Chapter 4

Planetary motion and few-body problems

In projectile motion discussed in Chapter 3, gravity is considered to be constant. This is an approximation, valid only for distances small compared to the size of Earth. Now imagine we stood atop Mt. Everest and fired a cannon horizontally. At low speed, the cannonball would travel just like projectile motion. With greater firing power and ever increasing speed, the cannonball would travel further and curve less. At some critical speed, it would fly all the way around Earth and come back from behind (ignoring drag, of course). This is exactly how the Moon moves around Earth, or the planets around the Sun. In a sense, planetary motion is akin to projectile motion in that a planet falls continuously toward the Sun. We can no longer treat gravity as a constant force, however.

The motion of planets, or the heavenly bodies as they were known, is a fascinating subject. Observations of planetary motion led to the discovery of Kepler's laws which, strictly speaking, are valid only for a two-body system (the planet and the Sun). The fact that these empirical laws could be successfully explained by

Newton's theory of gravity confirmed the validity and universality of the classical theory of gravity. However, beyond two-body systems, no analytic solutions are known for the general configuration, even though simple three-body systems are still at the heart of physics at all scales, such as helium (two electrons plus the nucleus), water (one oxygen and two hydrogen atoms), or celestial systems. These systems, part of the so-called few-body problems, prove to be a fertile ground for numerical investigation.

We discuss several interesting cases in celestial mechanics ranging from differential precession of Mercury including correction from Einstein's theory of relativity, to exoplanets and three-body problems.

4.1 Motion of a planet

Planetary motion is by and large governed by Newton's theory of gravity. The force F obeys the inverse-square law

$$F(r) = -\frac{k}{r^2} \text{ (force)}, \quad V(r) = -\frac{k}{r} \text{ (potential)}, \tag{4.1}$$

where r is the radial distance, and the constant k is

$$k = GMm, \ G = 6.67 \times 10^{-11} \text{ Nm}^2/\text{kg}^2. \tag{4.2}$$

Here, M and m are the masses.

From Newton's second law $\vec{a} = \vec{F}/m$, the equations of motion are

$$\frac{d\vec{r}}{dt} = \vec{v}, \quad \frac{d\vec{v}}{dt} = -GM\frac{\vec{r}}{r^3}. \tag{4.3}$$

Note that the mass of the planet, m, is canceled out, and the force components have $1/r^3$ dependence because of the $\vec{r}$ vector in the

numerator. Assuming the motion is confined in the *xy* plane, the vectors have only *x* and *y* components.

To see what a typical motion is like, run the following program.

Program listing 4.1: Planetary motion (`earth.py`)

```
 1    import ode, numpy as np      # get ODE solvers, numpy
      import visual as vp               # get VPython modules for
      animation

 3
      def earth(id, r, v, t):                    # return the eqns of
      motion
 5        if (id == 0): return v          # velocity, dr/dt
          s = vp.mag(r)                   # s = |r⃗|

 7        return −GM*r/(s*s*s)          # accel dv/dt, faster than
      s**3

 9   def go():
          r = np.array([1.017, 0.0])       # initial x,y position for
      earth
11       v = np.array([0.0, 6.179])      # initial vx, vy

13       # draw the scene, planet earth/path, sun/sunlight
          scene = vp.display(title ='Planetary motion',      #
      scene start
15                        background=(.2,.5,1), forward=(0,2,
      −1))
                        planet=  vp.sphere(pos=r,  radius=0.1,
      make_trail=True,
17                       material=vp.materials.earth, up=
      (0,0,1))
                  sun      = vp.sphere(pos=(0,0),  radius=0.2,
      color=vp.color.yellow,
19                       material=vp.materials.emissive)
                  sunlight = vp. local_light (pos=(0,0),
      color=vp.color.yellow) #scn end
21
```

198

```
        t, h = 0.0, 0.001
23      while True:
            vp.rate(200)    # limit animation speed
25          r, v = ode.leapfrog(earth, r, v, t, h)        # integrate
            planet.pos = r                                # move
    planet
27                          if (scene.kb.keys):  scene.kb.getkey(),
    scene.kb.getkey()    #pause

29  GM = 4*np.pi*np.pi    # G*Msun
    go()
```

We should see planet Earth orbiting the Sun as shown in Figure 4.1.
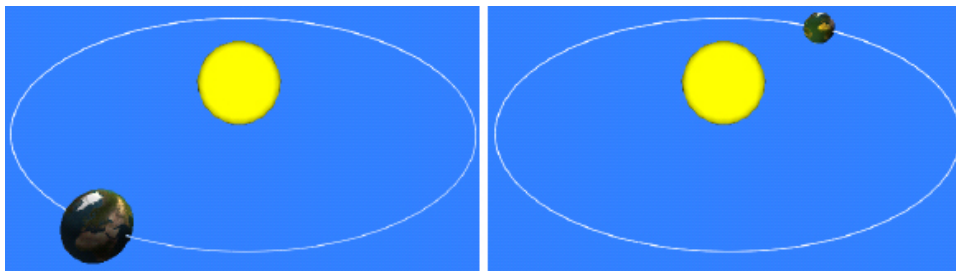

Figure 4.1: Motion of Earth around the Sun.

Do not be too concerned if Program 4.1 seems to have increased complexity in a hurry, for we have yet to explain key elements such as the model or the units in it. By the time we finish Section 4.3, you should understand everything about this program. We put the program upfront because it serves as a central problem around which most problems are built subsequently. We note, however, that the program follows the basic templates and structure of programs we have seen so far, especially Programs 2.7, 3.1, and 3.8. For instance, Program 3.1 first sets up the visualization scene, then goes into the "while" loop for computation. This is basically what the main function go() does here. The difference

199

is that the former uses Euler's method within the loop, while the latter calls the leapfrog method separately. As stated earlier, most programs generally follow this template, with problem specific differences.

In Program 4.1, the equations of motion (4.3) are computed by `earth()` in a form suitable to leapfrog integration (see Chapter 2, Section 2.4).[1] It returns either the velocity or the acceleration depending on the flag `id`, which is requested by the leapfrog integrator. The function `earth()` is similar to `oscillator()` in Program 2.7 but is two-dimensional, so it uses two-element NumPy arrays to represent the position, velocity, and acceleration vectors (lines 10, 11, and 7, respectively). As a result, the acceleration is computed vectorially in identical expression as is written in Eq. (4.3). Besides simplifying the computation, it is also conceptually clearer with vector operations. The advantage of this approach, last used in Program 3.8, will be seen frequently in subsequent problems (see Programs 4.3, 4.5, and 4.7).

In the main code `go()`, the initial position and velocity vectors are given as two-component vectors (lines 10 and 11). These initial values represent the motion of Earth at the point farthest away from the Sun (aphelion, see Table 4.1). The next few lines (14 to 20) add visualization effects much like those in Program 3.1. In order, they set the display window and viewing direction, draw an Earth-like planet via `material.earth` attribute, put a glowing (emissive) Sun at origin, and add a light source representing sunlight (when attributes such as position are given as (x, y) pairs, the z value defaults to zero). Only the line for the "planet" is required, the rest are just bells and whistles, if you will. The `make_trail` attribute for the planet will leave a trail marking the path of VPython objects.

After setting the time and step size, the program enters the main loop, which is functionally similar to the loop in Program 2.7. Because this is a Hamiltonian system, the leapfrog integrator is called to advance the solution in time (line 25). The position of the planet is updated (line 26) after every step, leaving a trail automatically. The program pauses if a key has been pressed (line 27). It then calls `scene.kb.getkey()` twice, first to read the key and then to wait for the second key press before continuing.[2] Finally, we define a constant `GM`$=4\pi^2$ which sets the unit of mass (to be discussed shortly) before executing the main function. Earth should move around the Sun repeatedly over the same path. In other words, the orbit is closed.

I encourage you to play around with Program 4.1, e.g., change the initial position or velocity slightly, turn off (commenting out) the sunlight, disable the trail, or add spin to the planet (see Program 3.8), etc.

4.2 Properties of planetary motion

Since planetary motion falls under the broader scope of two-body central field problems, we will briefly review some helpful properties of motion with central forces for easy reference [40].

4.2.1 REDUCTION TO AN EFFECTIVE ONE-BODY PROBLEM

A general two-body problem can be reduced to an equivalent one-body problem in the center of mass (CM) coordinates. Figure 4.2 illustrates this.

Let $\vec{r}_1$ and $\vec{r}_2$ be the positions of the particles with mass m and M, respectively. We define the coordinates of the CM, $\vec{R}$, and the

relative coordinates, $\vec{r}$, the usual way

$$\vec{R} = \frac{m\vec{r}_1 + M\vec{r}_2}{m + M}, \quad \vec{r} = \vec{r}_1 - \vec{r}_2. \tag{4.4}$$

It can be shown (see Exercise E4.1) that in the CM coordinate system, the equations of motion may be written as

$$\ddot{\vec{R}} = 0, \quad \mu\ddot{\vec{r}} = F\hat{r}, \quad \text{with } \mu = \frac{mM}{m + M}, \tag{4.5}$$

where the force F is given by Eq. (4.1). The center of mass moves at constant velocity (can be set to zero for simplicity), and the relative motion is a one-body problem with the reduced mass, μ. If one mass (say M) is large compared to the other, i.e., $M/m \gg 1$ as in the case of Sun/planet mass ratio, the reduced mass takes on the value of the smaller mass, $\mu = m$. In this case the bigger mass (Sun) can be regarded as infinitely heavy and frozen in space. Only the smaller mass (planet) moves. It is for this reason that the Earth-Sun system in Program 4.1 is effectively a one-body problem. Hereafter, we will use μ and m interchangeably except where confusion might occur.
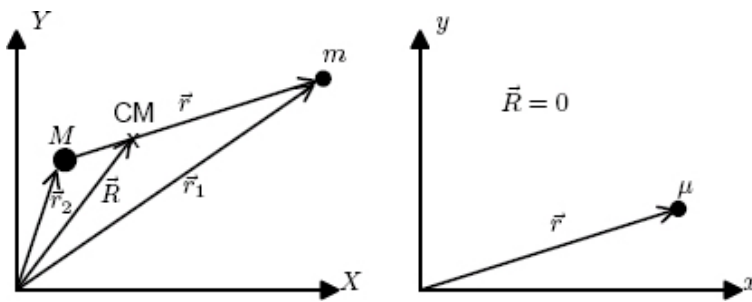

Figure 4.2: The reduction of a two-body problem to a one-body problem in the center of mass coordinates. Left: the space-fixed coordinate system; Right: the center of mass coordinate system. In either case, the relative coordinate $\vec{r} = \vec{r}_1 - \vec{r}_2$ is the same.

4.2.2 ANGULAR MOMENTUM AND EFFECTIVE POTENTIAL

Let $\dot{r}$ and $\dot{\theta}$ be the radial and angular velocities, respectively. The angular momentum $\vec{L} = \vec{r} \times (m\vec{v})$ will be in the z direction as

$$\vec{L} = L\hat{k}, \quad L = mr(r\dot{\theta}) = mr^2\dot{\theta} = \text{ constant.} \qquad (4.6)$$

Since the torque is zero for a central force, $\vec{L}$ is a constant of motion. This is the basis of Kepler's second law which states that the radial vector sweeps an equal area in equal time.

The second constant of motion is energy, of course. From Eq. (4.5), we have

$$E = \frac{1}{2}mv^2 + V(r) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + V(r). \qquad (4.7)$$

We can eliminate $\dot{\theta}$ in E by substituting $\dot{\theta} = L/mr^2$ into Eq. (4.7) as

$$E = \frac{1}{2}m\dot{r}^2 + V_{eff}(r), \quad V_{eff}(r) = \frac{L^2}{2mr^2} + V(r) = \text{ Effective potential.} \qquad (4.8)$$

The term $L^2/2mr^2$ is the centrifugal potential, and it keeps the particle from falling into the origin. The effective potential V_{eff} is the sum of the real potential plus the centrifugal potential. Equation (4.8) describes the motion as if it is one-dimensional along the radial direction only in the potential V_{eff}. For bound motion in an attractive potential, the radius r usually oscillates between $r_{min} \leq r \leq r_{max}$. The values r_{min} and r_{max} are called turning points because at those points, the radial velocity $\dot{r} = 0$.

4.2.3 KEPLER ORBITS

Of course, the actual motion is two-dimensional, say in radius r and angle θ. The orbit is most conveniently expressed as $r(\theta)$. For the inverse-square law (4.1), $r(\theta)$ is

$$r = \frac{L^2/mk}{1 + e\cos(\theta + \theta_0)},\qquad(4.9)$$

$$e = \sqrt{1 + \frac{2EL^2}{mk^2}}.\qquad(4.10)$$

The constant θ_0 determines the orientation of the orbit and can be chosen for convenience. The other constant e is called the eccentricity. Geometrically, Eq. (4.9) is the equation of a conic section (slice of a cone) with the focus at the origin.

The shape of the orbit is determined by the value of e as follows:

$$\begin{array}{llll} E > 0, & e > 1; & \text{hyperbola (unbound),} & \\ E = 0, & e = 1; & \text{parabola (borderline),} & \\ E < 0, & e < 1; & \text{ellipse (bound, Kepler's first law),} & (4.11) \\ E = -mk^2/2L^2, & e = 0; & \text{circle (bound).} & \end{array}$$

The unbound, hyperbolic motion corresponds to scattering of particles and will be discussed later (Chapter 12). For bound, elliptic motion, it is useful to introduce a, the semimajor axis, and rewrite e and L^2/mk as

$$a = -\frac{k}{2E}, \Rightarrow e = \sqrt{1 - \frac{L^2}{mka}}, \Rightarrow \frac{L^2}{mk} = a(1 - e^2).\qquad(4.12)$$

To be more definite, let us choose $\theta_0 = \pi$ in Eq. (4.9), and with the equalities (4.12), Eq. (4.9) becomes

$$r = \frac{a(1 - e^2)}{1 - e\cos\theta}.\qquad(4.13)$$

With this choice of θ_0,[3] it follows that r is bounded by

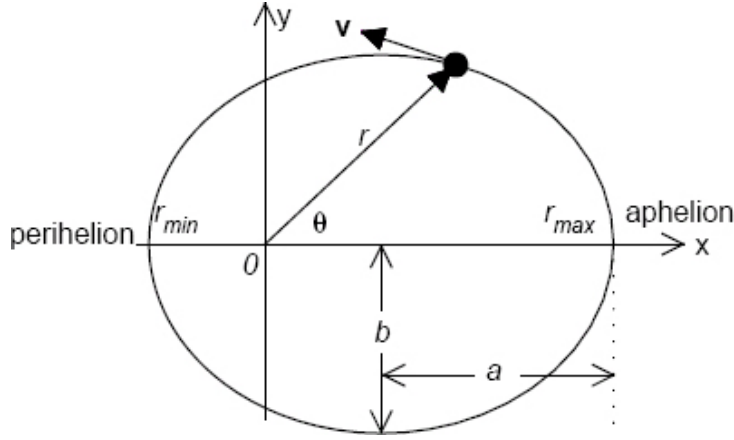$$r_{max} = a(1+e) \text{ (when } \theta = 0), \quad r_{min} = a(1-e) \text{ (when } \theta = \pi). \quad (4.14)$$

Figure 4.3: An elliptic orbit ($e = 0.4$) of a particle moving about the origin O (focus) with semimajor and semiminor axes a and b ($b = a\sqrt{1-e^2}$). The orbit is oriented such that $r = r_{max}$ at $\theta = 0$ (aphelion) and $r = r_{min}$ at $\theta = \pi$ (perihelion). Note $r_{max} + r_{min} = 2a$ by Eq. (4.14).

Figure 4.3 illustrates the parameters of an elliptic orbit. The particle moves around the center of force at the origin, which is also one focus of the ellipse. For planetary motion, the force center is at (very nearly) the Sun. The radial velocity ($\dot{r}$) is zero at the maximum distance (aphelion) and also at the minimum distance (perihelion) from the force center. The velocity at these two positions is perpendicular to the x-axis and may be obtained from energy conservation (4.7) and (4.12) as

$$v_{ap} = \left[\frac{k}{ma} \frac{1-e}{(1+e)} \right]^{1/2} \quad \text{(at aphelion, } r = r_{max}),$$

$$v_{peri} = \left[\frac{k}{ma} \frac{1+e}{(1-e)} \right]^{1/2} \quad \text{(at perihelion, } r = r_{min}). \quad (4.15)$$

We list in Table 4.1 some properties of Kepler's orbits of the planets.[4]

Table 4.1: Properties of Kepler's orbits of the planets.

Planet Name	e	a (AU)	Period (year)	r_{max} (AU)	v_{ap} (AU/year)	Mass $(M_\odot)$
Mercury	0.2056	0.3871	0.2408	0.4667	8.198	1.660×10^{-7}
Venus	0.0068	0.7233	0.6152	0.7282	7.338	2.448×10^{-6}
Earth	0.0167	1.0000	1.0000	1.0167	6.179	3.004×10^{-6}
Mars	0.0934	1.5237	1.8808	1.6660	4.635	3.227×10^{-7}
Jupiter	0.0484	5.2029	11.863	5.4547	2.624	9.551×10^{-4}
Saturn	0.0539	9.5367	29.447	10.051	1.928	2.859×10^{-4}
Uranus	0.0473	19.189	84.017	20.097	1.368	4.355×10^{-5}
Neptune	0.0086	30.070	164.79	30.329	1.136	5.178×10^{-5}
Pluto	0.2488	39.482	247.92	49.305	0.776	6.582×10^{-9}

4.2.4 UNITS FOR PLANETARY MOTION

As in Table 4.1, we will use the astronomical units (AU) for distance, 1 AU=1.496 $\times 10^{11}$ m. This is convenient for planetary motion in our solar system. It is equal to the semimajor axis of Earth's orbit, and represents the average distance between the Earth and the Sun. The natural unit of time is chosen to be one year (365.25 days), or 3.156 $\times 10^{7}$ s. This is the characteristic time scale for planet motion. The unit of speed is 1 AU/year = 4740 m/s.

