6 The Two-Body Problem and Kepler's Laws

Much of what we will discuss in this class involves orbiting objects: a moon around a planet, or several planets around their star, thousands of stars in a cluster, or billions in a galaxy. The first consideration of all these issues involves the so-called **two-body problem**, when two objects orbit each other due to their gravitational attraction. Observations of such systems are especially powerful because (as we will see) certain quantities — e.g. masses, radii, orbital periods — can be measured extremely precisely. For example, binary pulsars (two dead **neutron stars** orbiting each other) may have their masses measured with a precision of $10^{-3} M_\odot$ ($\sim 0.1\%$).

Our goal in the discussion below is to work through the gravitational two-body problem with an eye on features that are observationally testable, and on features specific to the $1/r^2$ nature of gravity. In the real world many "details" push us away from exact $1/r^2$ — e.g. physical sizes, non-spherical shapes, and general relativity.

6.1 *Kepler's Laws*

To set the stage, recall from introductory physics Kepler's three laws of orbital motion (not to be confused with Asimov's Three Laws of Robotics):

1. **Kepler's First Law**: objects orbit along elliptical trajectories, as shown in Fig. 4. The most relevant quantity here is the **semimajor axis** a, which will come up again and again. The **eccentricity** e is also important: if $e = 0$, the orbit is circular and $r(\phi) = a$ always. However, most orbits are at least slightly eccentric and so in general

$$r(\phi) = \frac{a\left(1 - e^2\right)}{1 + e \cos\phi}. \quad (37)$$

2. **Kepler's Second Law**: In a given time interval dt an object's orbit always sweeps out the same area dA across its orbital ellipse. That is, dA/dt is

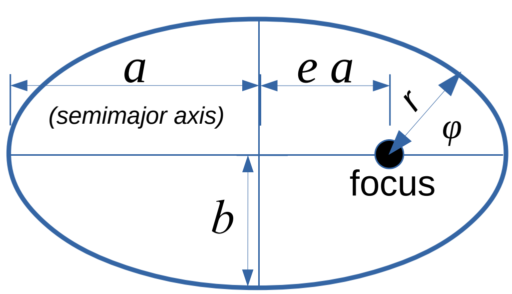

Figure 4: Elliptical trajectory of a two-body orbiting system. Important quantities are labeled, including the semimajor axis a, the orbital eccentricity e, and the polar coordinates r and ϕ as measured from the ellipse's focus.

constant; this will turn out to be

$$\frac{dA}{dt} = \frac{1}{2} r^2 \frac{d\phi}{dt}. \quad (38)$$

3. **Kepler's Third Law**: The most useful of all Kepler's Laws in many situations we will encounter. This relates a, the orbital period P, and the total mass M_{tot} of the orbiting bodies, as a^3/P^2 equal to a constant. You may have seen this before as some variation on:

$$P^2 = \left(\frac{4\pi^2}{GM} \right) a^3 \quad (39)$$

6.2 *Deriving Kepler's Laws*

Our goal in what follows is to rigorously derive Kepler's Laws and so understand how we can describe the positions and velocities of objects orbiting in a two-body arrangement. Initially this may seem daunting: since in 3D space each object has three position coordinates and three velocity components, we have twelve degrees of freedom that might have to be explained. We need to reduce this number to make things tractable!

Kepler's 2nd Law

First, we can reduce the dimensionality by half by recognizing that both objects will orbit around their common center of mass (see Fig. 5). Whatever arbitrary origin we choose for our coordinate system (such that our objects have 3D positions $\vec{r_1}$ and $\vec{r_2}$), the position $\vec{R}$ of the center of mass in this reference frame will be

$$\vec{R} = \frac{m_1 \vec{r_1} + m_2 \vec{r_2}}{m_1 + m_2}. \quad (40)$$

position vector from one to the other will always be

$$\vec{r} = \vec{r_2} - \vec{r_1}. \quad (41)$$

Remember from introductory physics that if there are no external forces on the system, then the center of mass experiences no accelerations: $\ddot{\vec{R}} = 0$. And since physics is the same in all inertial reference frames, we have freedom to choose the 'easy' inertial reference frame in which $\dot{\vec{R}} = 0$ too – so the position of the center of mass never changes. Since physics is also the same in all locations, we can again pick the 'easy' reference frame in which $\vec{R} = 0$ too.

