within the Lagrangian, which becomes

$$L = \frac{1}{2}\mu\big(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2\big) + \frac{G\mu m}{r}. \tag{3.145}$$

Another symmetry is revealed: the Lagrangian does not depend explicitly on ϕ, and the generalized momentum $p_\phi = \partial L/\partial_\phi = \mu r^2\sin^2\theta\,\dot{\phi}$ is a constant of the motion. Because ϕ is an angular coordinate, p_ϕ has the interpretation of an angular momentum, and it is in fact the component of the angular-momentum vector $\boldsymbol{L} = \mu\boldsymbol{r}\times\boldsymbol{v}$ in the z-direction. Because this direction was selected arbitrarily in the definition of the spherical coordinates, we can conclude that in fact, all components of the angular-momentum vector are conserved. This statement implies that the motion takes place in a fixed orbital plane, and this plane can be identified with the equatorial plane $\theta = \frac{\pi}{2}$ without loss of generality.

Setting $\theta = \frac{\pi}{2}$ and $\dot{\theta} = 0$ in the Lagrangian, we finally obtain

$$L = \frac{1}{2}\mu\big(\dot{r}^2 + r^2\dot{\phi}^2\big) + \frac{G\mu m}{r} \tag{3.146}$$

for the fully reduced Lagrangian that governs the planar motion of a two-body system. At this stage it is easy to show that the Euler–Lagrange equations produce the orbital equations first obtained in Sec. 3.2.3.

3.6.3 Lagrangian mechanics of a test mass

To conclude this brief tour of Lagrangian mechanics, we examine the dynamics of a test body in the gravitational field of other massive bodies. In this case the dynamical system consists only of the test body, which is given a mass m, position $\boldsymbol{r}(t)$, and velocity $\boldsymbol{v}(t)$. The gravity of the external bodies is described by their gravitational potential $U(t, \boldsymbol{x})$, which

is supposed to be known as a function of space and time. The particle's kinetic energy is $\frac{1}{2}mv^2$, and its potential energy – its interaction energy with the remaining bodies – is $-mU(\mathbf{r})$, in which the potential is evaluated at $\mathbf{x} = \mathbf{r}(t)$. The particle's Lagrangian is

$$L = \frac{1}{2}mv^2 + mU(\mathbf{r}).$$ (3.147)

In this case the Lagrangian possesses whatever symmetry is displayed by the gravitational potential U; in general there will be no particular symmetry.

The generalized momentum associated with $\mathbf{r}$ is $\mathbf{p} = \partial L/\partial \mathbf{v} = m\mathbf{v}$, which is recognized as the usual momentum vector. The Euler–Lagrange equations read

$$\frac{d\mathbf{p}}{dt} = m\nabla U(\mathbf{r}),$$ (3.148)

in which the potential is differentiated with respect to $\mathbf{r}$. The left-hand side is $m\mathbf{a} = md^2\mathbf{r}/dt^2$, and the right-hand side is recognized as the gravitational force exerted by the external bodies. We have recovered the familiar Newtonian equations of motion, $m\mathbf{a} = \mathbf{F}$ with $\mathbf{F} = m\nabla U$.

Thus ends our survey of Newtonian gravity. Now onward with relativity!

3.7 Bibliographical notes

Kepler's problem (Sec. 3.2) is a staple of Newtonian mechanics, and its solution is presented in most textbooks devoted to this topic, including French (1971). More complete presentations, including the three-dimensional aspects of Keplerian orbits (as reviewed in Sec. 3.2.5), are typically found in books on celestial mechanics; among our favorites are the old classic by Brouwer and Clemence (1961) and the more modern text by Murray and Dermott (2000).

The method of osculating orbital elements developed in Sec. 3.3 is also standard fare of books on celestial mechanics; these also contain many more applications of the formalism than the few presented in Sec. 3.4. The Kozai mechanism was discovered independently by Lidov (1962) and Kozai (1962). The first measurement of the solar oblateness by Dicke and Goldenberg is presented in their 1967 article. An accessible introduction to the rich field of helioseismology can be found in Narayanan (2012). The motion of the Moon is treated in great detail in Brown (1960). A broad overview of the gravitational N-body problem (Sec. 3.5), with applications to the solar system, stellar and galactic dynamics, and cosmology, is contained in the compilation of lectures edited by Steves and Maciejewski (2001).