We now have two of the three basic units. The third one should be about mass. Often we need the ratio $k/m = GM$ with M being the mass of the Sun, as in Eq. (4.3). If we assume circular motion such as the Earth around the Sun, then

$$\mu \frac{v^2}{r} = |F(r)| = \frac{GMm}{r^2}. \qquad (4.16)$$

With a mean radius $r = 1$ AU, the period is 1 year, and the speed is $v = 2\pi$ AU/year (29.77 km/s). Using the reduced mass given by (4.2), we get

206

$$G(M + m) \simeq GM = v^2 r = 4\pi^2 \ \text{AU}^3/\text{year}^2, \qquad (4.17)$$

where we have used the condition that $m/M \ll 1$. In effect, Eq. (4.17) defines the mass of the Sun $M_\odot$ as the unit of mass. We will use $GM = 4\pi^2$ for planetary motion in our units system, as in Program 4.1. In absolute units (rarely needed), $M_\odot = 1.99 \times 10^{30}$ kg, and $GM = 1.327 \times 10^{20}$ m^3/s^2.

Note that Eq. (4.17) is a special case of Kepler's third law

$$T^2 = \frac{4\pi^2}{G(m + M)} a^3 \simeq \frac{4\pi^2}{GM} a^3, \qquad (4.18)$$

which states that the square of period T is proportional to the cube of semimajor axis a. For the purpose of setting the value of GM in our unit system, Eq. (4.18) is equivalent to Eq. (4.17).

Finally, in the unit system thus defined, everything becomes dimensionless. Distance and time are to be compared with 1, and the time step must satisfy $h \ll 1$ in numerical integration, as was the case in Program 4.1.

4.2.5 OPEN AND CLOSED ORBITS

We have seen from running Program 4.1 and Eq. (4.13) that Earth moves in a fixed orbit, retracing itself over the same path periodically (Figure 4.1). The orbit is said to be closed.

Closed orbits may seem ordinary at first, but they are extraordinary. They are the exception rather than the rule. Gravity is one of only two forces supporting them. To wit, slightly modify Program 4.1 by changing the force to $1/r^{2+\delta}$ with $\delta \sim 0.1$, set the initial values for Mercury (so the effect is clearer), run it, and we should see an open orbit as shown in Figure 4.4. The small $\delta = 0.1$ is such that the force deviates from the inverse-square law (4.1)

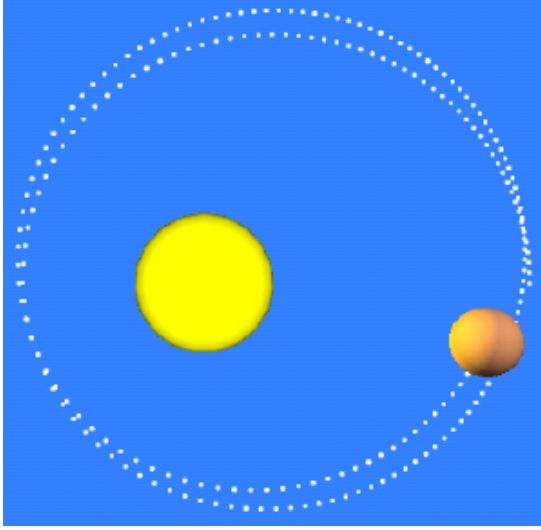only slightly. But significant differences are seen in the orbit after about two revolutions.

Figure 4.4: Open orbit of Mercury with a presumed force $\propto 1/r^{2+\delta}$, $\delta = 0.1$. The initial condition is $[x, y] = [0.4667, 0.0]$, $[v_x, v_y] = [0.0, 8.198]$.

We can understand qualitatively the nature of open and closed orbits by considering the stability of circular motion.[5] Circular motion is possible for any attractive potential. One just has to adjust the energy such that Eq. (4.16) holds. But circular motion may be stable or unstable. If, by increasing the energy slightly from the condition (4.16), the orbit deviates only a small amount from the circular orbit, the motion is said to be stable (with a slightly deformed orbit). Since the full motion consists of radial $r(t)$ and angular $\theta(t)$ motion, we can think of a deformed orbit as composed of radial oscillations and angular sweeps, illustrated in Figure 4.5.

The radial distance oscillates between the turning points $r_{min} \leq r \leq r_{max}$, so the motion is restricted to the region between the concentric circles shown in Figure 4.5 (bottom). If the radius com-

pletes n_r integer oscillations and the angle θ makes n_θ integer sweeps (revolutions) in the same time interval, then the orbit is closed, since n_r/n_θ is rational (commensurate). As it turns out, of all the possible forces in the universe, only $1/r$ (gravitational and Coulombic) and r^2 (harmonic, i.e., Hooke's law) potentials can produce closed orbits.

Other forms of potentials do not support stable circular motion, and an integer pair of (n_r, n_θ) cannot be found, i.e., n_r/n_θ is irrational. The orbit is open, as it does not repeat itself within a finite period of time. Furthermore, the orbit orientation rotates a bit with each revolution. This rotation is known as orbital precession, to be discussed next.
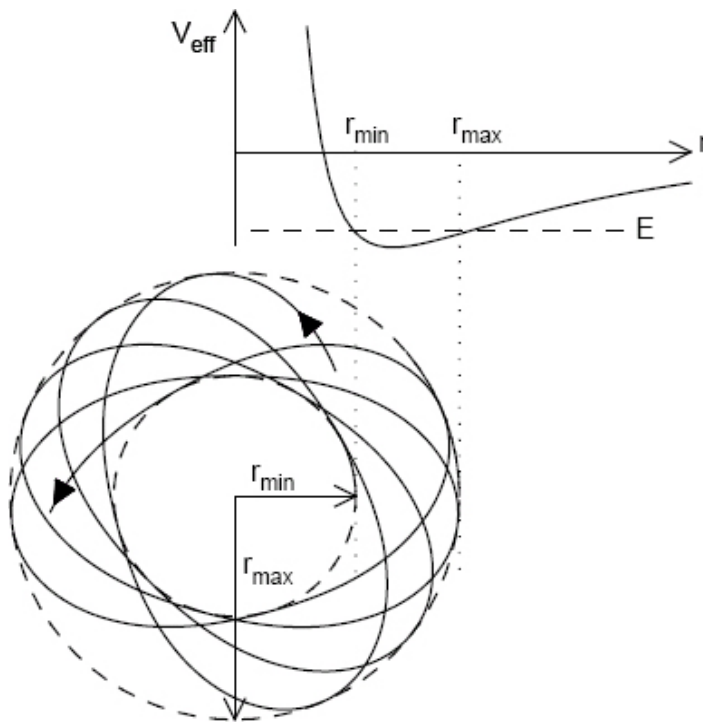

Figure 4.5: An open orbit for bound motion ($E < 0$) in an effective potential.

4.3 Precession of Mercury

Astronomers have long known that the orbits of planets precess with time [25]. Since observations are carried out on Earth which wobbles like a gyroscope (called precession of equinoxes, not our concern here), the data must be analyzed to separate the kinematic effects from true dynamic effects. This is equivalent to going to the fixed space frame where the Sun is at rest. Reduction of measurements of the inner-most planet, Mercury, indicates a precession rate of 574 seconds of arc per century in the space frame (1 arcsecond, or ″, is 1/3600°). In other words, the perihelion (or aphelion, Figure 4.3) of Mercury slowly rotates through an angle of 0.16 degrees per century.

4.3.1 RECONCILIATION OF 43″

From our earlier discussion, we understand that planetary orbits should not be exactly closed because they do not move in a pure $1/r^2$ force field with a single center. Besides the Sun, the other planets pull and tug each other all the time, not to mention their moons [72]. What planet would have the largest effect on the precession of Mercury? It is probably not a surprise that Venus does, because it is closest to Mercury, so its force is largest (see Table 4.2). Farther out radially, the effect decreases except for Jupiter which has the second largest effect because of its mass. Beyond Saturn, the effects are negligible.

The problem is (or was), taking into account all the known perturbations present on Mercury due to these planets (within Newtonian theory of gravity), and after careful number crunching, the precession is expected to be about 531″ for Mercury per century (Table 4.2). That would leave 43″ unaccounted for, or 0.012 degrees/century. This is a very small amount, but well beyond observational uncertainties to be ignored.

As it turns out, the discrepancy can be explained only with Einstein's theory of general relativity. The relativistic correction beautifully accounts for the difference of 43″ exactly [16, 85]. It is one of the major triumphs of Einstein's theory of gravity.

In general relativity, the gravitational field is modified and can be construed as an effective force given by

$$F(r) = -\frac{GMm}{r^2}\left(1 + \frac{\lambda}{r^2}\right), \quad (\lambda = 1.096 \times 10^{-8} \text{ AU}^2 \text{ for Mercury}). \quad (4.19)$$

The details of the ingredients going into producing Eq. (4.19) are rather involved, and beyond our scope [40]. Compared to Eq. (4.1), though, there is an additional $1/r^4$ term in the relativistically modified force. This term destroys the perfect inverse-square law. It causes the orbit to precess.

4.3.2 THE RUNGE-LENZ VECTOR

At first glance, it seems that we can calculate the precession of Mercury due to the relativistic effect by just tracking the rotation of the aphelion (Figure 4.5). While this is possible, it is harder than it sounds. Because the parameter λ in Eq. (4.19) is very small, it is like searching for a needle in a haystack. If we were to use the actual value of λ directly in our simulation, we would have found that, for any moderately practical step size h, the change of the aphelion after one revolution would be insignificant compared to the displacement in one time step. Careful fitting and extrapolation would be needed to find the exact position of the aphelion. Is there a better way? In this case, there is: the Runge-Lenz vector [39, 47].

For observational purposes, of course, it might be natural to want to track the change of the aphelion (or perihelion) in order to

discern any precession. This would be unnecessary in a simulation if we have another way to measure the precession. Fortunately, we do. There is a quantity known as the Runge-Lenz vector that serves our purpose well. The Runge-Lenz vector is defined as [40]

$$\vec{A} = \vec{p} \times \vec{L} - mk\hat{r}, \tag{4.20}$$

where $\vec{p} = m\vec{v}$ is the momentum, $\vec{L}$ the angular momentum as usual, m the mass and k the force constant given in Eq. (4.1).

We know energy E and $\vec{L}$ are constants of motion, a fact very useful to test if our programs are performing correctly (see Project P4.1). As it turns out, $\vec{A}$ is also a constant of motion in a pure inverse-square law of force, i.e., $d\vec{A}/dt = 0$ (see Exercise E4.4). Using vector product rules, we find that $\vec{A} = 0$ for circular orbits. For ellipses, $\vec{A}$ points in the direction from the origin (center of force) to the perihelion, see Figure 4.6.

When the force is not a pure inverse-square in r, as is the case in Eq. (4.19), the Runge-Lenz vector is no longer a constant of motion. Rather it will rotate with the perihelion. This helps us greatly: unlike finding the aphelion, calculating the Runge-Lenz vector involves no fitting. We can now just follow $\vec{A}$ directly in our simulation, and the angle it sweeps through gives the amount of precession. For example, suppose we start our simulation with $\vec{A}$ pointing in the negative x-direction. In time, it will develop a y-component. We can evaluate $\vec{A}$ (both components A_x and A_y) at any time using Eq. (4.20). The angle of precession can be found by $\theta = \tan^{-1}(A_y/A_x)$.
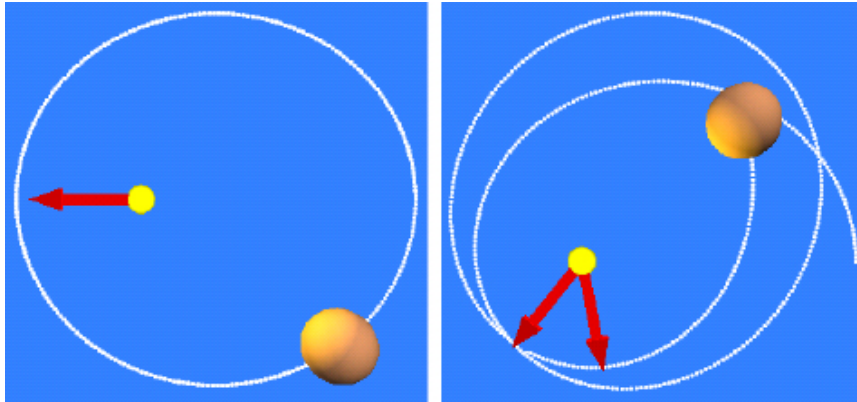
Figure 4.6: The Runge-Lenz vector (arrows). It is constant in a closed orbit in a pure $1/r^2$ force field (left). For other forces where the orbit is open, it rotates with the perihelion (right).

4.3.3 THE LEAPFROG METHOD WITH TIME TRANSFORMATION

If we went ahead and calculated the Runge-Lenz vector with the leapfrog method, we would be disappointed with the results: we would see mostly noise for any meaningfully practical step size h. Again, the problem of the needle-in-a-haystack bites us. Decreasing h is not a preferred solution since it causes more round off error. But there is something more. Because Mercury's relatively large eccentricity (Table 4.1), its orbit is quite elliptic. Close to the perihelion, the force and speed are much larger than the averages. This large variation is what causes the noise. The physical reason, and the insight we gain, is that the force has a singularity at $r = 0$ and the leapfrog integrator somehow "feels" it. In fact, it is possible to induce artificial precession in a pure $1/r^2$ force law if the orbit is elliptic enough and the step size is not small enough. We need to smooth out, or "regularize", the singularity.

An effective way to deal with the problem is by a time transformation. The idea is that the step size is adjusted so that near the closest approach (perihelion), we take smaller

steps. Let us consider the vector equations of motion equivalent to Eqs. (2.39) and (2.40)

$$\frac{d\vec{r}}{dt} = \vec{v},$$ (4.21)

$$\frac{d\vec{v}}{dt} = \vec{a}(\vec{r}).$$ (4.22)

We introduce a time transformation from the real time t to a fictitious time s

$$\frac{ds}{dt} = \Omega(r).$$ (4.23)

Now rewrite Eqs. (4.21) and (4.22) in terms of s using the chain rule $\frac{d}{dt} = \frac{ds}{dt}\frac{d}{ds} = \Omega\frac{d}{ds}$,

$$\frac{d\vec{r}}{ds} = \vec{v}/\Omega(r),$$ (4.24)

$$\frac{d\vec{v}}{ds} = \vec{a}(\vec{r})/\Omega(r).$$ (4.25)

Suppose we choose a function $\Omega(r)$. We would not be able to solve the above equations using the leapfrog method because the position also appears on the RHS of Eq. (4.24), not velocity alone (see Section 2.4). That means the method would not be area preserving.

What we need to do is to treat $\Omega(r)$ as a generalized "velocity" [63]. Let us introduce an auxiliary variable, W, as

$$W = \Omega(r).$$ (4.26)

We will regard W as a "velocity". Using $dW/ds = \nabla\Omega(r) \cdot \vec{v}/\Omega(r)$ and Eq. (4.23), we can add two auxiliary equations to Eqs. (4.24) and (4.25)

214

$$\frac{d\vec{r}}{ds} = \vec{v}/W, \tag{4.27a}$$

$$\frac{dt}{ds} = 1/W, \tag{4.27b}$$

$$\frac{d\vec{v}}{ds} = \vec{a}(\vec{r})/\Omega(r), \tag{4.27c}$$

$$\frac{dW}{ds} = \nabla\Omega(r) \cdot \vec{v}/\Omega(r). \tag{4.27d}$$

The above equations are in the proper form for the leapfrog method if we treat $\vec{r}$ and t as the generalized "coordinates", and $\vec{v}$ and W as the generalized "velocities". Equations (4.27a) and (4.27b) are the exact analogue to Eq. (2.44), and Eqs. (4.27c) and (4.27d) to Eq. (2.45). We are free to choose any form of $\Omega(r)$. A good choice for planetary motion (and other $1/r^n$ forces) is

$$\Omega(r) = \frac{1}{r}. \tag{4.28}$$

It ensures that when r is small, $\Omega(r)$ is large, and dt is small by Eq. (4.23). Furthermore, it is easy to calculate the gradient, $\nabla\Omega(r) = -\vec{r}/r^3$. With Eq. (4.28), we can write the leapfrog algorithm with time transformation analogous to Eqs. (2.41) to (2.43)

$$\vec{r}_{1/2} = \vec{r}_0 + \vec{v}_0 \frac{h}{2W_0}, \tag{4.29a}$$

$$t_{1/2} = t_0 + \frac{h}{2W_0}, \tag{4.29b}$$

$$\vec{v}_1 = \vec{v}_0 + \vec{a}(\vec{r}_{1/2})\, r_{1/2} h, \tag{4.29c}$$

$$W_1 = W_0 - \vec{r}_{1/2} \cdot \frac{\vec{v}_0 + \vec{v}_1}{2} \frac{h}{r_{1/2}^2}, \tag{4.29d}$$

$$\vec{r}_1 = \vec{r}_{1/2} + \vec{v}_1 \frac{h}{2W_1}, \tag{4.29e}$$

$$t_1 = t_{1/2} + \frac{h}{2W_1}. \qquad (4.29\text{f})$$

As usual, the subscripts 0 and 1 denote values at the beginning and the end of a time step h, respectively. However, here the time is the fictitious (transformed) time s, not the actual time t. The actual time is returned in t_1. The difference $t_1 - t_0$ is not constant, giving us the desired effect of smoothing out the singularity near the origin.

Here is the leapfrog method with time transformation, Eqs. (4.29a) to (4.29f).