So the center of mass is at the origin now, and since it isn't moving and isn't accelerating it will always be at the origin. From Eq. 40, this means that

$$m_1 \vec{r_1} + m_2 \vec{r_2} = 0. \quad (42)$$

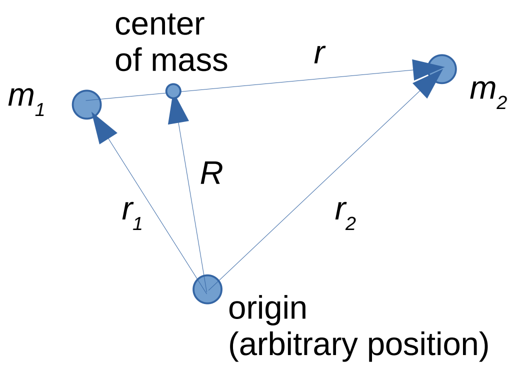

Figure 5: Coordinate system for two masses m_1 and m_2 separated by a distance r. Both objects will orbit around their common center-of-mass; in this case, $m_1 > m_2$ and so the center of mass is closer to m_1.

From Fig. 5 the position of one object relative to the other is just

$$\vec{r} = \vec{r_2} - \vec{r_1}. \tag{43}$$

Combining Eqs. 42 and 43 shows that both positions $\vec{r_1}$ and $\vec{r_2}$ are proportional to each other and also to the difference vector $\vec{r}$:

$$\vec{r_2} = \frac{\vec{r}}{1 + m_2/m_1} \tag{44}$$

and

$$\vec{r_1} = -\frac{m_2}{m_1}\vec{r_2}. \tag{45}$$

This means that in terms of the number of unknown quantities, we've reduced the two-body problem to an effective one-body problem with 'only' one set of unknown position and velocity coordinates.

We can go further: since gravity is a radially-acting force, $\vec{F_G} \parallel \vec{r}$. This means that torque will be zero (since $\vec{\tau} = \vec{r} \times \vec{F}$), and thus the **angular momentum** $\vec{L} = \vec{r} \times \vec{p}$ of the system will be constant in magnitude and direction: thus the orbit must be constrained to a (2D) plane, and so we're justified in using just r and ϕ to describe the orbit.

The area dA swept out by the orbit in a time interval dt is given by

$$\begin{aligned} dA = & \frac{1}{2}\left|\vec{r} \times \vec{dr}\right| \\ = & \frac{1}{2}\left|\vec{r} \times \vec{v}dt\right| \\ = & \frac{dt}{2m}\left|\vec{r} \times m\vec{v}\right| \\ = & \frac{1}{2}\frac{L}{m}dt = \text{constant}. \end{aligned}$$

So an equal area is swept out in any equal time interval – that's Kepler's Second Law.

One open question in the derivation immediately above is what mass we should use for m — since in fact we have two masses orbiting each other, m_1 and m_2. It turns out that this should be written in terms of the so-called **reduced mass** μ of the two-body system, where

$$\mu \equiv \frac{m_1 m_2}{m_1 + m_2} \tag{46}$$

and so

$$\frac{dA}{dt} = \frac{1}{2}\frac{L}{\mu} = \text{constant}. \tag{47}$$

Kepler's 3rd Law

Deriving this law involves rather more than we want to deal with in this class, but we can get close with some fairly basic approximations. Consider an object of mass m_2 in a simple, circular orbit with around a larger mass m_1 (we're neglecting that actually both objects orbit around their common center of mass). The orbit has semimajor axis a, which is also the (constant) separation r between the objects.