Our presentation of Lagrangian mechanics in Sec. 3.6 does not do justice to this most elegant formulation of the laws of mechanics, and we add insult to injury by not even mentioning the Hamiltonian and Hamilton–Jacobi versions of the theory. These topics, fortunately, are thoroughly treated in many textbooks. A very terse but elegant presentation can be found in the classic *Mechanics* by Landau and Lifshitz (1976), and a more leisurely treatment in Goldstein, Poole, and Safko (2001).

             3.8 Exercises

3.8 Exercises

3.1 In this problem we examine the conic sections of Kepler's problem.

(a) For an orbit with $e \neq 1$ and $\omega = 0$, consider a coordinate transformation to new variables $\bar{x}$ and $\bar{y}$, of the form $\bar{x} = r \cos \phi + p\alpha$, $\bar{y} = r \sin \phi$, where $r = p/(1 + e \cos \phi)$ and α is a constant. Find the value of α that converts the curve to an ellipse or hyperbola, described by the equation

$$\frac{\bar{x}^2}{A^2} \pm \frac{\bar{y}^2}{B^2} = 1,$$

where the positive and negative signs correspond to an ellipse and hyperbola, respectively. Determine A and B in terms of the semi-latus rectum p and eccentricity e.

(b) For $e = 1$, find the transformation that converts the curve into a parabola described by $\bar{x} = C\bar{y}^2$. Determine C in terms of p.

3.2 Suppose that the solar system is filled with a uniform distribution of dark matter with constant mass density ρ. Taking this distribution into account, calculate the modified gravitational potential of the Sun, and find the perturbing force f acting on a planetary orbit. Find the relation between orbital period P and semi-major axis a for a circular orbit, and calculate the secular changes in the planet's orbital elements. Place a bound on ρ using suitable solar-system data.

3.3 In some relativistic theories of gravity, the "graviton" is not massless as in general relativity, but possesses a mass m_g. In the Newtonian limit, the graviton mass gives rise to a modified Poisson equation of the form

$$\left(\nabla^2 + \frac{1}{\lambda^2} \right) U = -4\pi G \rho,$$

in which $\lambda = h/(m_g c)$ is the Compton wavelength of the graviton. Show that the spherically symmetric potential of a body of mass m is given by $U = (Gm/r)e^{-r/\lambda}$. Applying this to the Sun as in the preceding problem, place a bound on λ using solar-system data.

3.4 A test body of mass μ orbits a body of mass m, radius R, and dimensionless quadrupole moment J_2 relative to a symmetry axis e; all other J_ℓs are assumed to vanish. Prove that the following quantities are constants of the orbital motion:

(a) The total energy, given by

$$E = \frac{1}{2}\mu v^2 - \frac{G\mu m}{r} + \frac{1}{2}\frac{G\mu m J_2 R^2}{r^3}\left[3(\boldsymbol{n} \cdot \boldsymbol{e})^2 - 1 \right].$$

(b) The angular momentum along e, given by $L_e = \mu \boldsymbol{h} \cdot \boldsymbol{e}$, where $\boldsymbol{h} := \boldsymbol{r} \times \boldsymbol{v}$.

(c) A third quantity, constant to first order in J_2, given by

$$C = h^2 + J_2 R^2 \left[(\boldsymbol{e} \cdot \boldsymbol{v})^2 - 2\frac{Gm}{r}(\boldsymbol{e} \cdot \boldsymbol{n})^2 \right],$$

where $n := r/r$. This third constant is analogous to the "Carter constant" in the Kerr geometry of a rotating black hole.

3.5 Consider a spherical body on an inclined, circular orbit about an axisymmetric body of radius R and even multipole moments J_ℓ, with $\ell = 2, 4, 6$, and so on. To first order in perturbation theory, calculate the secular changes in the relevant orbital elements. In particular, show that:

(a) the inclination is constant, that is, $\Delta\iota = 0$;

(b) the line of nodes changes by an amount

$$\Delta\Omega = -3\pi \cos\iota \sum_{\ell=2}^{\infty} J_\ell \left(\frac{R}{p}\right)^\ell C_\ell,$$

where $C_2 = 1$, $C_4 = -\frac{5}{2}(1 - \frac{7}{4}\sin^2\iota)$, and $C_6 = \frac{35}{8}(1 - \frac{9}{2}\sin^2\iota + \frac{33}{8}\sin^4\iota)$.