Program listing 4.2: Leapfrog with time transformation (`leapfrog_tt.py`)

```
1   def leapfrog_tt (lfdiffeq, r0, v0, t0, w0, h):
        """ vectorized leapfrog_tt with time transformation,
3           Omega=1/r, that solves general (r,v) ODEs as:
            dr[i]/dt = f[i](v), and dv[i]/dt = g[i](r).
5           User supplied lfdiffeq (id, r, v, t) returns
            f [i](r) if id=0, or g[i](v) if id=1 """
7       # 1st step: calc r at h/2
        hw = h/(2.0*w0)                          #
    half h/2w0
9       t1 = t0 + hw
        r1 = r0 + hw*lfdiffeq(0, r0, v0, t0)     #
    id=0, get r_1/2
11      r2 = np.dot(r1, r1)                      #
    get r^2=x*x+y*y+z*z
        r12 = np.sqrt(r2)                        #
    r_1/2

13
        # 2nd step: calc v1 using r at h/2
15      v1 = v0 + h*r12*lfdiffeq(1, r1, v0, t1)  #
    id=1 for g(r) at h/2
        rdotv = np.dot(r1, v0+v1)/2.             #
    r . v_1/2
```

216

```
17        w1 = w0 − rdotv*h/r2                          # w₀
```
$- \vec{r} \cdot \vec{v}_{1/2} h/r^2$

```
19        # 3rd step: calc r by another 1/2 step using v1
          hw = h/(2.0*w1)
21        t1 = t1 + hw
          r1 = r1 + hw*lfdiffeq(0, r1, v1, t1)          # get
```
$\vec{r}_1$ *at t+h*
```
23        return r1, v1, t1, w1
```

The module `leapfrog_tt()` uses the same `lfdiffeq` format as required by the standard `leapfrog` (Section 2.4). Except for additional elements of time transformation and calculations of dot products with the NumPy function `np.dot`, it works very much the same way also. We will use the transformed leapfrog integrator for sensitive problems such as the precession of Mercury. As discussed in later chapters, it is also a very useful method for *N*-body systems, including molecular dynamics and atomic reactions (Section S:12.5.1). We include `leapfrog_tt` in the library file `ode.py` with other ODE solvers.

4.3.4 SIMULATION OF DIFFERENTIAL PRECESSION

Precession due to the relativistic correction

Finally we have all the necessary parts to calculate the relativistic correction to the precession of Mercury. Putting it all together, we give the complete simulation in Program 4.3 (Section 4.D).

Run the simulation with animation on, and enter a λ large enough, say 0.01, to see exaggerated effect of precession. It should be similar to Figure 4.6. You should also see that the motion of the Runge-Lenz vector is not uniform: it rocks back and

forth, with the mean motion forward (counterclockwise), in the same direction as the circulation of Mercury.
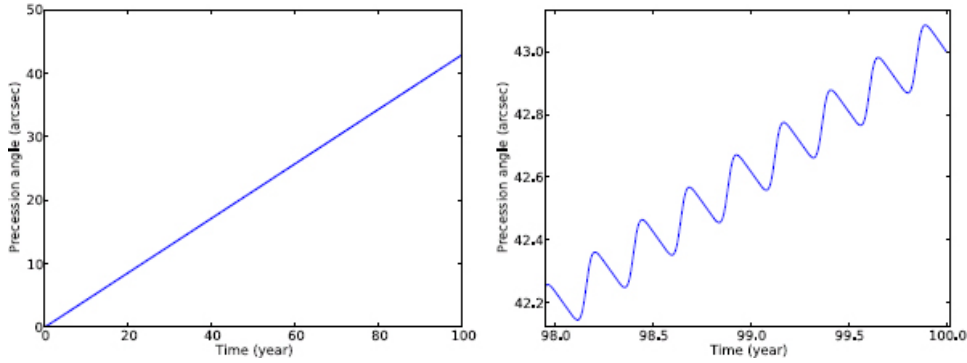

Figure 4.7: Differential precession of Mercury, i.e., the rotation of the Runge-Lenz vector. A zoom-in of the last two years is shown on the right.

Now set `animation=False`, and enter the true value $\lambda = 1.1 \times 10^{-8}$ when prompted. After the simulation ends, we should see a plot showing the angle of precession differential in time over one century in Figure 4.7. It is seen that right at the century mark, the precession is exactly $43''$, a remarkable success of Einstein's theory of gravity. Equally remarkable is the precision of observational astronomy which allows the detection of this tiny amount. Numerically, the leapfrog method with time transformation is sufficiently accurate to pick up this amount, and at the same time, preserves the phase space properties. It is also quite robust. For example, we could increase the step size by a factor of 10 and still get reasonably accurate results.

The full curve (Figure 4.7, left) is not as straight as it looks. It has small-scale oscillations which become visible in the enlarged figure. Because these oscillations are so small (a fraction of an arcsecond), we should ask ourselves, are they real? Could they be artifacts of numerics due to such factors as limited accuracy or incorrect algorithm, for example?

In numerical work, we always have to be mindful of potential pitfalls which could produce artificial effects not present in the actual system. In our case, we can check whether the effect remains for a larger perturbation. As seen from the animation with $\lambda \sim 0.01$ discussed earlier, the back-and-forth rocking motion of the Runge-Lenz vector shows that the oscillations are not restricted to small perturbations. Other checks including conservation laws and step sizes can be performed to see if the oscillations are affected. The answer is no (see Project P4.1). In all likelihood, we can exclude numerical error (noise) as the cause.

Once we are certain that the oscillations are real, what is the physical reason for them? We note that there are about four oscillations per year, or the period is about 1/4 years. This is just the orbital period of Mercury (Table 4.1). Therefore, the oscillations might be connected to the orbital period. With animation turned on and with an artificially large $\lambda \sim 0.01$, we can observe the phenomenon visually. Let us assume that Mercury starts at the aphelion going counterclockwise, with the Runge-Lenz vector $\vec{A}$ pointing at the perihelion ($A_y = 0$). The y-component, A_y, has two terms from Eq. (4.86), $A_y/m^2 = -v_x L - GMy/r$.

Roughly,[6] when v_x turns negative, A_y increases, and when y increases, A_y decreases. Near the aphelion, Mercury is farthest away from the Sun, and the attractive relativistic term causes v_x to turn negative faster than the increase in y (more precisely y/r). The net effect is that A_y increases. This causes the clockwise (retrograde) rotation of $\vec{A}$. Near the perihelion (past the $x = 0$ plane), the opposite occurs, and $\vec{A}$ rotates counterclockwise. Because the perturbation is greater near the perihelion, the counterclockwise rotation is larger than the clockwise rotation, resulting in a net precession in the same direction as Mercury's circulation (ccw,

prograde). It then follows that if the sign of the perturbation is reversed (repulsive), the precession will be in the opposite direction. This can be verified through an exercise.

Central field approximation and precession due to other planets

Mercury interacts with other planets through pair-wise forces that are not central relative to the Sun. A general description would need to take into account all the planets simultaneously (a many-body problem). But if the effects are small, and if we are interested in only averaged values like precession over a period of time, we could approximate each pair-wise force as a central force and treat it individually. The idea is that Mercury moves around the origin (Sun) rapidly and the average force points in the radial direction. In such cases, we speak of central field (or mean field) approximations which are very useful and important in many areas of physics.

If we can construct the force between Mercury and a planet in a central field approximation, we can calculate the precession of Mercury due to that force with the same methodology as the relativistic effect discussed above. One way to obtain the force is as follows. Imagine the mass of a planet, M_p, is distributed uniformly over a coplanar ring of radius a (semimajor axis) [72]. By symmetry, the force between Mercury and the ring is central, i.e., $\vec{F} = F(r)\hat{r}$. The potential is given by an elliptic integral [49] (see Exercise E4.6)

$$V(r) = -\frac{GM_p m}{2\pi a} \int_0^{2\pi} \frac{d\theta}{\sqrt{1 + \rho^2 - 2\rho\cos\theta}}, \quad \rho = r/a. \qquad (4.30)$$

Because Mercury is the inner-most planet and $\rho < 1$, this elliptic integral can be expanded in a power series. The first few terms are

$$V(r) = -\frac{GM_p m}{a} \left[1 + \frac{1}{4}\rho^2 + \frac{9}{64}\rho^4 + \frac{25}{256}\rho^6 + \cdots \right]. \qquad (4.31)$$

The force can be found via the gradient $-\nabla V$ as

$$\vec{F} = \frac{GM_p m}{2a^3} \left[1 + \frac{9}{8}\frac{r^2}{a^2} + \frac{75}{64}\frac{r^4}{a^4} \right] \vec{r}. \qquad (4.32)$$

This force is repulsive (radially outward), and the leading term behaves like Hooke's law with a negative spring constant. Two factors determine the magnitude of the force, the mass of the planet M_p and its average radius a. Accordingly, with everything being equal, we expect either a larger mass or a smaller semimajor axis will cause a greater precession. The actual amount depends on the interplay between the two parameters.

Using the force in the central field approximation (4.32) in Program 4.3 with the parameters from Table 4.1, we obtain the results of Table 4.2.

Table 4.2: Precession of Mercury due to the planets (arcseconds/century).

Venus	Earth	Mars	Jupiter	Saturn	Uranus	Neptune	Total
267.4	94.7	2.38	157.2	7.58	0.141	0.0437	529.4

The results vary over three orders of magnitude. Venus, the closest one, has the largest effect, followed by Jupiter, the most massive one. In between, the order reflects the balance between the radius and mass. Note that the results for Earth are obtained with the combined mass of Earth and the moon. Neptune, the outer-most planet, gives the smallest contribution, $0.0437''$, because

of its large radius and moderate mass. This is three orders of magnitude smaller than the relativistic contribution (43″). If 43″ was a needle in a hay stack, Neptune's contribution might be a drop in a lake. Yet, the leapfrog method with time transformation is robust enough to crunch it out accurately and efficiently. Also, although the central field approximated force (4.32) is positive, the precession is still counterclockwise, the same as the relativistic correction even though the force there (4.19) is negative. We will investigate these questions and related aspects in Project P4.3.

Scaling law in the central field approximation

As the numbers listed in Table 4.2 jump up and down with different planets, they are not particularly insightful for discovering trends or revealing more general properties. As Eq. (3.33) from Chapter 3 showed, scaling laws are very useful in this regard. Whenever possible, we should try to find scaled variables and plot the results in these variables to show the general trends.

What is a good scaled variable to use in this case? The amount of precession depends on the combination of the mass and radius. The prefactor in the potential (4.31) suggests M_p/a, but the force (4.32) points to M_p/a^3. Since the force is directly involved in the simulation, the latter is more appropriate. Therefore, for small perturbations, we expect precession to scale linearly with M_p/a^3, i.e., a linear scaling law.[7]

When we go ahead and calculate the scaled variable M_p/a^3 for the planets using Table 4.1, we find that it also varies over three orders of magnitude similar to the precession results in Table 4.2. If we were to display precession versus the scaled variable as linear plots, the data points would be clustered at two locations, the lower-left and upper-right corners. Thus, to more evenly spread

out the data points in the whole range and to help visually identify any trends, we should plot the data using log-log scales as we did with the semilog plot in Figure 2.7, for example. The results are shown in Figure 4.8.
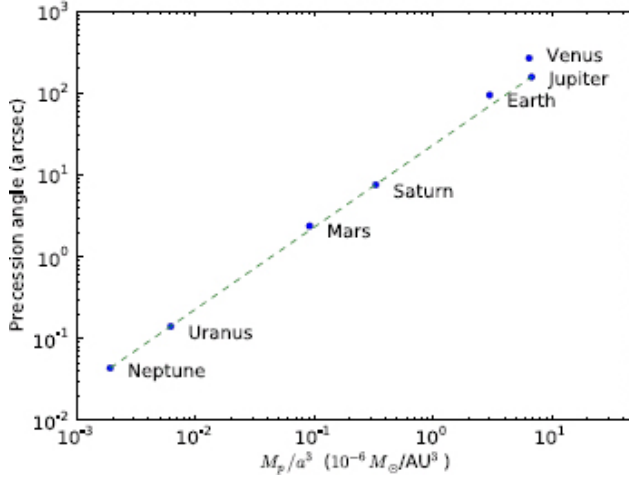

Figure 4.8: The precession of Mercury due to other planets as a function of the scaled variable M_p/a^3. The dashed trend line indicates a linear scaling law expected for small perturbations in the central field approximation.

Most data points in Figure 4.8 fall on a straight line (trend line), clearly showing a linear scaling law. Small deviations are seen for Venus, Earth, and Mars. This is due to their relative proximity to Mercury, causing the second order term in Eq. (4.32) to have a non-negligible effect on the deviation. Nonetheless, Venus and Jupiter, the closest and the most massive planets, are vastly different and yet nearly identical in the scaled variable, 6.47 vs. 6.78, respectively (Figure 4.8). Therefore, they cause similar amount of precession according to the scaling law. The predicative power and the physical insight of scaling laws make them very useful in practice and in our understanding.

Since each planet produces precession in the same counter-clockwise direction, the total precession due to all planets is addi-

tive, and is equal to 529.4″ (Table 4.2). The result is in very good agreement with the accepted observational data of 531″ [72, 85]. The relative error is about 0.3%, with most of it coming from Venus. What are the sources of error? We have assumed the planets are coplanar with Mercury, which is only approximate. The inclination angles of the planets differ from each other by a few degrees. Another source is the assumption of circular orbits (rings). From the eccentricities in Table 4.1, except for Pluto which has negligible effects and Mercury itself, the perturbing planets are nearly, but not perfectly, circular. Lastly, we have included only a few terms in the series expansion (4.31). This will affect only the close planets though, such as Venus and Earth.

4.4 Star wobbles and exoplanets

We have seen from Section 4.2 that a two-body problem can be reduced to an effective one-body problem in the center of mass frame. The actual motion, of course, still consists of two bodies pulling on each other (with forces of equal magnitude and opposite direction) and revolving around the center of mass. In the case of a star-planet system, the star barely wobbles due to the large mass ratio.

Nonetheless, the star's wobbling is important to a current topic of considerable interest: the discovery and detection of extrasolar planets, or exoplanets, including Earth-like planets and planets in the habitable zone supporting an atmosphere and liquid water.[8] Unlike planets at the doorstep in our solar system, direct observation of exoplanets is very difficult because they are at astronomical distances, small and faint, and in sharp contrast to the brightness of their host stars. Indirect methods of detection are neces-

sary, including measuring the wobbles of stars (radial velocity), the light blocked off when exoplanets pass in front of the stars (transiting planets), and light bending and brightening due to gravitational fields (microlensing), among others.

4.4.1 RADIAL VELOCITY METHOD

We discuss the method of detecting the wobbles, known as the radial velocity (RV) method. The RV method relies on the fact that light from stars moving about the center of mass will be Doppler-shifted. The Doppler shifts depend on the velocity of the star along the line of sight, i.e., the radial velocity from the observer's perspective. Light from the star will be blue-shifted moving toward us, and red-shifted moving away. By measuring the Doppler profile, one can infer the velocity profile of orbital motion. From there, we can determine the orbital parameters of the exoplanets, including the period, eccentricity, and the lower limit of the mass. The RV method is responsible for a majority of exoplanets discovered early on [14].[9]

Referring to the position vectors in Figure 4.2 and Eq. (4.4) (with the results from Exercise E4.1), the wobbling velocity of the star, $\vec{v}_* \equiv \dot{\vec{r}}_\mathrm{P}$, is related to the orbital velocity $\vec{v} = \dot{\vec{r}}$ as

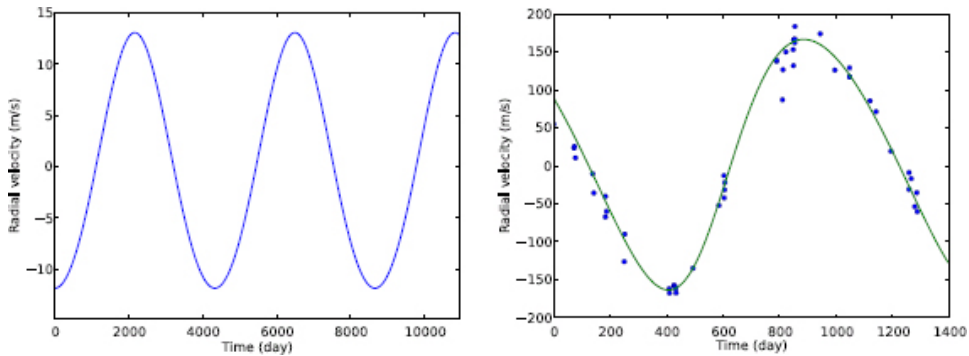$$\vec{v}_* = -\frac{m}{m + M}\vec{v}. \qquad (4.33)$$

Figure 4.9: The wobbling velocity of the Sun due to Jupiter (left) and of HD 139357 due to its exoplanet (right). The solid curve is a fit described in text.

For a given star-planet system, we can obtain $\vec{v}$ from Program 4.1 and therefore $\vec{v}_*$. The result for the Sun-Jupiter system is shown in Figure 4.9. One component of the Sun's velocity is plotted, $\vec{v}_{*y}$ in this case, which is equivalent to the edge-on view of the orbital motion. The velocity is periodic ($\sim$ 4330 days, Jupiter's period), and nearly sinusoidal because the orbit is almost circular (e = 0.0484, see Table 4.1). It would be sinusoidal for circular orbits. The magnitude of the velocity is roughly 10 m/s, typical of the pull by planets of Jupiter's size on a star of about one solar mass. For smaller planets like Earth, the magnitude is about 10 cm/s.

Also shown in Figure 4.9 is an observational RV dataset for a star in the Henry Draper catalog, HD 139357. It shows an almost sinusoidal curve, clearly indicating the presence of an exoplanet moving in a nearly circular orbit around the star, with a period $\sim$ 1100 days. This exoplanet was discovered in 2009. From the RV dataset, the lower limit of the exoplanet's mass is determined to be about 9.8 Jupiter mass. The fact that only a mass limit can be established, and not the mass itself, has to do with observing geometry and information on the inclination angle, or lack thereof.
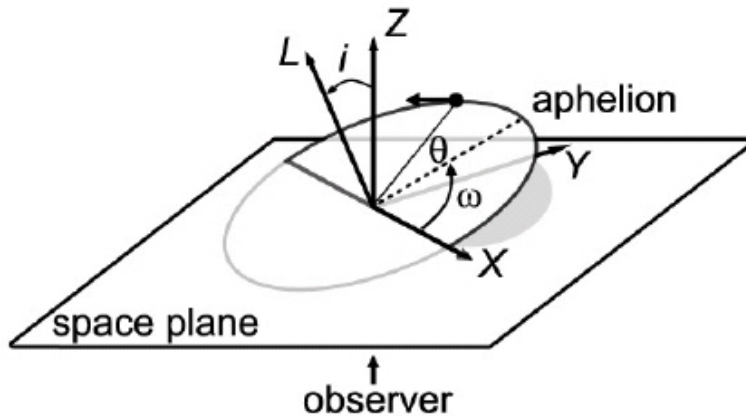

226

Figure 4.10: Observing geometry for radial velocity
measurements.

