

## 10.1: The Inverse-Square Law

Up to this point, all I have told you about gravity is that, near the surface of the Earth, the gravitational force exerted by the Earth on an object of mass  $m$  is  $F^G = mg$ . This is, indeed, a pretty good approximation, but it does not really tell you anything about what the gravitational force is where other objects or distances are involved.

The first comprehensive theory of gravity, formulated by Isaac Newton in the late 17th century, postulates that any two “particles” with masses  $m_1$  and  $m_2$  will exert an attractive force (a “pull”) on each other, whose magnitude is proportional to the product of the masses, and inversely proportional to the square of the distance between them. Mathematically, we write

$$F_{12}^G = \frac{Gm_1m_2}{r_{12}^2}. \quad (10.1.1)$$

Here,  $r_{12}$  is just the magnitude of the position vector of particle 2 relative to particle 1 (so  $r_{12}$  is, indeed, the distance between the two particles), and  $G$  is a constant, known as “Newton’s constant” or the *gravitational constant*, which at the time of Newton still had not been determined experimentally. You can see from Equation (10.1.1) that  $G$  is simply the magnitude, in newtons, of the attractive force between two 1-kg masses a distance of 1 m apart. This turns out to have the ridiculously small value  $G = 6.674 \times 10^{-11} \text{ N m}^2/\text{kg}^2$  (or, as is more commonly written,  $\text{m}^3\text{kg}^{-1}\text{s}^{-2}$ ). It was first measured by Henry Cavendish in 1798, in what was, without a doubt, an experimental tour de force for that time (more on that later, but you may peek at the “Advanced Topics” section of next chapter if you are curious already). As you can see, gravity as a force between any two ordinary objects is absolutely insignificant, and it takes the mass of a planet to make it into something you can feel.

Equation (10.1.1), as stated, applies to particles, that is to say, in practice, to any objects that are very small compared to the distance between them. The net force between extended masses can be obtained using calculus, by breaking up the two objects into very small pieces and adding (vectorially!) the force exerted by every small part of one object on every small part of the other object. This requires the use of integral calculus at a fairly advanced level, and for irregularly shaped objects can only be computed numerically. For spherically-symmetric objects, however, it turns out that the result (10.1.1) still holds exactly, provided the quantity  $r_{12}$  is taken to be the distance between the center of the spheres. The same result also holds for the force between a finite-size sphere and a “particle,” again with the distance to the particle measured from the center of the sphere.

For the rest of this chapter, we will simply assume that Equation (10.1.1) is a good approximation to be used in any of the problems we will encounter involving extended objects. You can estimate visually how good it may be, for instance, when applied to the Earth-moon system (as we will do in a moment), from a look at Figure 10.1.1 below:



Figure 10.1.1: The moon, the earth, and the distance between them, all approximately to scale.

Based on this picture it seems that it might be OK to treat the moon as a “particle,” but that it would not do, in general, to neglect the radius of the Earth; that is to say, we should use for  $r_{12}$  the distance from the moon to the center of the Earth, not just to its surface.

Before we get there, however, let us start closer to home and see what happens if we try to use Equation (10.1.1) to calculate the force exerted by the Earth on an object of mass  $m$  near its surface—say, at a height  $h$  above the ground. Clearly, if the radius of the Earth is  $R_E$ , the distance of the object to the center of the Earth is  $R_E + h$ , and this is what we should use for  $r_{12}$  in Equation (10.1.1). However, noting that the radius of the Earth is about 6,000 km (more precisely,  $R_E = 6.37 \times 10^6 \text{ m}$ ), and the tallest mountain peak is only about 9 km above sea level, you can see that it will almost always be a very good approximation to set  $r_{12}$  equal to just  $R_E$ , which results in a force

$$F_{E,o}^G \simeq \frac{GM_E m}{R_E^2} = m \frac{6.674 \times 10^{-11} \text{ m}^3\text{kg}^{-1}\text{s}^{-2} \times 5.97 \times 10^{24} \text{ kg}}{(6.37 \times 10^6 \text{ m})^2} = m \times 9.82 \frac{\text{m}}{\text{s}^2} \quad (10.1.2)$$

where I have used the currently known value  $M_E = 5.97 \times 10^{24} \text{ kg}$  for the mass of the Earth. As you can see, we recover the familiar result  $F^G = mg$ , with  $g \simeq 9.8 \text{ m/s}^2$ , which we have been using all semester. We can rewrite this result (canceling the mass  $m$ ) in the form

$$g_E = \frac{GM_E}{R_E^2}. \quad (10.1.3)$$

Here I have put a subscript  $E$  on  $g$  to emphasize that this is the acceleration of gravity near the surface of the Earth, and that the same formula could be used to find the acceleration of gravity near the surface of any other planet or moon, just replacing  $M_E$  and  $R_E$  by the mass and radius of the planet or moon in question. Thus, we could speak of  $g_{\text{moon}}$ ,  $g_{\text{Mars}}$ , etc., and in some homework problems you will be asked to calculate these quantities. Clearly, besides telling you how fast things fall on a given planet, the quantity  $g_{\text{planet}}$  allows you to figure out how much something would weigh on that planet's surface (just multiply  $g_{\text{planet}}$  by the mass of the object); alternatively, the ratio of  $g_{\text{planet}}$  to  $g_E$  will be the ratio of the object's weight on that planet to its weight here on Earth.