3.6 The equations that govern the evolution of the osculating orbital elements have singularities when $e = 0$ and $\iota = 0$.

(a) Show that the bad behavior of the equations when $e = 0$ can be cured by implementing the transformation $\alpha = e\cos\omega$ and $\beta = e\sin\omega$. Obtain the osculating equations for the new elements α and β.

(b) Show that the bad behavior of the equations when $\iota = 0$ can be cured by implementing the transformation $\mu = \tan\iota\cos\Omega$ and $\nu = \tan\iota\sin\Omega$. Obtain the osculating equations for the new elements μ and ν.

3.7 Show that the period of oscillation of the Kozai resonance is given by Eq. (3.89), and that the orbital elements ω, ι, and e vary according to

$$\omega = \frac{\pi}{2} + A\sin(2\pi t/P_{\text{Kozai}}),$$

$$\iota = \iota_c - \frac{1}{2}Ae_c\cos(2\pi t/P_{\text{Kozai}}),$$

$$e = e_c + \frac{5}{6}A\sin\iota_c\cos\iota_c\cos(2\pi t/P_{\text{Kozai}}),$$

on a secular time scale; here A is an arbitrary amplitude for the oscillations.

3.8 From the equations for Δe, $\Delta\omega$, and $\Delta\iota$ in the Kozai mechanism, show that

$$e^2\cos^2\omega\sin^2\iota - \frac{3}{5}e^2 + \cos^2\iota = \text{constant}.$$

3.9 (a) Consider a two-body system on a circular orbit with separation r and angular velocity $\Omega = (Gm/r^3)^{1/2}$, where $m = m_1 + m_2$ is the total mass. Body 2 may be treated as a point mass, but body 1 is tidally distorted by body 2. Because of tidal dissipation in body 1, r changes over one orbit by

$$\Delta r = -24\pi(k_2)_1 \frac{m_2}{m_1} \frac{R_1^5}{r^4}(\Omega - \omega_1)\tau_1,$$

where R_1, ω_1, $(k_2)_1$, and τ_1 are the radius, rotational angular velocity, Love number, and viscous delay of body 1, respectively – refer to Eq. (3.105). Find the

rate of change of orbital angular momentum, $L = \mu r^2 \Omega$. Comparing your result with Eq. (2.287), what can you conclude about the total angular momentum of the system, $J = L + S_1$?

(b) The Earth–Moon distance is known to be increasing at a rate of about 3.8 cm/yr. Assuming a value $k_2 \approx 0.15$ for the Love number of the Earth, estimate the angle between the Earth's tidal bulge and the Earth–Moon direction. Does the bulge lie ahead of or behind that line?

(b) Show that when the Earth eventually becomes tidally locked with the Moon, the length of the day (and of the month) will be about 48 current days. You may ignore the rotational angular momentum of the Moon, and you may assume that to a good approximation, the moment of inertia of the Earth about its rotation axis is given by $0.33 M_\oplus R_\oplus^2$.

3.10 Consider a point at a position $\bar{\boldsymbol{r}}(t)$ on a circular orbit of radius r around a central body of mass m, orbiting with angular velocity Ω, with $\Omega^2 = Gm/r^3$. Consider also a test body moving on nearby orbit, at a position $\delta\bar{\boldsymbol{r}}(t)$ relative to the point on the circular orbit. Assume that $\delta r \ll r$.

(a) In a coordinate system that rotates around the central body with angular velocity Ω, show that the equations of motion of the test body are given by

$$\frac{d^2}{dt^2}\delta\boldsymbol{r} + 2\boldsymbol{\Omega} \times \frac{d}{dt}\delta\boldsymbol{r} = \Omega^2\Big[3(\boldsymbol{n}\cdot\delta\boldsymbol{r})\boldsymbol{n} - (\boldsymbol{e}_z\cdot\delta\boldsymbol{r})\boldsymbol{e}_z\Big]$$

to first order in $\delta\boldsymbol{r}$; here $\boldsymbol{\Omega} := \Omega\boldsymbol{e}_z$ is the angular-velocity vector, and $\boldsymbol{n} := \boldsymbol{r}/r$.