The geometry is illustrated in Figure 4.10. The space-fixed co-ordinates (the sky) are denoted by *XYZ,* where the *XY* plane defines the space plane. We choose the *X*-axis along the line where the Kepler orbit of the exoplanet intersects the space plane (ascending node in astronomy nomenclature). The inclination angle, *i,* is the angle between the orbital plane and the space plane. It is equal to the angle between the *Z*-axis and the normal of the orbit (direction of angular momentum $\vec{L}$). We use ω to denote the angle between the aphelion and the *X*-axis. Except for the azimuthal angle around *Z* (not shown), which is unimportant due to rotational symmetry, *i* and ω determine the orientation of the Kepler orbit. As usual, θ represents the angle between the radial vector of the exoplanet and the aphelion, the same as in Figure 4.3.

The observer's line of sight to the star runs parallel to the *Z*-axis. Therefore, only the *Z* component of the star's velocity, v_{*z}, is relevant and measurable by Doppler shift methods. From the observer's perspective, v_{*z} appears as a radial velocity (away or toward), thus the name of the RV method.[10] We have to determine the relationship between the wobbling velocity v_{*z} and the orbital velocity $\vec{v}$ in terms of orbital parameters and the angles *i* and ω.

The radial velocity of the star is derived in Section 4.C. For a star with a single planet, the result from Eq. (4.85) is

$$v_{*z} = -\mathcal{V}[\cos(\omega + \theta) - e\cos\omega] + C, \quad \mathcal{V} = \frac{m\sin i}{m+M}\frac{2\pi a}{T\sqrt{1-e^2}}, \quad (4.34)$$

where *C* is a constant representing the velocity of the center of mass, *T* the period, and *a* the semimajor axis. The first term indicates that v_{*z} oscillates with a period *T* and an amplitude $|\mathcal{V}|$.

227

From fitting a given RV dataset like the one shown in Figure 4.9 according to Eq. (4.34), parameters such as T, e, and V can be obtained. If the mass of the star M is independently known, the mass of the exoplanet m can be determined from

$$\frac{m \sin i}{m + M} \simeq \frac{m \sin i}{M} = \sqrt{1 - e^2} \left(\frac{T}{2\pi GM}\right)^{1/3} \mathcal{V}. \qquad (4.35)$$

We have assumed $m/M \ll 1$ and used Kepler's third law (4.18) to eliminate a. Equation (4.35) shows that we can determine the product $m \sin i$ only. Because the inclination angle i is generally unknown, the RV method yields only a lower limit for the exoplanet mass ($\sin i \le 1$).

If the period T is given in years, M in solar masses, and V in m/s, the numerical value for the exoplanet mass in Earth masses is

$$m \sin i = 11\sqrt{1 - e^2} \, (TM^2)^{1/3} \mathcal{V}. \quad \text{(in Earth mass)} \qquad (4.36)$$

This formula is useful for a rough estimate of an exoplanet's mass.

Take for instance the exoplanet of the star HD 139357. The star is known to be about 400 light-years away and a mass $M \sim 1.3$ $M_\odot$. We can estimate the other parameters by inspecting Figure 4.9 to obtain: $T \sim 1100$ days $= 3$ years, and $V \sim 160$ m/s. The RV curve is almost sinusoidal, so the orbit is nearly circular, and $e \sim 0$. Substituting these numbers into Eq. (4.36), we obtain $m \sin i \sim 3000$ Earth mass, or ~ 9.5 Jupiter mass. More careful data fitting gives $m \sin i = 9.8$ Jupiter mass, which means that the exoplanet is at least 9.8 times more massive than Jupiter.

4.4.2 MODELING RV DATASETS

To accurately model observational RV datasets according to Eq. (4.34), we need to know θ as a function of time t, $\theta(t)$. However,

this is not as straightforward as one might expect. Except for circular orbit where θ is linear in t, no simple closed-form solutions exist. We could obtain $\theta(t)$ numerically from simulation codes such as Program 4.3, but that would be the hard way. Fortunately, it is a two-way street here, a hard way and an easy way.

We first write $\dot{r}$ using conservation of energy from Eq. (4.8) and obtain t as

$$t = -\sqrt{\frac{m}{2}} \int_{r_{max}}^{r} \frac{dr}{\sqrt{E - V_{eff}(r)}}. \tag{4.37}$$

This equation can be solved by introducing a new variable ψ through

$$r = a(1 + e\cos\psi), \tag{4.38}$$

where ψ is called the eccentric anomaly.[11] Substituting Eq. (4.38) into (4.37), and after some algebraic details left to Exercise E4.7, we obtain

$$\frac{t}{T} = \frac{1}{2\pi}(\psi + e\sin\psi). \tag{4.39}$$

This is the celebrated Kepler's equation relating t to ψ. We can solve for θ in terms of ψ by combining Eqs. (4.38) and (4.13) to get

$$\cos\theta = \frac{e + \cos\psi}{1 + e\cos\psi}. \tag{4.40}$$

We see from Eq. (4.40) that ψ has the same range as θ, i.e., when θ changes from 0 to 2π, so does ψ. Furthermore, the two variables are equal at 0, π, and 2π, no matter the eccentricity e.

The pair of equations (4.39) and (4.40) give the relationship between t and θ. For a given t, Eq. (4.39) can be solved for ψ,

which can be substituted into Eq. (4.40) to obtain θ. The relationship for several eccentricities is depicted in Figure 4.11. The dependence is linear for $e = 0$ (circular orbit) as expected. For non-zero e, it becomes increasingly nonlinear. The figure was generated without actually solving Eq. (4.39). The trick is this: first generate a mesh for ψ in $[0, 2\pi]$; second, compute t from Eq. (4.39) for each mesh point and store it in an array; third, compute θ from Eq. (4.40) and store it in another array; finally plot the two arrays. This trick is easy and fast, but the drawback, of course, is that the t values calculated are not pre-determined and non-uniform. If we wanted to generate θ for a given t, we would have to solve Eq. (4.39) using a root solver. It's not necessary for our purpose here.
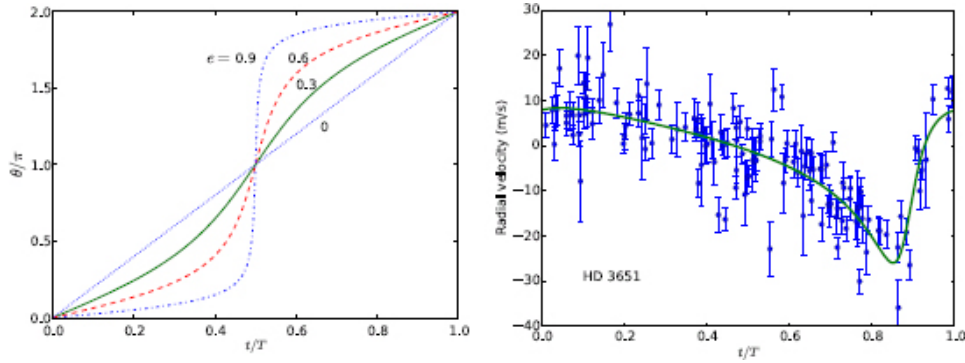

Figure 4.11: Left: The relationship between θ and t for four eccentricities. Right: an actual RV dataset of Fischer *et al.* [30] and a fit.

The effect of eccentricity on radial velocities is also shown in Figure 4.11. The dataset is for HD 3651 with an exoplanet reported in 2003 [30]. The curve is not sinusoidal at all. The sharp dip in the radial velocity can be reproduced only with a significant value of eccentricity. Physically, the planet moves much faster near the perihelion than anywhere else in a highly eccentric orbit. This causes a dip (or spike) in the velocity of the recoiling star.

A general best fit of Eq. (4.34) to a given RV dataset is not trivial because of the number of parameters involved. We use least square fitting, `leastsq`, from SciPy as a blackbox. It requires only an error function that returns the difference between our model and the actual data points. The details are given in Program 4.4, yielding the results in Figure 4.11.

The blackbox method is easy to use and gives satisfactory results, but leaves the modeling process a bit opaque. To see various factors at play, we outline an iterative fitting process by simple visual inspection that also works well for our purpose. Starting with an initial guess of parameters T, V, e, ω, C:

1. generate a mesh grid for ψ
2. adjust the appropriate parameters T, V, e, ω, C for better visual fit
3. compute t from Eq. (4.39) and θ from Eq. (4.40)
4. calculate v_{*z} from Eq. (4.34) and plot the results
5. inspect the match of the curve and the data, repeat step 2 as necessary

With code segments from Program 4.4, we can begin to model any datasets and visualize the process. We leave modeling of the dataset shown in Figure 4.11 to Project P4.6. Many other RV datasets can be found at the exoplanet archive [66].

4.5 Planar three-body problems

So far we have studied motion of two-body systems reducible to effective one-body motion. If we add just one more body to a two-body system, we have a three-body problem, the simplest of gen-

eral *N*-body problems. Even if the pair-wise forces obey the pure inverse-square law (4.1), general analytic solutions to the three-body problem are still not possible at present, making simulations indispensable.

Since the early development of Newtonian mechanics, efforts have been made to reduce the intractability of the three-body problem. There are nine degrees of freedom in total, three for each body. Going to the center of mass frame reduces that number by three. Further reduction is possible by invoking conservation laws (energy, angular momentum, etc.), but the mathematical difficulties involved are still insurmountable. Special attention, therefore, has been devoted to finding solutions of the planar three-body problem: the motion of all bodies is confined to a two-dimensional plane. This is still an area of ongoing research, especially in the search of periodic orbits. We will discuss several numerical examples to get a glimpse of the planar three-body problem and its intricacies.

4.5.1 EQUATIONS OF MOTION IN THE THREE-BODY PROBLEM

Qualitatively, the motion of three bodies is no more difficult to simulate than one-body motion. Let the three bodies be numbered as $i = 1, 2$, and 3. Their masses, coordinates, and velocities are m_i, $\vec{r}_i$, and $\vec{v}_i$, respectively. The equations of motion are

$$\frac{d\vec{r}_i}{dt} = \vec{v}_i, \tag{4.41}$$

$$\frac{d\vec{v}_i}{dt} = -Gm_j\frac{\vec{r}_{ij}}{r_{ij}^3} - Gm_k\frac{\vec{r}_{ik}}{r_{ik}^3}. \tag{4.42}$$

It is understood that $i \neq j \neq k$, and $\{ijk\}$ is a permutation of $\{123\}$. The relative coordinates between two bodies are $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$, and $r_{ij} = |\vec{r}_i - \vec{r}_j|$. Note that $\vec{r}_{ij} = -\vec{r}_{ji}$, a relation useful in programming. Because the gravitational constant G always appears in the product Gm_i, we can set $G = 1$ in our program. It is equivalent to setting the scale of mass.

Equations (4.41) and (4.42) are 12 first order differential equations of a Hamiltonian system in 2D. Given the initial condition, we can solve them using the leapfrog method. Of course, the more elaborate leapfrog method with time transformation discussed earlier may be extended to the three-body problem as well.[12] However, we are not interested in very small or sensitive processes like precession of Mercury, so its use is unnecessary here. In addition, most of the periodic orbits are physically unstable, and there is little a numerical algorithm can do to rectify that. What remains is to determine the initial condition for a given numerical solution.

4.5.2 EULER'S COLLINEAR MOTION

One of the simplest and most interesting systems is Euler's collinear motion of the three-body problem. It is illustrated in Figure 4.12.

At any given time, the three bodies are on a straight line, with one body, say m_2, lying in-between the other two bodies, m_1 and m_3. Each moves in an ellipse with the same period. When they reach the aphelion simultaneously on the axis of the ellipses, the distances among them are the largest, and when they pass the perihelion, the distances are the closest. At other times, the distances oscillate between the extrema. The motion is periodic, but unstable. That is to say, if there is a slight deviation from the perfect configuration, the system flies apart from collinear motion.
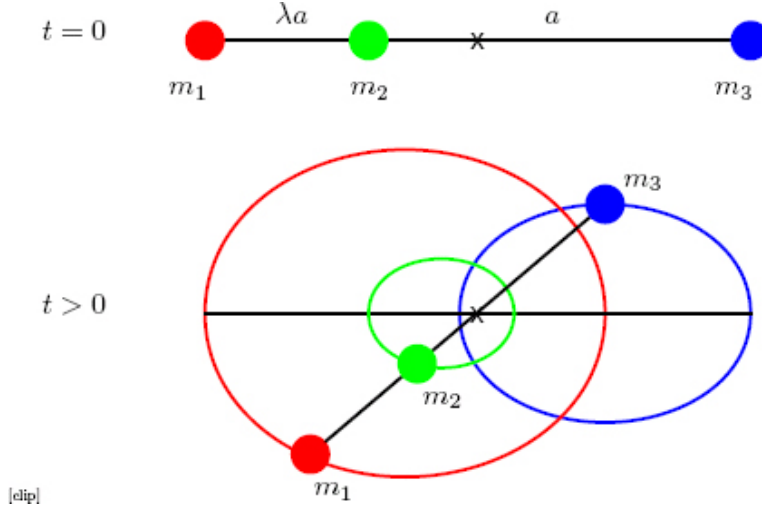
Figure 4.12: Euler's collinear motion of the three-body problem.

To determine the initial condition, let us assume the bodies start from the *x*-axis, thus $y_i = 0$, and only the horizontal positions x_i are needed. To execute Euler's collinear motion, the distances between the bodies must obey certain proportions. Referring to Figure 4.12 (top), let *a* be the distance between m_2 and m_3, and λa between m_1 and m_2. It follows that

$$x_3 - x_2 = a, \ x_2 - x_1 = \lambda a, \ x_3 - x_1 = (1 + \lambda)a. \qquad (4.43)$$

The equation of motion for a given body, say m_1, is

$$m_1 \frac{d^2 x_1}{dt^2} = -m_1 m_2 \frac{x_1 - x_2}{|x_1 - x_2|^3} - m_1 m_3 \frac{x_1 - x_3}{|x_1 - x_3|^3},$$

where we have set $G = 1$ for convenience. Because of the oscillating nature of the distances, we seek a solution of the form $d^2 x_1/dt^2 = -\omega^2 x_1$ (simple harmonic oscillator). Substituting this into the above equation and dividing $-m_1 \omega^2$ on both sides, we obtain

$$x_1 = \frac{1}{\omega^2} \left[m_2 \frac{x_1 - x_2}{|x_1 - x_2|^3} + m_3 \frac{x_1 - x_3}{|x_1 - x_3|^3} \right]. \qquad (4.44)$$

234

Similarly, we have for the other two bodies

$$x_2 = \frac{1}{\omega^2}\left[m_1\frac{x_2 - x_1}{|x_2 - x_1|^3} + m_3\frac{x_2 - x_3}{|x_2 - x_3|^3}\right], \quad (4.45)$$

$$x_3 = \frac{1}{\omega^2}\left[m_1\frac{x_3 - x_1}{|x_3 - x_1|^3} + m_2\frac{x_3 - x_2}{|x_3 - x_2|^3}\right]. \quad (4.46)$$

We can combine Eqs. (4.43) to (4.46) to obtain the parameter λ first,

$$\lambda^5(m_2 + m_3) + \lambda^4(2m_2 + 3m_3) + \lambda^3(m_2 + 3m_3)$$
$$+ \lambda^2(-3m_1 - m_2) + \lambda(-3m_1 - 2m_2) + (-m_1 - m_2) = 0. \quad (4.47)$$

This is known as Euler's quintic equation. We leave as an exercise the verification of Eq. (4.47) and the results below, Eqs. (4.48) to (4.50). Given a set of masses, there is one positive root. The variable a in terms of λ is

$$a^3 = \frac{1}{\omega^2}\left[m_2 + m_3 - \frac{m_1(1 + 2\lambda)}{\lambda^2(1 + \lambda)^2}\right]. \quad (4.48)$$

With both λ and a found, we can find the initial positions as

$$x_2 = \frac{1}{\omega^2 a^2}\left[\frac{m_1}{\lambda^2} - m_3\right], \quad x_1 = x_2 - \lambda a, \quad x_3 = -\frac{m_1 x_1 + m_2 x_2}{m_3}. \quad (4.49)$$

We recognize that the last term means that the center of mass is the origin. The initial velocities (all along y-axis) are

$$v_{1y} = \omega x_1, \quad v_{2y} = \omega x_2, \quad v_{3y} = \omega x_3. \quad (4.50)$$

The procedure to simulate Euler's collinear motion is as follows. First, solve Eq. (4.47) for λ using a root finder such as bisection or Newton's methods in Chapter 3. Then compute the initial condition from Eqs. (4.49) and (4.50). Finally, solve the equations of motion (4.41) and (4.42) with the leapfrog method.

The code implementing this strategy is given in Program 4.5 listed in Section 4.D. The program animates the collinear motion of the three-body problem and snapshots are shown in Figure 4.13.

Watching the motion, we should see that at first the bodies move perfectly (to the naked eye) in their own ellipses, retracing over them periodically (Figure 4.13, left). After a while, typically three periods or so depending on the energy, two bodies will be so attracted to each other that they quickly move off the periodic orbits, forming a binary system. The binary partners usually perform choreographically beautiful (albeit mechanical) "dances" (Figure 4.13, right) as they move along. The third body, attracted to the binary system acting roughly as a single body, often turns around. If it does and moves in close, there is a high probability that the dancing partners switch, forming a new binary system. This may continue for a while, but eventually, the final outcome is that a tight binary is formed, and the third body is ejected, moving indefinitely away from the binary system. Gravitationally bound classical three-body systems are mostly unstable, and often so sensitive to the initial conditions that a small difference there can lead to totally different outcomes. This behavior is known as classical chaos, a topic we will discuss next in Chapter 5. Our knowledge at present seems to indicate that common many-body systems held by gravity are chaotic, including our solar system [58].
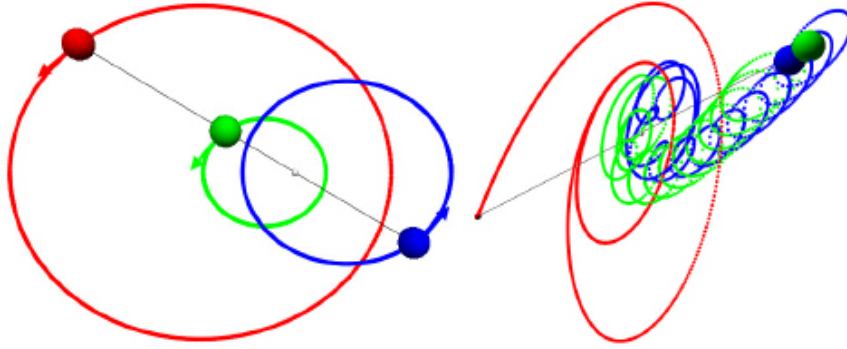
Figure 4.13: Euler's collinear motion of the three-body problem at different times. The camera in the right figure is tilted to keep the third body in view.