By Newton's first law, we must have$F_{\text{external}} = ma_r$ where (for uniform circular motion) a_r is the centripetal acceleration (not to be confused with a, the semimajor axis). So for a gravitational orbit, we have

$$\frac{Gm_1}{m_2}a^2 = \frac{m_2 v_2^2}{a^2}. \tag{48}$$

The orbital velocity is related to the orbital period by $v = 2\pi a/P$. Plugging that in for v and rearranging, we find the almost-correct form:

$$P^2 \frac{Gm_1}{4\pi^2} = a^3.$$

This is not quite right because we neglected the movement of the larger m_1 around the common center of mass. This isn't a big effect if m_2 is relatively small; a full derivation would show us that we should be using the total mass of the system, $M_{\text{tot}} = m_1 + m_2$. This gives us the "physicists's version" of Kepler's Third Law:

$$P^2 \frac{GM_{\text{tot}}}{4\pi^2} = a^3. \tag{49}$$

This isn't too bad, but we can make it even easier to use in many situations. We know that the Earth takes one year to orbit the Sun at a distance of 1 au, and that the total mass of the Earth-Sun system is $M_\odot + M_\oplus \approx M_\odot$. This means we can turn Eq. 49 into a set of ratios for the "astronomer's version" of

Kepler's Third Law:

$$\left(\frac{P}{1\ \text{yr}}\right)^2 \left(\frac{M_{\text{tot}}}{M_\odot}\right) = \left(\frac{a}{1\ \text{au}}\right)^3 \tag{50}$$

Either Eq. 49 or Eq. 50 let astronomers calculate an object's mass merely by observing its orbital motion (i.e., a and P) — and either expression will give the correct answer when applied correctly.

6.3 *Introducing Energy Diagrams*

Energy Diagrams, or potential energy plots, are useful tools that we will use in a variety of situations in this class. Their main benefit is to tell us when various parameters are permitted (or prohibited) from taking on certain values due to the energy considerations of the system. These are most easily employed when we just have a single variable (e.g., energy is only a function of x), though they can be used in other situations too.

The steps are fairly straightforward:

1. Write down an expression for the potential energy U in your system.
2. Plot U vs. your dependent spatial variable (e.g., $U(x)$ vs. x).
3. Assume some amount of total mechanical energy (i.e., the sum of U and kinetic energy K).
4. Overplot a horizontal line indicating the total E_{mech} of the system.
5. Consider your plot and gain insight.

The insights gained should be the following:

- Forbidden regions: for any areas on the plot where $U > E_{\text{mech}}$, this would imply $K < 0$ (which is impossible). Thus your system will never be located at these regions!
- Permitted regions: the opposite of forbidden regions, wherever $U \leq E_{\text{mech}}$. The system could potentially be located anywhere in these regions. There could be one or multiple permitted regions, and they need not all be contiguous.
- Kinetic Energy: Since $K = E_{\text{mech}} - U$, the difference between the horizontal E_{mech} and more complicated U curves gives the amount of kinetic energy at that position – and thus also how fast the system is moving.
- Turning points: Wherever $U = E_{\text{mech}}$ exactly, the object has zero velocity. These are the points where the object would turn around if it had been in the adjacent Permitted Region.

Fig. 6 gives an example for a simple harmonic oscillator with $U(x) = 1/2kx^2$, with some nonzero amount of mechanical energy E_{mech}. At the indicated position x_1, there is some nonzero kinetic energy: so if the oscillator

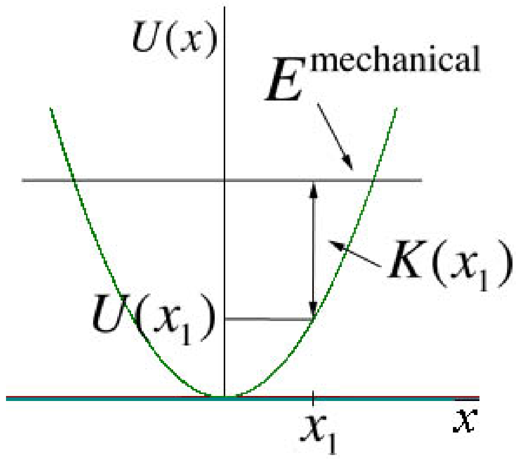

Figure 6: Energy diagram for the case of a simple harmonic oscillator, $U(x) = 1/2kx^2$, with some nonzero amount of mechanical energy E_{mech}.

were located at x_1, it would be moving – but more slowly than if it were at $x = 0$, since $E_{\text{mech}} - U(x)$ is less at x_1 than at $x = 0$. Regions with very large x are forbidden; this should make intuitive sense, since if you start a spring bouncing you don't expect it to stretch out to infinity all of a sudden.