(b) Prove that the general solution to the equations of motion takes the form of the linear superposition $\delta\boldsymbol{r} = c_1\delta\boldsymbol{r}_1 + c_2\delta\boldsymbol{r}_2 + c_3\delta\boldsymbol{r}_3 + c_4\delta\boldsymbol{r}_4$, where c_n are arbitrary constants, and

$$\delta\boldsymbol{r}_1 = \cos(\Omega t - \chi_1)\,\boldsymbol{n} - 2\sin(\Omega t - \chi_1)\,\boldsymbol{\lambda},$$

$$\delta\boldsymbol{r}_2 = \boldsymbol{n} - \frac{3}{2}\Omega t\,\boldsymbol{\lambda},$$

$$\delta\boldsymbol{r}_3 = \boldsymbol{\lambda},$$

$$\delta\boldsymbol{r}_4 = \cos(\Omega t - \chi_4)\,\boldsymbol{e}_z,$$

are the four eigenmodes of the perturbed orbit; χ_n are arbitrary phases.

(c) Describe the motion that corresponds to each mode, and show that each mode is generated by a perturbation in the orbital elements (p, e, ι, Ω) relative to the unperturbed, circular orbit. Relate the constants c_n to the variations of the orbital elements.

(d) Find a solution with $c_2 = c_3 = 0$, but with $c_1 \neq 0$ and $c_4 \neq 0$, describing a relative orbit that is circular, with a constant radius δr. What is the angle between the plane of the relative orbit and that of the original, unperturbed orbit?

(e) Now find a solution describing three satellites that are moving on the same circular relative orbit, such that initially they are placed at the vertices of an equilateral triangle. Show that as each satellite follows its orbit, the constellation maintains the shape of an equilateral triangle. This configuration was adopted for the three

satellites making up the Laser Interferometer Space Antenna (LISA), a proposed space-based gravitational-wave detector.

3.11 In this problem we explore further the general three-body problem introduced in Sec. 3.5.1.

(a) Show that there exist planar solutions for which $r_{12} = r_{13} = r_{23}$.

(b) Show that the general bound solution for $r := r_{AB}$ is given by

$$r = p/(1 + e \cos \phi),$$

$$r^2 \frac{d\phi}{dt} = (GMp)^{1/2},$$

$$\boldsymbol{n}_{AB} = \cos(\phi - \psi_{AB})\, \boldsymbol{e}_X + \sin(\phi - \psi_{AB})\, \boldsymbol{e}_Y,$$

with an orbital period given by $P = 2\pi a^{3/2}/(GM)^{1/2}$, where $a := p/(1 - e^2)$ and $M := m_1 + m_2 + m_3$, and where the three phases ψ_{AB} differ by $\frac{\pi}{3}$.

(c) Show that a_A, the semi-major axes of the individual orbits, are given by

$$a_1 = \frac{(m_2^2 + m_2 m_3 + m_3^2)^{1/2}}{M}\, a,$$

with the others obtained by suitable permutations.

3.12 Consider a small displacement $\delta\boldsymbol{\rho}$ about the Lagrange point L_4 or L_5 of the restricted three body problem. Expand $\nabla\Phi$ to first order in the displacement, and include the Coriolis term in the equations of motion. Substitute $\delta\boldsymbol{\rho} = \delta\boldsymbol{\rho}_0 e^{-ipt}$ and show that the solutions for the frequency p are real, and thus that the Lagrange points are stable, if and only if the mass ratio q satisfies the criterion of Eq. (3.132).

3.13 Assuming that $q := m_1/m_2 \ll 1$, show that the unstable Lagrange points L_1 and L_2 lie on either side of body 1, at a distance d given by $d \simeq r(q/3)^{1/3}$, where r is the distance between bodies 1 and 2. Calculate d for the Earth–Sun system.