4.5.3 THREE-BODY CHOREOGRAPHY

There are more elaborate periodic orbits that have been found. For example, the following initial condition [84] produces a heart-shaped choreographic orbit shown in Figure 4.14.

$$
\begin{aligned}
&m_1 = m_2 = m_3 = 1/3, \\
&x_1 = -0.30805788, y_1 = 0, v_{1x} = 0, v_{1y} = -1.015378093, \\
&x_2 = 0.15402894, y_2 = -0.09324743, v_{2x} = 0.963502817, v_{2y} = 0.507689046, \\
&x_3 = x_2, y_3 = -y_2, v_{3x} = -v_{2x}, v_{3y} = v_{2y}.
\end{aligned} \tag{4.51}
$$

The three bodies rotate counterclockwise, so there is a net angular momentum. The starting position of each body is approximately at the position of the body immediately behind it (clockwise) in the left panel of Figure 4.14. The most delicate part occurs near the center of mass (CM, white dot) where the speed is the lowest. Referring to the animation (or Figure 4.14, right), as the red body (closest to the CM) moves in on the upper part of the inner orbit, it slows down, coming to a temporary stop right at the tip (the protrusion) of the inner orbit. At the same time, the other two bodies (light/green and dark/blue) are almost diametrically positioned on the outer orbit, so the net force on the red body is

237

very small. But, there is a slight bend to the diametrical which provides just enough net force that allows the red body to step back from the tip and continue on to the lower part of the inner orbit (where the body immediately ahead (blue) was previously). However, the motion is unstable. After about two periods, the system breaks up, suffering the same fate of most three-body systems: a tight binary plus a runaway body (not shown in Figure 4.14). But, if we reversed the velocities at any time, the system would run backward to its original configuration precisely like rewinding a videotape (see Exercise E4.9). This is because the leapfrog method preserves time reversibility of the system, which may be unstable but not indeterministic.
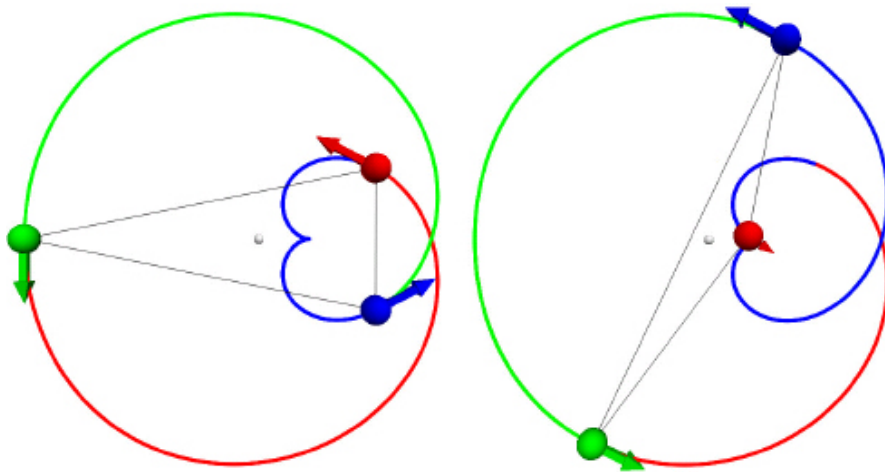

Figure 4.14: Choreographic orbit of the three-body problem with the initial condition (4.51). The arrows are proportional to the velocities.

Another interesting orbit is a figure-8 orbit. It is shown in Figure 4.15 produced from the following initial condition [15],

$$m_1 = m_2 = m_3 = 1,$$
$$x_1 = 0.97000436, y_1 = -0.24308753, v_{1x} = -v_{3x}/2, v_{1y} = -v_{3y}/2,$$
$$x_2 = -x_1, y_2 = -y_1, v_{2x} = v_{1x}, v_{2y} = v_{1y},$$
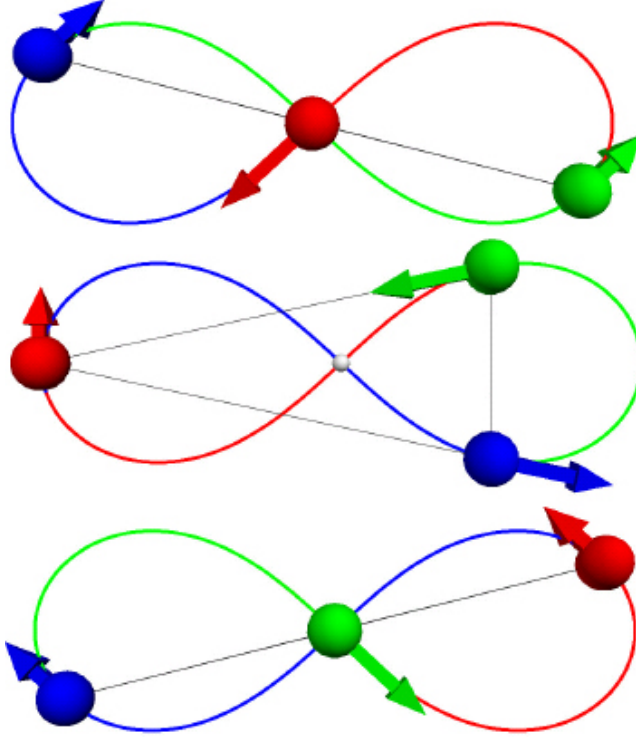$$x_3 = 0, y_3 = 0, v_{3x} = -0.93240737, v_{3y} = -0.86473146. \tag{4.52}$$

Figure 4.15: Choreographic orbit of the three-body problem with the initial condition (4.52). The arrows are proportional to the velocities.

This orbit is symmetric both horizontally and vertically. Unlike the previous configuration, the bodies do not circulate uniformly in one direction. The motion is clockwise on the left half of the orbit (Figure 4.15), and counterclockwise on the right half. The combination is such that there is no net angular momentum. When one body is at the origin, the other two are symmetrically situated on the orbit, all on a straight (slanted) line. In this sense it is like Euler's collinear configuration (Figure 4.13), although the three bodies never line up horizontally. Whenever one body is at

the edge of the figure-8 (Figure 4.15, middle), the other two are on the opposite half lined up vertically, forming an isosceles triangle. At this moment, the velocity of the first body is vertical, and the velocities of the other two bodies are such that their x components cancel each other and the sum of their y components cancels the y component of the first body, so that the center of mass velocity remains zero.

Numerical experiments show a rare but remarkable property of this orbit: it is stable. If the orbit is perturbed slightly, the motion remains close to the figure-8 shape. For example, if the position or the velocity of a body (or even its mass) is changed by a small amount, the orbit may be slightly deformed, or it may even precess, but the three-body system does not disintegrate (Exercise E4.9).

4.6 The restricted three-body problem

The planar three-body problem does not have any restrictions other than that the motion is confined to a plane. We impose now two additional constraints: the masses of two primary bodies are much larger than the mass of the third body, and the two primaries move in a circular orbit about their center of mass. As a result, the simplified system is called the restricted three-body problem. It is illustrated in Figure 4.16.
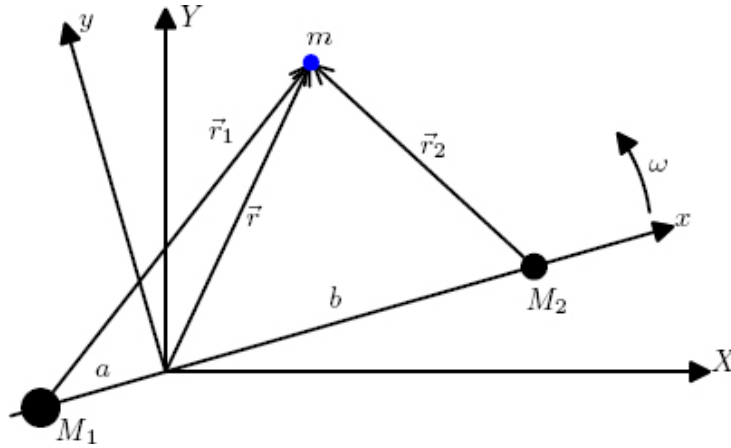
Figure 4.16: The restricted three-body coordinate system.

Let us call the third body of mass m the test body. Its mass is considered so small compared to the two primaries, i.e., $m/M_{1,2} \sim 0$, that it has no effect on the motion of the primaries, which rotate around the center of mass at constant angular velocity ω and constant separation. We are only concerned with the motion of the test body. Though the restricted three-body problem is an idealized situation and no actual systems behave exactly like it, it is still a very good approximation for many systems, including the motion of a satellite (test body) in the Earth-Moon system, or of the asteroids in the Sun-Jupiter system. The simplification allows us to gain much insight from studying such systems.

The motion of the test body may be best understood in the rotating coordinate system of the primaries. For example, if a and b are the respective distances of M_1 and M_2 from the center of mass (the origin, Figure 4.16), then the primaries appear as stationary points at $(-a, 0)$ and $(b, 0)$, respectively, in the rotating coordinates (x-y), though they are rotating in space-fixed coordinates (X-Y). The test body will move around the stationary force centers in the rotating coordinates.

4.6.1 EQUATIONS OF MOTION IN THE ROTATING SYSTEM

The force on the test body is

$$\vec{F} = -GM_1 m \frac{\vec{r}_1}{r_1^3} - GM_2 m \frac{\vec{r}_2}{r_2^3}, \qquad (4.53)$$

$$r_1 = \sqrt{(x+a)^2 + y^2}, \ r_2 = \sqrt{(x-b)^2 + y^2},$$

where (x, y) are the coordinates of the test body in the rotating system. The calculation of velocity or acceleration must take the rotating coordinates into account. Details are given in Appendix 4.A. For example, according to Eq. (4.72) the acceleration is,

$$\vec{a} = \vec{F}/m - 2\vec{\omega} \times \vec{v} - \vec{\omega} \times (\vec{\omega} \times \vec{r}). \quad \text{(rotating frame)} \qquad (4.54)$$

Note that we have dropped the prime (′) according to our coordinate system in Figure 4.16 for clarity. Of course, all variables $\vec{r}$, $\vec{v}$, and $\vec{a}$ refer to the rotating coordinates.

Using the fact $\vec{\omega} = \omega \hat{k}$ and $\vec{A} \times (\vec{B} \times \vec{C}) = (\vec{A} \cdot \vec{C})\vec{B} - (\vec{A} \cdot \vec{B})\vec{C}$, we can simplify the last term in Eq. (4.54) to

$$\vec{\omega} \times (\vec{\omega} \times \vec{r}) = (\vec{\omega} \cdot \vec{r})\vec{\omega} - (\vec{\omega} \cdot \vec{\omega})\vec{r} = -\omega^2 \vec{r}.$$

We have used $\vec{\omega} \cdot \vec{r} = 0$ because $\vec{\omega}$ is perpendicular to the plane of motion (x-y plane). Substituting these results and Eq. (4.53) into (4.54), we obtain

$$\vec{a} = -GM_1 \frac{\vec{r}_1}{r_1^3} - GM_2 \frac{\vec{r}_2}{r_2^3} + \omega^2 \vec{r} - 2\vec{\omega} \times \vec{v}. \qquad (4.55)$$

The system can be readily integrated with our ODE solvers. However, before doing that, we would like to examine the rich struc-

ture contained within the effective potential which will help us better understand the motion.

4.6.2 THE EFFECTIVE POTENTIAL

There are four terms in Eq. (4.55). We recognize that the first two terms are actual forces due to the primaries M_1 and M_2, the third term due to the centrifugal force (or effect), a pseudo-force, and the last term due to the Coriolis effect. Except for the velocity-dependent Coriolis term, the other three terms may be obtained from the gradient of a scalar function of coordinates only, say $V(x, y)$, such that the acceleration can be written as

$$\vec{a} = -\nabla V - 2\vec{\omega} \times \vec{v}. \qquad (4.56)$$

We call V the effective potential given by

$$
\begin{aligned}
V(x, y) &= -\frac{GM_1}{r_1} - \frac{GM_2}{r_2} - \frac{1}{2}\omega^2 r^2 \\
&= -\frac{GM_1}{\sqrt{(x+a)^2 + y^2}} - \frac{GM_2}{\sqrt{(x-b)^2 + y^2}} - \frac{1}{2}\omega^2(x^2 + y^2). \quad (4.57)
\end{aligned}
$$

The effective potential consists of the actual gravitational potentials, and a centrifugal potential which resembles the harmonic oscillator potential but with a negative spring constant (see Eq. (4.32)). The force is therefore repulsive and pushes radially outward. This is because the motion is being described in the rotating system. We could imagine that, to keep the test body stationary in the rotating system, a real force toward the center, i.e., a centripetal force, would be needed to balance an extra (unseen) force to keep it in place. That extra force is called the centrifugal force.

Units for restricted three-body problems

The effective potential (4.57) (and equations of motion) are valid for any unit system. To further analyze its properties quantitatively, we should adopt a unit system most convenient for the restricted three-body (RTB) problem at hand. Analogous to planetary motion, we choose the RTB units as shown in Table 4.3.

Table 4.3: RTB unit system for the restricted three-body problem.

Mass	$M = M_1 + M_2$, total mass of the system
Length	$R = a + b$, the separation between the primaries
Time	T, rotational period of the primaries

Consequently, it follows by definition that $\omega = 2\pi/T$, and from Eqs. (4.16) and (4.17) that

$$GM = G(M_1 + M_2) = 4\pi^2 \frac{R^3}{T^2}. \tag{4.58}$$

This means that in actual calculations employing the RTB unit system ($M = R = T = 1$), we can set angular velocity $\omega = 2\pi$ and $GM = 4\pi^2$, just as in the planetary unit system.

For the distances, because the center of mass is chosen as the origin (Figure 4.16), we have

$$M_1 a = M_2 b, \ R = a + b.$$

Solving for a and b and introducing the mass parameter α, we have

$$a = \frac{M_2}{M_1 + M_2} R = \alpha R, \ b = \frac{M_1}{M_1 + M_2} R = (1 - \alpha)R, \ \alpha = \frac{M_2}{M_1 + M_2}. \tag{4.59}$$

It is customarily assumed that $M_2 \leq M_1$, so $0 < \alpha \leq 0.5$. In terms of α, the primaries' masses in our unit system can be expressed as

$$GM_1 = \frac{GM_1}{GM} GM = 4\pi^2(1-\alpha), \ GM_2 = \frac{GM_2}{GM} GM = 4\pi^2\alpha. \quad (4.60)$$

Consider the Earth-Moon system as an example. The primaries' masses are $M_1 = 5.974 \times 10^{24}$ kg and $M_2 = 7.348 \times 10^{22}$ kg, respectively. The distance between them is 3.844×10^8 m, and the rotational period is about one month (27.3 days). The absolute units and the RTB units are listed in Table 4.4.

Table 4.4: RTB units of the Earth-Moon system.

Quantity	Absolute units	RTB units
Unit of mass	$M_1 + M_2 = 6.047 \times 10^{24}$ kg	1
Unit of length	$R = 3.844 \times 10^8$ m	1
Unit of time	$T = 2.36 \times 10^6$ s (27.3 days)	1
Unit of ω	2.66×10^{-6} rad/s	2π
Unit of velocity	163 m/s	1
Unit of acceleration	6.91×10^{-5} m/s^2	1
Mass parameter α	$M_2/(M_1 + M_2) = 0.0121$	0.0121
Earth's mass	$GM_1 = 3.98 \times 10^{14}$ m^3/s^2	$0.988 \times 4\pi^2$
Moon's mass	$GM_2 = 4.90 \times 10^{12}$ m^3/s^2	$0.0121 \times 4\pi^2$
Earth's position	$-a = -4.67 \times 10^6$ m	-0.0121
Moon's position	$b = 3.80 \times 10^8$ m	0.988

We note that the RTB units are elastic. In other words, when the restricted three-body system changes, e.g., to the Sun-Jupiter system, the units of mass, length, etc., have different numbers on the absolute scale, even though they may all be 1 in the RTB units.

The Lagrange points

We can simplify the effective potential (4.57) by rewriting it in the newly introduced RTB unit system as

$$V(x,y) = -4\pi^2 \left[\frac{1-\alpha}{\sqrt{(x+\alpha)^2 + y^2}} + \frac{\alpha}{\sqrt{(x-1+\alpha)^2 + y^2}} + \frac{x^2 + y^2}{2} \right].$$
$$(4.61)$$

One single mass parameter α fully determines the effective potential. The potential surface for $\alpha = 0.121$ is shown in Figure 4.17. This α is about ten times larger than that of the actual Earth-Moon system in order to show some fine features.

Figure 4.17 is produced with Program 4.6. The program is explained in detail there. However, the block (lines 21 to 34) can be used as a template for making surface and contour plots.