6.4 *Energy of the Two-Body System*

In any system, the total mechanical energy is just $E_{\text{mecH}} = U + K$. In the two-body system in particular, our dependent variable is r and $U(r)$ takes its familiar form

$$U_g(r) = \frac{-Gm_1m_2}{r}. \tag{51}$$

Our kinetic energy is slightly more complicated. On an elliptical orbit the velocity $\vec{v}$ is typically neither totally radial nor azimuthal; in general

$$\begin{aligned} \vec{v} &= \vec{v}_r + \vec{v}_\phi \\ &= \dot{\vec{r}} + \dot{\phi}\vec{r}. \end{aligned}$$

So the total kinetic energy actually depends on r, inasmuch as

$$K = \frac{1}{2}\mu\dot{r}^2 + \frac{1}{2}\mu r^2\dot{\phi}^2. \tag{52}$$

We can write the second part of this expression for K in terms of the an-

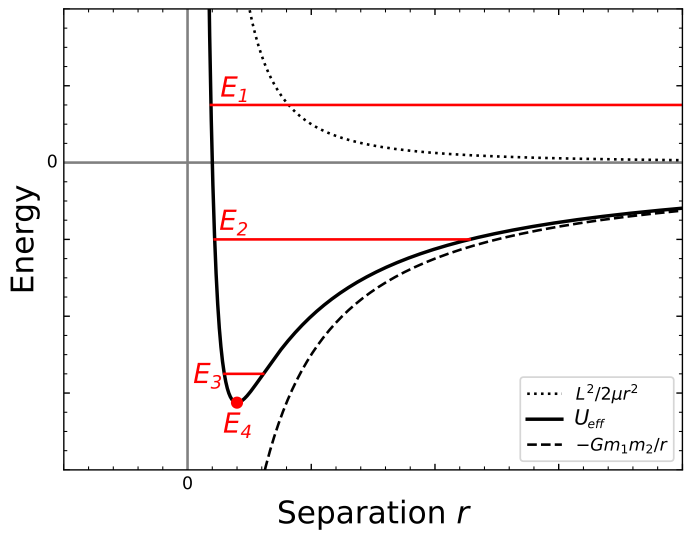

Figure 7: Energy diagram for the the two-body problem, showing the contributions of both the centripetal and gravitational terms to the final effective potential U_{eff}. Various permitted energy levels are indicated by E_1, E_2, etc.

gular momentum, which we determined earlier to be constant. We have

$$L = |\vec{r} \times \mu\vec{v}|$$
$$= \mu r v$$
$$= \mu r^2 \dot{\phi}.$$

So in our expression for $K(r)$, $\dot{\phi}^2 = L^2/\mu^2 r^4$ and thus Eq. 52 becomes

$$K = \frac{1}{2}\mu\dot{r}^2 + \frac{L^2}{2\mu r^2}. \tag{53}$$

Our expression for the total energy of the two-body system is then

$$E = \frac{1}{2}\mu\dot{r}^2 + \frac{L^2}{2\mu r^2} - \frac{Gm_1}{m_2}r, \tag{54}$$

and we typically combine the last two terms together into an **effective potential** U_{eff}.

Fig. 7 shows the final energy diagram for two-body problem (compare to Fig. 6). In the extreme case that we have zero angular momentum, this means that the full potential would be the dashed gravitational term: $L = 0$ means that the objects are heading for a collision at $r = 0$. On the other hand, this means that so long as $L > 0$ there is always a forbidden zone at small r. Just because the Earth is gravitationally attracted to the Sun doesn't mean that we can just fall into it!

For nonzero L, the red lines show various permitted energy levels. As we know, a gravitationally bound system has negative total energy (Eq. 51). So if a two-body system has $U_{\rm eff} > 0$ (such as E_1 in the figure) the system is **unbound**: the objects can escape from each other away toward infinity. At the other extreme, there is some minimum possible energy (labeled here as E_4). At this energy, only a single separation r is permitted: thus the distance between the objects is always the same; they're on a circular orbit. As energy is added to a circular orbit, wider and wider ranges of r are accessible to the system (e.g. E_3 and E_2: adding energy to a circular orbit increases its eccentricity. When $U_{\rm eff} = 0$ the elliptical orbit becomes an unclosed parabola, and for any larger energy (e.g. E_1) the orbit becomes hyperbolic.