The surface plot (Figure 4.17, top) shows several expected features. The potential has two holes (singularities) at the primaries, M_1 at $(-a, 0)$ and M_2 at $(b, 0)$. It also goes to $-\infty$ at large r due the centrifugal potential $-r^2/2$. In between, there is a broad C-shaped ridge surrounding the bigger body M_1. The formation of the ridge is due to the balance of the gravitational and centrifugal potentials. As the ridge nears the other body M_2, a hole is "burned" into it, and a saddle point is created between M_1 and M_2.
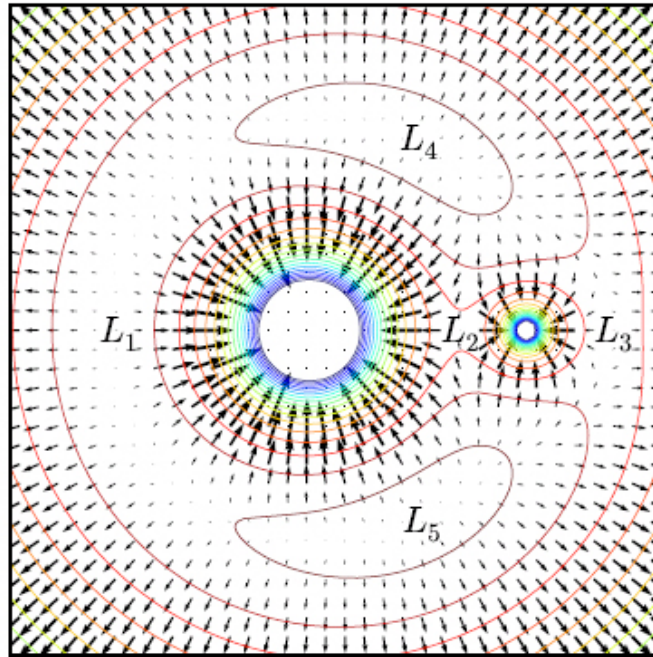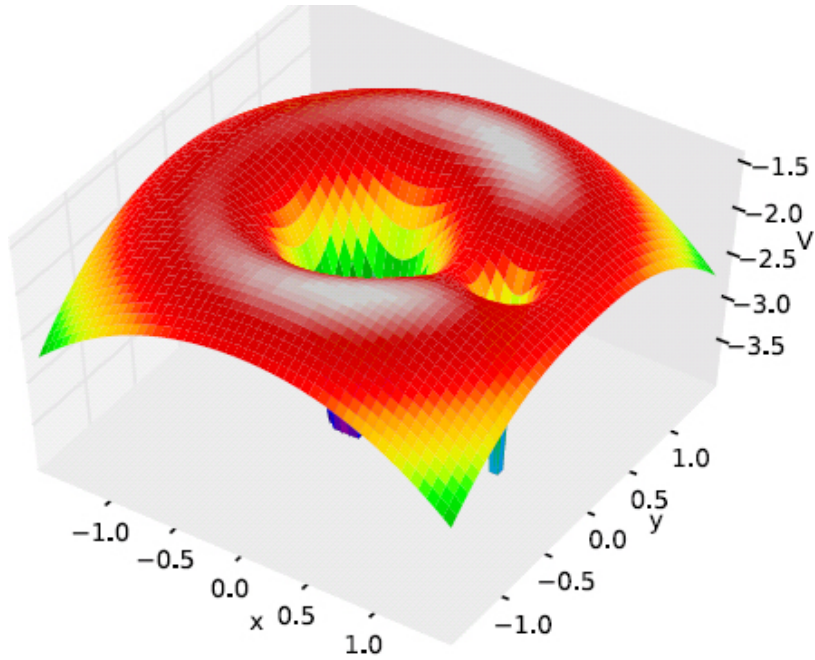
Figure 4.17: The effective potential of a restricted three-body system with $\alpha = 0.121$. The arrows on the contours indicate the force from the potential.

The contour plot (Figure 4.17, bottom) shows that there are finer structures embedded along the ridge. There are five points

247

where the potential is locally maximum. They are known as the Lagrange points, labeled L_1 to L_5.[13] Three of them, L_1, L_2 and L_3, are located on the x-axis, and the other two, L_4 and L_5, are off the axis but symmetrically located about it, one above and one below. On closer look, the locations of the primaries and L_4 (or L_5) form an equilateral triangle.

The gradient of the potential is graphically illustrated by the arrows. Since $\vec{F} = -\nabla V$, the arrows are proportional to the force (hence acceleration) on the test body. We see large gravitational forces around each of the primaries, as well as large centrifugal forces at large distances. Near the Lagrange points, the forces are small. Right at the Lagrange points, the forces are exactly zero. If we were to place a test body at any point L_1 to L_5 exactly and perfectly at rest (in the rotating system), it would remain there indefinitely. However, the Lagrange points are unstable, because forces in the neighborhood point away from them. Any slight deviation (or perturbation) would most likely lead the test body to move away. The existence of Lagrange points is a result of balance between the attractive gravitational and the repulsive centrifugal forces. Without rotation, all Lagrange points but one would vanish. Only L_2, the saddle point, would remain. The effect of the potential surface is wholly dynamic.

For small α, all five Lagrange points lie close to the unit circle $r = 1$. While the locations of L_4 and L_5 can be found analytically, the locations of L_1, L_2 and L_3 need to be solved numerically. But for $\alpha \ll 1$, the following series expansion can be used (Exercise E4.11)

$$L_1 : \left(-1 - \frac{5\alpha}{12}, 0\right), \quad L_2, L_3 : \left(1 \mp \sqrt[3]{\frac{\alpha}{3}}, 0\right),$$

$$L_4, L_5 : \left(\frac{1}{2} - \alpha, \pm\frac{\sqrt{3}}{2}\right). \quad \text{(exact)} \tag{4.62}$$

Because the existence of the Lagrange points are due to the balance between gravity and rotation, it might appear curious that the locations (4.62) depend on α only, and not on other parameters such as the angular velocity ω. The reasons are two-fold. First, the mass parameter α as defined by Eq. (4.59) depends on the masses M_1 and M_2. Because ω also depends on the masses via Eq. (4.16), α is related to ω indirectly. Second, as discussed earlier, the RTB unit system (Table 4.3) is elastic. This hides some parameters which would be explicit if we used absolute units.

4.6.3 MOTION AROUND THE LAGRANGE POINTS

As discussed above, with the effective potential alone, none of the Lagrange points are stable. However, we must also consider the Coriolis effect (4.56), $-2\vec{\omega} \times \vec{v}$, when the test body is not stationary. The combination of the force from the effective potential and the Coriolis effect makes it possible to have stable periodic orbits in the neighborhood of the Lagrange points.[14]

Let us examine that possibility by focusing on the equipotential island surrounding L_4 in Figure 4.17 (bottom). All around the island, the force from the effective potential (the arrows) points away from L_4. Suppose the test body moves clockwise around the edges. On the outer edge, it will be moving to the right. Recall that $\vec{\omega}$ points out of the page, and the Coriolis effect, $-2\vec{\omega} \times \vec{v}$, points inward. It is thus possible for the Coriolis effect to cancel the outward force, such that the net acceleration is toward L_4. On

the inner edge, the velocity is reversed. So does the Coriolis effect. The net result is the same: it counter balances the inward force from the effective potential. Because no energy is added by the Coriolis effect, $(\vec{\omega} \times \vec{v}) \cdot \vec{v} = 0$, it should be possible for the test body to move about L_4 periodically with the right initial conditions. As it turns out, stable periodic motion does happen if $\alpha <$ 0.03852 [79].

The Trojan asteroids

Both the Earth-Moon and the Sun-Jupiter systems support stable periodic orbits. We start with an actual example for the motion of asteroids in the Sun-Jupiter system. Most asteroids in our solar system are in the so-called asteroid belt. But there are two groups of asteroids that revolve around the Sun at about the same radius and with the same period as Jupiter (Figure 4.18, left). The two groups are symmetrically located with respect to the Sun-Jupiter radial vector: one group 60° ahead and another 60° behind. They are known as the Trojan asteroids (or simply Trojans) [95]. Because their locations form equilateral triangles with the Sun and Jupiter just like the Lagrange points L_4 and L_5 (Figure 4.18, right), we suspect there is a connection.
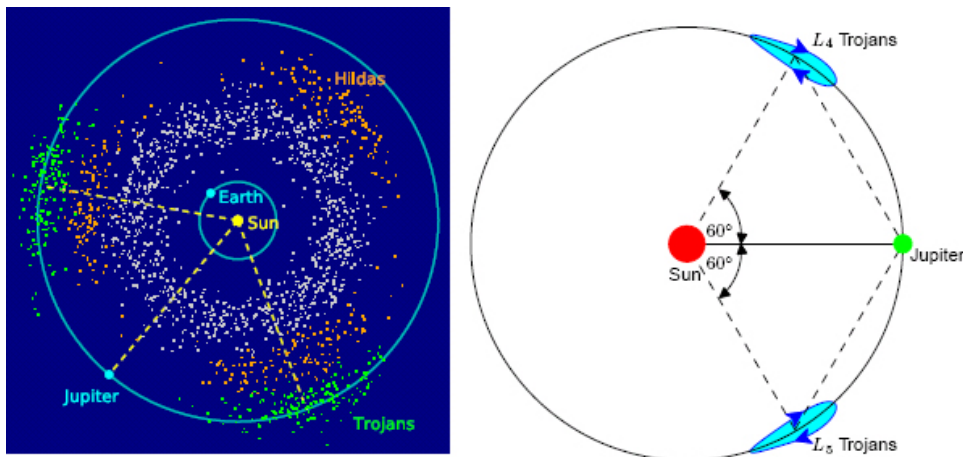
Figure 4.18: The Trojan asteroids and their connection to the
Lagrange points L_4 and L_5.

Though Jupiter is the most massive planet, the mass parameter for the Sun-Jupiter system is $\alpha = 0.0009542$ (see Table 4.1), still well within the limit for stable orbits. If the Trojans move in clockwise orbits around L_4 and L_5 as indicated in Figure 4.18 (right), stable orbits can be formed.

Program 4.7 (Section 4.D) simulates stable orbits around Lagrange points. Three orbits are shown in Figure 4.19. The first orbit (top left), also the smallest, corresponds to the initial condition as given in the program. It is a stable periodic orbit with a period of about 13 time units (Jupiter orbital periods), or about 150 years (using Table 4.1). The second orbit (top right) is produced with the initial condition

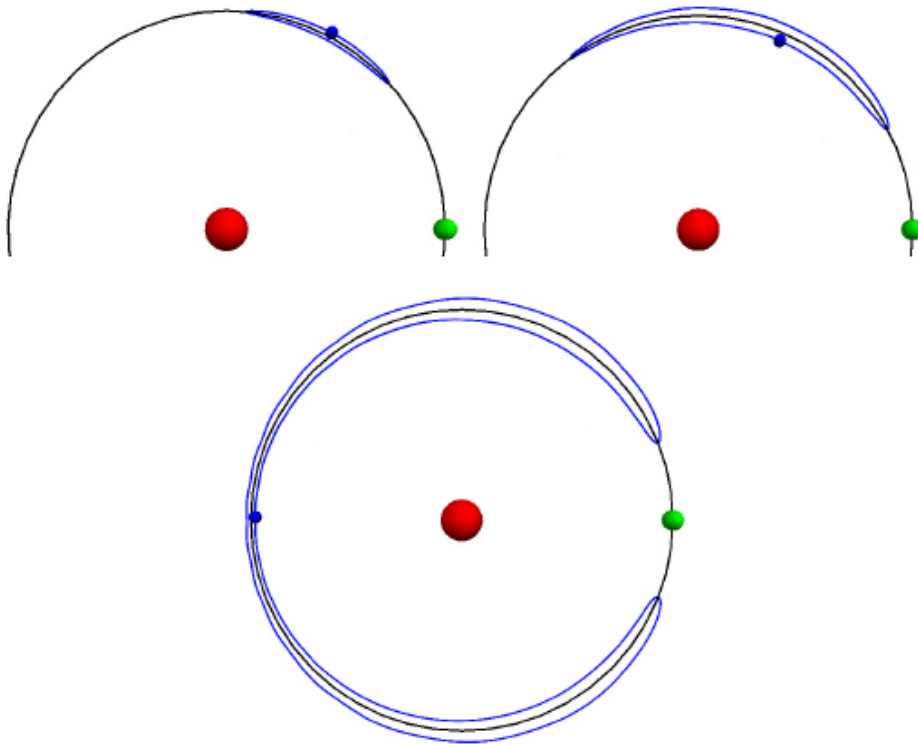$$\vec{r}, \ \vec{v} = [0.0, 1.035], \ [0.331, 0].$$

Figure 4.19: The orbits of Trojan astroids near the Lagrange point L_4 (top) and along the C-ridge (bottom).

It is approximately periodic and has a period of about 15 (~ 180 years). Both orbits are within the Trojan group. We have seen that Jupiter has a significant effect on the precession of Mercury because it is the most massive planet (Table 4.2). Here it shows its influence again, only this time its lasting effect is on the asteroids.

The largest orbit (Figure 4.19, bottom) goes all the way around the C-shaped ridge. Its initial condition is $\vec{r},\ \vec{v} = [0.524, 0.92], [0.49, -0.27]$. It appears to be periodic with a period of 33 (~ 390 years). The motion is very delicate, navigating along the equipotential lines. It does not seem to be connected to any asteroid motion at present.

As for the Earth-Moon system, although there is no evidence like the Trojan asteroids, numerical results show that stable periodic orbits do exist (see Ref. [73] and Exercise E4.12). The mass parameter is $\alpha = 0.0121$, much larger than the Sun-Jupiter mass parameter but still within the limit for the existence of stable periodic orbits. Therefore, it is theoretically possible to put a satellite or space station around L_4 or L_5. However, with the Sun being very close and perturbing the motion, the long term stability of these orbits is questionable.

All three orbits move clockwise (↻) around the potential ridge. Motion in the restricted three-body system is highly directional. Reversing the velocity of a stable periodic orbit in the restricted three-body problem results in unstable and nonperiodic motion (see Project S:P4.2). In contrast, reversing the velocity would have no effect on a Kepler orbit.

4.6.4 ORBITAL RESONANCE

As viewed in the space-fixed frame where the Sun is at rest, Jupiter and the Trojan asteroids have the same orbital period. We say their periods are commensurate. We can think of Jupiter's large effect on the Trojan asteroids as a result of the commensurability, or orbital resonance. If we consider the ratio of the orbital periods, we will find that orbital resonance exists if that ratio is a low rational number. In the case of Trojans and Jupiter, the ratio is 1/1. When orbits are in resonance, there is a periodic interaction ("kicks") between the orbits, which does not get averaged out and can lead to enhanced effects in the long run.

Trojans are not the only asteroids in resonance with Jupiter, though they are the most prominent. There is another group of asteroids known as the Hilda group which has a period ratio of roughly 2/3 to Jupiter. The Hildas are in three clusters forming a triangle about the Sun (see Project P4.7).

Here we discuss another interesting case of orbital resonance between Pluto and Neptune.[15] The mass parameter of the primaries, namely the Sun and Neptune, is very small, $\alpha = 0.0000515$. The ratio of Pluto's period to Neptune's is very close to 3/2 (see Table 4.1). For every three revolutions of Neptune around the Sun, Pluto makes two revolutions. Like the Trojan asteroids, the motion of Pluto is mainly determined by the Sun. However, considering the low commensurability ratio, we expect significant orbital resonance effect from the perturbation of Neptune. Also, because of Pluto's relatively high eccentricity (see Table 4.1), its orbit near the perihelion is inside Neptune's orbit (Figure 4.20, top). The question arises as to whether Pluto's orbit is stable, and if an "accidental" collision may occur between Pluto and Neptune that might knock Pluto out.
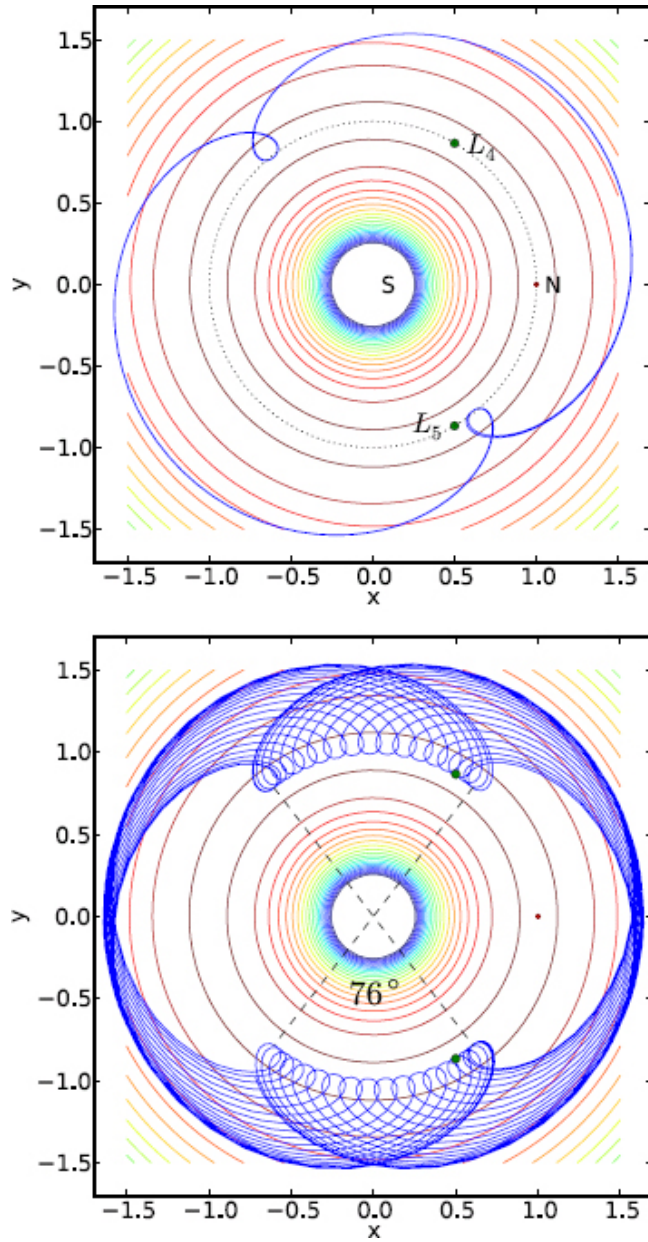
Figure 4.20: Orbits of Pluto in the Sun-Neptune system, $\alpha =$ 0.0000515, overlaid on the equipotential contours. Top: after several periods. The dotted line is the circular orbit of Neptune (viewed in a nonrotating frame). Bottom: after one half of libration (60 time units, or $\sim$ 10,000 years).

Let us investigate these questions by running our simulation (see details in Project S:P4.3). We can run Program 4.7 as is, with a different α and initial conditions, of course. But if we combine it

with Program 4.6, we can observe the orbit in relation to the potential contours and the Lagrange points, which will be helpful in understanding what is going on. All we need to do is to copy parts of the code that does the numerical integration `accel()` and `r3-body()`, integrate with RK4 as usual and record the positions x and y, and plot them on the same figure as the potential contours when the integration ends. The results are displayed in Figure 4.20. The initial condition used is [48]

$$\vec{r},\ \vec{v} = [0.607396, -0.777497],\ [0.680684, 0.531809].$$

This starting position is close to L_5.

The top figure shows the orbit for four Pluto periods. Starting near L_5 and the perihelion, we can see that Pluto's orbit dips inside Neptune's orbit by a small amount. It is about 0.013 length units when Pluto is closest to the Sun. The length unit is the distance between Neptune and the Sun, or 30 AU from Table 4.1. That means Pluto penetrates Neptune's orbit by about 0.39 AU, or nearly 40% the orbital radius of Earth. After one Pluto period, Pluto is again closest to the Sun but at the crossover loop diagonally opposite from the starting position near L_5. It takes two periods for Pluto to return to the starting position. If we look carefully at the crossover loop near L_5, we find that the orbit is not closed. The loop has rotated to the left. During this time, Neptune has revolved around the Sun three times.

Letting the program run for longer times, we obtain the complete orbit shown in Figure 4.20 (bottom). We see the continued pattern of the rotation of the orbit, or precession. The precession occurs over an angle of 76°, and remarkably, it reverses direction after that, and returns to the starting position to complete the cycle, which then repeats itself. The oscillatory back-and-forth pre-

cession is also called libration [17]. Half a libration cycle takes about 60 time units, or 10,000 years, making a full libration cycle approximately 20,000 years.

We can understand the reason of the precession as follows. When Pluto is at the crossover loop nearer to Neptune, say at the Lagrange point L_5, it is behind Neptune (taking counterclockwise direction as the forward). The forward pull from Neptune increases Pluto's energy slightly. As a result, its period also increases. The next time Pluto comes back (two revolutions around the Sun), it will be at a later time, and Neptune will be farther ahead, so the distance between Neptune and Pluto's perihelion (tip of the crossover loop) increases. Thus, Pluto appears to precess to the left, or rotate clockwise. In effect, Neptune repels Pluto, or more accurately the nearer loop, away. Eventually, Pluto gets ahead of Neptune at the perihelion because of the commensurate orbital periods (the other crossover loop near L_4). Then the backward pull from Neptune will decrease Pluto's energy and its period, which means that Neptune will be farther back, again effectively repelling the nearer loop. The cycles continues, giving rise to back and forth oscillations, or libration. The energy also oscillates with the libration period and an amplitude of about 0.4% (see Project S:P4.3).

Because of librational motion, at the closest approach to the Sun, Pluto gets no closer than 52° to Neptune in angular separation. Of course, they do line up radially when passing each other, for example, at (1.6, 0) in Figure 4.20 (bottom). But when that happens, they are radially far apart (~ 0.6 length units, or 18 AU). As a result, if Pluto and Neptune are close radially, they are far apart in angular separation, and vice versa. Therefore, a collision between them would not occur. In this case, orbital resonance

adds dynamic stability to Pluto, and Pluto will stay in its current orbit for the foreseeable future.

We may then ask, how did Pluto end up in its current orbit? It is possible that Pluto, at the time it was formed, happened to be right where it is now, locked in a stable configuration due to Neptune's perturbation. It is also possible, arguably more likely, that Pluto evolved into its current orbit over a long period of time (billions of years). There is evidence which suggests that Pluto's motion may be chaotic [88]. It may well be that we are seeing a Pluto that has been marching on its chaotic path all along.

Chapter summary

Starting with an animated motion of a planet, we briefly reviewed useful properties of central field motion and Kepler's orbits. We then discussed precession of Mercury due to corrections of the general theory of relativity. Because the effect is small, we introduced the Runge-Lenz vector as a sensitive indicator of precession. An accurate and area-preserving method, the leapfrog method with time-transformation, was implemented to simulate this process. We found an interesting effect that precession is oscillatory, not monotonic. Precession of classical origin due to other planets was also discussed within the framework of central field approximation.

We also described the radial velocity method for the detection of exoplanets. We presented two ways to model RV datasets, least square and visual fit, and applied them to observational datasets.

The planar three-body problem has been discussed, including Euler's collinear motion and several interesting choreographic solutions of gravitational systems. We studied the restricted three-

body problem, and analyzed in detail the structure of the Lagrange points. Motion around the Lagrange points was simulated. We presented the application to the Trojan asteroids in the Jupiter-Sun system. We also discussed orbital resonance and the stability of Pluto in the Neptune-Sun system.

Throughout this chapter, we have applied several visualization techniques using the 3D elements of VPython and Matplotlib. We have consistently presented the results as dynamic animations or informative graphs with these techniques. They are contained in blocks of sample codes that can be used as basic templates for integrating simple but effective visualization elements into other programs.

4.7 Exercises and Projects

EXERCISES

E4.1 With the coordinate transformation (4.4), show that (a) the equations of motion for two bodies interacting with the internal force (4.1) are given by Eq. (4.5); and (b)*

the two-body Hamiltonian $\mathcal{H} = \frac{1}{2}mv_1^2 +$

$\frac{1}{2}Mv_2^2 + V(|\vec{r}_1 - \vec{r}_2|)$ can be reduced to $\mathcal{H} = H_{CM} + H$,

where $H_{CM} = \frac{1}{2}(m + M)v_{CM}^2$ is the center of mass

motion, and H is the effective one-body Hamiltonian

given by $H = \frac{1}{2}\mu v^2 + V(r)$.

E4.2 Write a program to plot the effective potential of the

Earth-Sun system, Eq. (4.8). Use Table 4.1 to determine the angular momentum L. Use planetary motion units. Adjust the scales so the structure near the minimum is clearly displayed.

E4.3 (a) Modify Program 4.1 to simulate the motion of a planet under the force $F = -k/r^{2+\delta}$. Run the program using the same initial condition as in Figure 4.4, and with $\delta = 0.2$ and 0.5. Compare your results with Figure 4.4. Optionally, add an arrow object to the animation of the planet to indicate instantaneous velocity using the `arrow()` function of VPython.

(b) When $\delta \neq 0$, open orbits are real. However, numerical problems can cause unphysical artifacts, including open orbits that should be closed. Let us use the same program, but set $\delta = 0$ and the initial condition to $\vec{r}$,

$\vec{v} = [0.8, 0], [0, 3.5]$ (for exaggerated effect). First run it with $h = 0.001$, then with $h = 0.004$. Discuss your observations. Should we always "trust" the computer results?

E4.4 (a) Prove that the Runge-Lenz vector is a constant, i.e., $d\vec{A}/dt = 0$, in the inverse-square law of force (4.1).

(b) Consider an additional perturbation force, $\vec{f}$, such that the total force is $\vec{F}_{tot} = \vec{F} + \vec{f}$, where $\vec{F}$ is given by Eq. (4.1). Show that the Runge-Lenz vector obeys

$$\frac{d\vec{A}}{dt} = \vec{f} \times \vec{L} + \vec{p} \times (\vec{r} \times \vec{f}).$$

As usual, $\vec{p}$ and $\vec{L}$ are the momentum and angular momentum, respectively.

E4.5 Suppose that, instead of $\Omega(r) = 1/r$, we choose $\Omega(r) = 1/r^2$ as the time transformation. Rewrite the corresponding equations of the leapfrog method to Eqs. (4.29a) to (4.29f). Discuss the merits and shortcomings of this time transformation.

E4.6 (a) A total mass M is uniformly distributed over a ring of radius a. A point particle of mass m is placed at a distance r from the center and in the plane of the ring. Show that the potential energy between the particle and the ring is given by Eq. (4.30).

(b) Expand the potential for $r/a < 1$ and $r/a > 1$, keeping four terms in each case. Verify Eq. (4.31), and explain why only even terms appear. You may wish to use SymPy to expand the integrand,

```
In [1]:  from sympy import *
In [2]:   init_printing (pretty_print=True)  #
neater output
In [3]:  r, x = symbols('rho x')
In [4]:   f=series (1/sqrt(1+r*r−2*r*cos(x)), r,
0,7) # expand to 7th
```

Inspect f, then integrate $x \in [0, 2\pi]$ in SymPy (see Exercise E2.7).

(c) Compute the value M_p/a^3 for each planet in Table 4.2 using the parameters in Table 4.1. Plot precession

260

vs. M_p/a^3 on linear scales. Compare with Figure 4.8 and discuss the differences.

E4.7 (a) Obtain from Eq. (4.35) a relation similar to Eq. (4.36) but with $m \sin i$ given in Jupiter masses, T in days, M in solar masses, and V in m/s.

(b) Carry out the algebraic details to prove Eq. (4.39) from (4.37).

(c) Verify Eq. (4.79), and reduce it to Eq. (4.80) using conservation of angular momentum (4.6) and Kepler's third law (4.18).

E4.8 (a) Solve Eqs. (4.44) to (4.46) to obtain Euler's quintic equation (4.47).

(b) Write a program to plot Eq. (4.47) for a few sets of masses, say $[m_1, m_2, m_3] = [1, 1, 1]$ and $[1, 2, 3]$. How many positive roots are there in each case? Numerically solve for the positive root(s) in each case using either the bisection or Newton's methods discussed in Chapter 3.

E4.9 (a) The periodic orbit shown in Figure 4.14 is unstable but reversible. First find its period. One way to do so is to add a text box to Program 4.5 using the `label()` function of VPython, see Program 4.3 for an example. Output the time information, and record the period T when the three bodies return to their initial positions (4.51). To make it easier, you can pause the program by intercepting key presses with `scene.kb.keys` and `scene.kb.getkey()`. The system should start to come apart after roughly two periods. Wait until the bodies are well off their orbits, and reverse the velocities of all three bodies. This may be done as a conditional statement based on time or by detecting certain key strokes. You should observe that the bodies retrace

their paths and go back to the initial positions in the same amount of time.

(b) The figure-8 orbit shown in Figure 4.15 is stable. It means that if we start the motion from a slightly different initial condition, the orbit will remain close to the original orbit. Verify that this is true. Run Program 4.5 with the initial condition (4.52). Next change the initial condition by a small amount, say $x_1' \to x_1 + 0.1$, and run the program again. Finally, repeat with $y_3' \to 0.05$. Explain your observations.

E4.10 Make a table similar to Table 4.4 showing the units of various quantities in the restricted three-body units and in the absolute units for the Neptune-Sun system. You may need the data from Table 4.1.

E4.11 (a) Set up the equations for the Lagrange points where $\nabla V = 0$ from Eq. (4.61). Solve these equations to obtain the lowest order expansions for the collinear ones L_1, L_2 and L_3 and the exact expressions for L_4 and L_5. Show that they agree with Eq. (4.62).

(b) Write a program to solve the equations you obtained in part (a) for the collinear Lagrange points L_1, L_2 and L_3. Use the Earth-Moon system as an example. Compare the numerical results with the series expansions.

E4.12 The Earth-Moon system supports stable periodic orbits around the Lagrange points L_4 and L_5. Explore these orbits with the following initial conditions,

$$\vec{r}, \vec{v} = [0.49786, 0.88335], \ [0.16698, -0.09218];$$
$$[0.51286, 0.90933], \ [0.42897, -0.20988].$$

(a) For each orbit, find the period, and estimate its length along the arc and its width where it is widest. Convert them to absolute units. From your estimates, find the approximate area of the smaller orbit. If that area were to be colonized, what would be the population density if the entire population on Earth (7 billion) were to be moved? Compare it with the current population density.

(b) Check the stability of the orbits by varying the initial positions slightly, say to ±10%. You may also play around with the velocities. Refer to [73] for a general strategy to finding initial conditions of periodic orbits.

PROJECTS

P4.1 (a) The artificial precession as observed in Exercise E4.3 above could be eliminated by using a smaller h, which would lead to slower program speed and more round-off error. Though speed is not a problem in this case, it could be crucial in more complex situations. A better choice is to use a leapfrog method with time transformation. Repeat part (b) of Exercise E4.3 with `leapfrog_tt()`. Remember to initialize W_0 before entering the loop (see Program 4.3). Try different step sizes. You should see that, no matter how large h is, there is no precession. Explain why.

(b) Perform checks to make sure that the program works correctly and the oscillations in Figure 4.7 are not numerical artifacts. First, modify Program 4.3 to calculate and record the energy and angular momentum in the main loop. Ignore mass since it is an overall scale factor (see Eq. (4.86)), and use a large $\lambda \sim 10^{-2}$ for ex-

aggerated effects. It is best to define a separate function for energy. Run the program for $t < 10$, and plot the energy and the z-component (L_z) of angular momentum as a function of t. Make sure your potential is correct.

To get a better view of the fluctuations, plot the relative error $|E(t)-E_0|/|E_0|$ and $|L_z(t) - L_0|/L_0$, where E_0 and L_0 are the initial values computed before the loop. The output should be analogous to Figure 2.7. Are they conserved as expected?

(c) On the precession figure, use the zoom-in feature to expand the scale so the oscillations are clearly visible. Measure the amplitudes and periods. Repeat the simulation by doubling and halving h. Check whether the amplitudes or the periods change.

Perform the same checks by changing the sign of λ, i.e., enter the number -1.1×10^{-8} when prompted. Does the direction of precession change? Describe your observations and conclusions.

P4.2 Though Mercury has the largest precession due to general relativity, other planets also experience this effect to a lesser degree. For example, the parameter for Venus is $\lambda = 2.1 \times 10^{-8}$ AU^2 (see Eq. (4.19)). Change Program 4.3 accordingly to calculate the precession of Venus per century due to relativistic correction. Refer to Table 4.1 for the initial condition at the aphelion. Plot the results in a similar fashion to Figure 4.7. Note the amount of precession and the amplitude and the period of oscillations. Expand the horizontal scale to observe the oscillations. They should be sinusoidal, so count the period carefully (two crests, for example).

For Earth, the parameter is $\lambda = 2.9 \times 10^{-8}$ AU^2. Based on the above results for Venus and for Mercury (Figure

4.7), make a prediction about the amount of precession and the period of oscillations we would expect. Write down your predictions. Repeat the simulation for Earth. How do the results compare with your predictions? Briefly explain.

Compare the results for Venus and Earth with Mercury. The parameters λ for Venus and Earth are larger than for Mercury by a factor of 2 and 3, respectively. Why are the results not proportional to λ? Also discuss the differences in the amplitudes and periods of the oscillations. Explain why the oscillations are sharper (more sinusoidal) for Venus and Earth than for Mercury. Offer any other observations you may have.

P4.3 As we have discussed, the main contribution to the precession of Mercury is from other planets. This amount must be subtracted from the total precession in order to find the contribution from general relativity.

(a) Calculate the precession due to each planet in the central field approximation using the force (4.32) with the parameters from Table 4.1. Reproduce the results of Table 4.2. Make sure your results are converged with respect to the step size h. One way to check for convergence is to run the simulation with $h = 0.002$, for example, and then reduce it by half to $h = 0.001$. The results should agree to with certain error you set, say 1% for instance.

(b) Observe that the net precession is counterclockwise, which is in the same direction for the relativistic case. However, the force (4.32) in the central field approximation is repulsive, while the relativistic effect (4.19) is

attractive. Investigate why forces of opposite signs produce precession in the same direction, because only then can we add them up. Do this by switching on the animation, and watch carefully how the Runge-Lenz vector swings in relations to where Mercury is in the orbit. Explain your observations. The results in Exercise E4.4 may help.

P4.4 If the force on a planet does not obey the inverse-square law, even for only a part of the trajectory, the orbit is open and will precess. Let us investigate this effect by modeling the Sun as a uniform sphere of radius R instead of a point source. If the planet is outside the sphere, the force is the usual gravity. But inside the sphere, the force will change to a form different from $1/r^2$.

(a) Show that the gravitational potential and force between a point mass m and a uniform sphere of radius R and mass M are given by

$$V = \begin{cases} \frac{GMm}{2R}(r^2/R^2 - 3), & r \le R, \\ -GMm/r, & r > R, \end{cases} \quad ; \quad F = \begin{cases} -GMmr/R^3, & r \le R, \\ -GMm/r^2, & r > R. \end{cases}$$

(b) Modify the force in `mercury()` in Program 4.3 according to the expression above. Set $R = 0.3$, leave everything else the same and run the program. Does the Runge-Lenz vector move? Why?

(c) Compute the perihelion distance of Mercury r_{min} from Eq. (4.14). Set R just above r_{min} and run the program. Does the Runge-Lenz vector move now? Experiment with R values slightly above r_{min} and determine the value that produces the same precession (43″) as the

relativistic effect. Discuss your findings. Does it surprise you?

(d) Let us have fun making a Runge-Lenz clock. With the same initial conditions, find the value of R such that the hour hand (Runge-Lenz vector) ticks exactly 12 times in a full circle. Alternatively, you can achieve the same effect by varying the eccentricity of the orbit while keeping R fixed. Set $R = 0.2$ and vary the initial vertical velocity v_y. Find the value of v_y when the hour hand moves 12 times in one revolution. What would be the eccentricity of the corresponding Kepler orbit?

P4.5 Simulate the wobbling motion of a star-planet system with animation of the two-body problem, and discover a mystery planet.

(a) With Program 4.1 as a basic template, add the Sun as a movable object. Do not include the Sun as a dynamic body in the numerical integration. We still assume an effective one-body problem. We only need to obtain the positions of the Sun and the planet ($\vec{r}_2$ and $\vec{r}_1$ in Figure 4.2) in the center of mass (CM) from the relative planet position stored in r, which is already available in line 26 of Program 4.1.

Assume $\vec{R} = 0$, and solve Eq. (4.4) for $\vec{r}_1$ and $\vec{r}_2$ in the CM (see Exercise E4.1). Animate the Sun and the planet by updating their positions in the CM. Make sure the modified code runs correctly for the Sun-Earth system.

(b) Note that due to the large Sun-Earth mass ratio, the Sun barely moves. Now change the mass of the planet,

267

say 0.4 solar mass. Run the code again and observe the visible wobbles. You may also need to change the initial conditions depending on the mass ratio.

(c) Plot the radial velocity of the Sun, v_{*z}, in the actual Sun-Earth system similar to Figure 4.9, assuming edge-on view. Discuss the results, such as the shape and magnitude of the curve.

(d) Make your own mystery planet by assuming certain mass, say anywhere from 1 to 10,000 Earth mass, and initial condition (For added challenge, choose your initial condition such that the orbit has a high eccentricity). Run your program and generate an RV dataset accordingly. Exchange your dataset with another classmate (or team), and try to visually fit the dataset according to Eq. (4.34). Calculate the planet mass from your fit. Compare your value with the value your classmate used to generate the dataset. Do they match? Why or why not?

P4.6 A subset of the radial velocity dataset for HD 3651 plotted in Figure 4.11 is given in the table below.

Table 4.5: Radial velocity (m/s) vs. time (day) for HD 3651.

t	6.3	6.8	11.3	12.5	23.5	25.7	27.6	29.4
v_{*z}	13.9	7.4	9.6	2.2	5.7	1.3	0.3	−5.4
29.4	32.3	34.7	36.2	36.7	38.3	38.6	39.1	39.7
1.2	−3.3	1.5	10.9	−3.1	0.4	−5.9	−1.2	−4.5
40.6	42.2	43.6	50.8	54.1	55.5	56.3	56.8	61.3
−12.1	−5.2	−12.8	−18.8	−15.6	−26.4	−0.5	−1.2	5.9

The star has 0.8 solar masses, and the wobbling period is assumed to be 62 days. The data points were observations from Keck Observatory spread over the years. They are converted (phased) to fall within one period.

(a) Follow the procedure outlined in Section 4.4.2 to generate a visual fit similar to the curve shown in Figure 4.11. Use Program 4.4, but turn off least square fit by commenting out line 38 and set time shift to 0 (line 27). Vary V, e, ω, and C but keep T and time shift fixed. As the parameters are being adjusted, watch the effect of each parameter, and try to isolate them to achieve a better match. Once the curve takes the shape of the dataset, adjust the parameters separately and in small steps to visually fine tune the fit, focus on the shape and ignore an apparent horizontal shift for now. Describe the effect of each parameter.

(b) You should see that the fitted curve is shifted from the data, e.g., the dip should be at a different location, even though the overall shapes match very well. If you would just shift your curve left or right, it would look quite like the fit in Figure 4.11. Explain the physical reason for the shift. Adjust the time shift to get the best fit. Write down the parameters, and calculate the lower limit of the exoplanet's mass, in Earth and in Saturn masses. Compare with observed values [30].

(c) Run Program 4.4 in full to obtain a least square fit. Compare the best results between visual and automatic fits. Optionally, switch to the reduced dataset of Table

4.5 or the self-generated dataset from Project P4.5, and discuss any difference in the fit.

P4.7 Similar to the Trojans, there is another group of asteroids known as the Hildas whose motion around the Sun is heavily influenced by Jupiter. The Hildas (Figure 4.18) do not move around the Lagrange points L_4 or L_5. Rather, like Pluto, they are influenced by orbital resonance with Jupiter. The Hildas have a period ratio of roughly 2/3 to Jupiter.

(a) Explore the motion of the Hilda asteroids with the following initial condition

$$\vec{r}, \ \vec{v} = [0.64772, 0], \ [0, 4.2902].$$

Find the period of motion. First plot the orbit in the rotating frame, then plot it in the nonrotating space frame where the Sun is at rest. You would need to make a coordinate transformation from the rotating frame to the space frame using the rotation matrix (4.74). Compare the orbits between the frames.

(b) Normally the orbit would also have a librational motion. The following initial condition will show substantial libration,

$$\vec{r}, \ \vec{v} = [0.78997, 0], \ [-.87208, 1.9879].$$

Run the program for 10 time units. Note the angles of libration. Modify the initial condition slightly, and describe how the librational motion changes.

4.A Rotating frames and rate of change of vectors

Consider the coordinate systems of a space-fixed reference frame and a reference frame rotating at constant angular velocity ω about the common z axis. Let the coordinates be (x, y) and (x', y') in these reference frames, respectively, as shown in Figure 4.21. We assume that at $t = 0$ the axes (e.g., x and x') coincide, so the angle of rotation is $\theta = \omega t$ at later times.
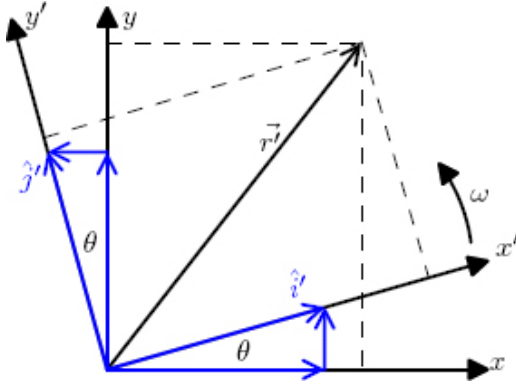

Figure 4.21: The space reference frame and the rotating reference frame.

From Figure 4.21, the relationship between the primed unit vectors in the rotating frame and the unprimed ones in the space frame can be written as

$$\hat{i}' = \cos\theta\hat{i} + \sin\theta\hat{j}, \ \hat{j}' = -\sin\theta\hat{i} + \cos\theta\hat{j}. \tag{4.63}$$

Because of rotation, the primed unit vectors $\hat{i}'$ and $\hat{j}'$ change with time. By differentiating the above equation with the help of $d\theta/dt = \omega$, the rate of change is

$$\frac{d\hat{i}'}{dt} = \omega\hat{j}', \quad \frac{d\hat{j}'}{dt} = -\omega\hat{i}'. \tag{4.64}$$

Now, let us consider a position vector $\vec{r'}$ in the rotating frame with components x' and y'. The same vector in the space frame will be denoted as $\vec{r}$. Since we are talking about the same vector, only different notations, we have

$$\vec{r} = \vec{r'} = x'\hat{i'} + y'\hat{j'}.$$

To find the rate of change, we differentiate both sides simultaneously

$$\frac{d\vec{r}}{dt} = \frac{dx'}{dt}\hat{i'} + \frac{dy'}{dt}\hat{j'} + x'\frac{d\hat{i'}}{dt} + y'\frac{d\hat{j'}}{dt}. \tag{4.65}$$

The left side refers to the rate of change in the space frame. The first two terms on the right hand side are the rate of change due to the change of coordinates x' and y' as measured in the rotating frame, and the last two terms are due to the change of the primed unit vectors. Let us denote this more clearly by rewriting Eq. (4.65) as

$$\left(\frac{d\vec{r}}{dt}\right)_{\text{space}} = \left(\frac{d\vec{r'}}{dt}\right)_{\text{rot}} + \omega(x'\hat{j'} - y'\hat{i'}), \tag{4.66}$$

where $(d\vec{r}/dt)$space is the velocity in the space frame $(\vec{v})$, and $(d\vec{r'}/dt)_{\text{rot}} = dx'/dt\,\hat{i'} + dy'/dt\,\hat{j'}$ the velocity in the rotating frame $(\vec{v'})$. We have also used Eq. (4.64). We can write the second term in Eq. (4.66) as a vector product by introducing $\vec{\omega} = \omega\hat{k'}$. This leads to a more compact form

$$\left(\frac{d\vec{r}}{dt}\right)_{\text{space}} = \left(\frac{d\vec{r'}}{dt}\right)_{\text{rot}} + \vec{\omega} \times \vec{r'}, \tag{4.67}$$

or equivalently and in a more familiar form,

$$\vec{v} = \vec{v'} + \vec{\omega} \times \vec{r'}. \tag{4.68}$$

272

In deriving Eq. (4.67), we have used the position vector as an example and restricted the angular velocity in the z direction. It is actually valid for an arbitrary vector and angular velocity. The more general rule is

$$\left(\frac{d}{dt}\right)_{\text{space}} = \left(\frac{d}{dt}\right)_{\text{rot}} + \vec{\omega} \times . \tag{4.69}$$

Equation (4.69) should be read this way: the derivative of a vector in the space frame is equal to the derivative of the (same) vector in the rotating frame plus the cross product of the angular velocity and the vector. This is a very useful rule.

For instance, suppose we wish to calculate acceleration from Eq. (4.68). We have

$$\vec{a} = \frac{d\vec{v}}{dt} = \left(\frac{d\vec{v'}}{dt}\right)_{\text{space}} + \left(\frac{d(\vec{\omega} \times \vec{r'})}{dt}\right)_{\text{space}}. \tag{4.70}$$

Applying Eq. (4.69) to each of the two terms on the right hand side, we have

$$\left(\frac{d\vec{v'}}{dt}\right)_{\text{space}} = \left(\frac{d\vec{v'}}{dt}\right)_{\text{rot}} + \vec{\omega} \times \vec{v'},$$

$$\left(\frac{d(\vec{\omega} \times \vec{r'})}{dt}\right)_{\text{space}} = \left(\frac{d(\vec{\omega} \times \vec{r'})}{dt}\right)_{\text{rot}} + \vec{\omega} \times (\vec{\omega} \times \vec{r'})$$

$$= \vec{\omega} \times \vec{v'} + \vec{\omega} \times (\vec{\omega} \times \vec{r'}).$$

We have used $d\vec{\omega}/dt = 0$ and $(d\vec{r'}/dt)_{\text{rot}} = \vec{v'}$. Substituting them into Eq. (4.70) and using the fact that $(d\vec{v'}/dt)_{\text{rot}} = \vec{a'}$ is the acceleration in the rotating frame, we obtain the relationship for $\vec{a}$ and $\vec{a'}$ as

$$\vec{a} = \vec{a'} + 2\vec{\omega} \times \vec{v'} + \vec{\omega} \times (\vec{\omega} \times \vec{r'}). \tag{4.71}$$

Let $\vec{F}$ be the net force on a particle of mass m. The equation of motion in the space frame (inertial) is $\vec{F} = m\vec{a}$. Expanded in the rotating frame (noninertial), it becomes

$$\vec{F} - 2m\vec{\omega} \times \vec{v'} - m\vec{\omega} \times (\vec{\omega} \times \vec{r'}) = m\vec{a'}. \qquad (4.72)$$

Besides the actual force, this equation has two additional force-like terms. The velocity-dependent term, $-2m\vec{\omega} \times \vec{v'}$, is the Coriolis effect, and the other term, $-m\vec{\omega} \times (\vec{\omega} \times \vec{r'})$, is the centrifugal force.

4.B Rotation matrices

Sometimes it is necessary to transform the coordinates from the rotating frame to the space frame. This can be done using $(x, y) = (\vec{r'} \cdot \hat{i}, \vec{r'} \cdot \hat{j})$ and Eq. (4.63). The results may be expressed conveniently in the form of a rotation matrix as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}. \qquad (4.73)$$

The inverse transformation is

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \qquad (4.74)$$

These rotation matrices work on any vectors, not just position vectors.

Denoting the rotation matrix by A, we can write Eq. (4.74) as

$$(\vec{r})' = A\vec{r}, \quad A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}. \qquad (4.75)$$

Note that we are careful with the notation $(\vec{r})'$ rather than $\vec{r'}$. The reason is that the transformation has dual interpretations. Here, we

are following the interpretation that the operator A transforms the components of the vector in the space frame (unprimed) into the components of the same vector in the rotating frame (primed). In other words, A is interpreted to act on the coordinate system, the vector itself is unchanged, and only the components are transformed from the unprimed frame to the primed frame which is rotated counterclockwise by an angle θ.

The second interpretation is that the operator A acts on the vector $\vec{r}$, transforming it into a new vector $\vec{r}'$, both in the same reference frame. The new vector represents a clockwise rotation of the original vector by an angle θ. We would represent this operation as $\vec{r}' = A\vec{r}$. Both interpretations involve identical mathematical operations. In the current context, we follow the first interpretation for convenience, i.e., the operator A relates the components of the same vector between rotated coordinates.

In three dimensions, the rotation matrices about the x, y, and z axes follow straightforwardly from Eq. (4.74),

$$
A_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}, \quad A_y = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix},
$$

$$
A_z = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4.76}
$$

The inverse transformation can be obtained by changing the sign of the angle $\theta \to -\theta$, because a rotation by an angle θ is exactly canceled by a reverse rotation by $-\theta$, resulting in no change (the identity transformation). If A_i^{-1} represents the inverse transformation, we can write $A_i^{-1}(\theta) = A_i(-\theta)$, and $A_i^{-1}A_i = A_iA_i^{-1} = \mathbf{1}$, $i = x, y$, or z.

Successive rotations can be represented by a product of rotation matrices. For example, a rotation of α about the x-axis followed by a second rotation of β about the z-axis can be written as

$$A = A_z(\beta) A_x(\alpha). \tag{4.77}$$

The order is important for finite rotations, and the final matrix is built from right to left in the order of rotations. The inverse matrix is $A^{-1} = A_x^{-1}(\alpha) A_z^{-1}(\beta) = A_x(-\alpha) A_z(-\beta)$.

4.C Radial velocity transformation

Our goal is to transform the velocity in the orbital plane from the center of mass frame to the space-fixed frame where observation is made. As Figure 4.3 shows, the relative position in the orbital plane is

$$x = r \cos\theta, \quad y = r \sin\theta. \tag{4.78}$$

Using Eq. (4.13), the velocity components can be obtained as,

$$v_x = \dot{x} = -r^2 \dot{\theta} \frac{\sin\theta}{a(1-e^2)}, \quad v_y = \dot{y} = r^2 \dot{\theta} \frac{\cos\theta - e}{a(1-e^2)}. \tag{4.79}$$

With the help of Eqs. (4.6) and (4.12), being careful to replace m by reduced mass μ, and using the exact form of Kepler's third law (4.18), we can reduce the orbital (relative) velocity to

$$v_x = -\frac{2\pi a}{T\sqrt{1-e^2}} \sin\theta, \quad v_y = \frac{2\pi a}{T\sqrt{1-e^2}} (\cos\theta - e), \tag{4.80}$$

where, as before, T is the period, a the semimajor axis, and e the eccentricity.

Let $\vec{v}'$ and $\vec{v}'_*$ denote the velocities of the planet and the star in the center of mass, respectively. Then $\vec{v}' - \vec{v}'_* = \vec{v}$, $M\vec{v}'_* + m\vec{v}' = 0$.

Solving for $\vec{v}'_*$, we have

$$\vec{v}'_* = -\frac{m}{m+M}\,\vec{v}.\qquad(4.81)$$

It is identical to Eq. (4.33) except for the primed notation which will be convenient for the transformation below. Substituting Eq. (4.80) into (4.81), we obtain the wobbling velocity of the star in the center of mass frame

$$v'_{*x} = \frac{m}{m+M}\frac{2\pi a}{T\sqrt{1-e^2}}\sin\theta,$$
$$v'_{*y} = -\frac{m}{m+M}\frac{2\pi a}{T\sqrt{1-e^2}}\,(\cos\theta - e).\qquad(4.82)$$

Next, we need to transform the wobbling velocity from the center of mass frame (primed) to the space-fixed frame (un-primed). This can be done conveniently by a rotation matrix. As Figure 4.10 illustrates, the orbit orientation is the result of two successive rotations. Imagine the aphelion (x'-axis) and the normal (the $\vec{L}$ vector, or z'-axis) of the orbital plane are initially aligned with X and Z, respectively. The final orientation of the orbit can be obtained as: a first rotation by an angle i about the x'-axis, and a second rotation by an angle ω about the z'-axis. The rotation matrix follows from Eq. (4.77), $A = A_z(\omega)A_x(i)$. Recall that A will transform a vector from unprimed frame (space) to the primed frame (center of mass). What we need is the inverse transformation from the center of mass frame to the space frame, i.e., $A^{-1} = A_x^{-1}(i)A_z^{-1}(\omega) = A_x(-i)A_z(-\omega)$. Using the rotation matrices (4.76), we have the inverse transformation

$$A^{-1} = A_x(-i)A_z(-\omega) = \begin{bmatrix} \cos\omega & -\sin\omega & 0 \\ \sin\omega\cos i & \cos\omega\cos i & -\sin i \\ \sin\omega\sin i & \cos\omega\sin i & \cos i \end{bmatrix}.\qquad(4.83)$$

Transforming the velocity vector in Eq. (4.82) by A^{-1}, we obtain the velocity of the star in the space-fixed frame

$$\vec{v}_* = A^{-1}\vec{v}'_* = \frac{m}{m+M}\frac{2\pi a}{T\sqrt{1-e^2}}\begin{bmatrix} \sin(\omega+\theta)-e\sin\omega \\ -\cos i\,[\cos(\omega+\theta)-e\cos\omega] \\ -\sin i\,[\cos(\omega+\theta)-e\cos\omega] \end{bmatrix}. \quad (4.84)$$

This is the desired result. We pick up a Z component in the process because the orbit is tilted due to the inclination angle. The inclination angle $i = 90°$ corresponds to the edge-on view, and $i = 0°$ the face-on view.

We can make two extensions to the result. First, if the center of mass is moving with velocity $\vec{V}_{CM}$, we must add it to $\vec{v}_*$. Second, if the star has multiple planets, and if planet-planet interaction is negligible,[16] then the total wobbling velocity is a vector sum of contributions from all the planets. For example, for the Z component v_{*z} of a star with N planets, we have from Eq. (4.84)

$$v_{*z} = -\frac{1}{M_T}\sum_{j=1}^{N}\frac{2\pi a_j m_j \sin i_j}{T_j\sqrt{1-e_j^2}}[\cos(\omega_j+\theta_j)-e_j\cos\omega_j] + C, \quad (4.85)$$

where $M_T = M + \Sigma_j\, m_j$ is the total mass, and $C = V_{CM,z}$ is the velocity of the center of mass along Z. Equation (4.85) can be used to fit the Doppler velocity profile to obtain the orbital parameters.

4.D Program listings and descriptions

Program listing 4.3: Precession of Mercury (`mercury.py`)

```
    import ode, numpy as np              # get ODE solvers,
    numpy
2   import visual as vp                  # get VPython
    modules for animation
```