Of course, historically, this is not what Newton and his contemporaries would have done: they had measurements of objects in free fall (or sliding on inclined planes) that would have given them the value of  $g$ , and they even had a pretty good idea of the radius of the Earth<sup>1</sup>, but they did not know either  $G$  or the mass of the Earth, so all they could get from Equation (10.1.3) was the value of the product  $GM_E$ . It was only a century later, when Cavendish measured  $G$ , that they could get from that the mass of the earth (as a result of which, he became known as “the man who weighed the earth”!)

What Newton could do, however, with just this knowledge of the value of  $GM_E$ , was something that, for its time, was even more dramatic and far-reaching: namely, he could “prove” his intuition that the same fundamental interaction—gravity—that causes an apple to fall near the surface of the Earth, reaching out hundreds of thousands of miles away into space, *also* provides the force needed to keep the moon on its orbit. This brought together Earth-bound science (physics) and “celestial” science (astronomy) in a way that no one had ever dreamed of before.

To see how this works, let us start by assuming that the moon does move in a circle, with the Earth motionless at the center (these are all approximations, as we shall see later, but they give the right order of magnitude at the end, which is all that Newton could have hoped for anyway). The force  $F_{E,m}^G$  then has to provide the centripetal force  $F_c = M_{\text{moon}}\omega^2 r_{e,m}$ , where  $\omega$  is the moon's angular velocity. We can cancel the moon's mass from both terms and write this condition as

$$\frac{GM_E}{r_{e,m}^2} = \omega^2 r_{e,m}. \quad (10.1.4)$$

The moon revolves around the earth once about every 29 days, which is about  $29 \times 24 \times 3600 = 2.5 \times 10^6$  s. So  $\omega$  is  $2\pi$  radians per 2.5 million seconds, or  $\omega = 2.5 \times 10^{-6}$  rad/s. Substituting this into Equation (10.1.4), as well as the result  $GM_E = gR_E^2$  (note that, as stated above, we do not need to know  $G$  and  $M_E$  separately), we get  $r_{e,m} = 3.99 \times 10^8$  m, pretty close to today's accepted value of the average Earth-moon distance, which is  $3.84 \times 10^8$  km. Newton would not have known  $r_{e,m}$  to such an accuracy, but he would still have had a pretty good idea that this was, indeed, the correct order of magnitude<sup>2</sup>.

<sup>1</sup>The radius of the earth was known, at least approximately, since ancient times. The Greeks first figured it out by watching the way tall objects disappear below the horizon when viewed from a ship at sea. Next time somebody tries to tell you that people used to believe the earth was flat, ask them just which “people” they are talking about!

<sup>2</sup>And, yes, the first estimates of the distance between the earth and the moon also go back to the ancient Greeks! You can figure it out, if you know the radius of the earth, by looking at how long it takes the moon to transit across the earth's shadow during a lunar eclipse.

## Gravitational Potential Energy

Ever since I introduced the concept of potential energy in Chapter 5, I have been using  $U^G = mgy$  for the gravitational potential energy of the system formed by the Earth and an object of mass  $m$  a height  $y$  above the Earth's surface. This works well as long as the force of gravity is approximately constant, which is to say, as long as  $y$  is much smaller than the radius of the earth, but obviously it must break down at some point.

Recall that, if the interaction between two objects can be described by a potential energy function of the objects' coordinates,  $U(x_1, x_2)$ , then (in one dimension) the force exerted by object 1 on object 2 could be written as  $F_{12} = -dU/dx_2$ . Since the force of gravity does lie along the line joining the two particles, we can cheat a bit and treat this as a one-dimensional problem, with  $(F_{12}^G)_x = -Gm_1m_2/(x_1 - x_2)^2$  (I've put a minus sign there under the assumption that particle 1 is to the left of particle 2, that is,  $x_1 < x_2$ , and the force on 2 is to the left), and find a potential energy function whose derivative gives that. The answer is clearly

$$U^G(x_1, x_2) = -\frac{Gm_1m_2}{x_2 - x_1} + C \quad (10.1.5)$$

where  $C$  is an arbitrary constant. (Please take a moment to verify for yourself that, indeed,  $-dU^G/dx_2 = -(F_{12}^G)_x$ , and also  $-dU^G/dx_1 = -(F_{21}^G)_x$ .)

Since I have assumed  $x_2 > x_1$ , the denominator in Equation (10.1.5) is just the distance between the two particles, and the potential energy function could be written, in three dimensions, as

$$U^G(r_{12}) = -\frac{Gm_1m_2}{r_{12}} \quad (10.1.6)$$

where  $r_{12} = |\vec{r}_2 - \vec{r}_1|$ , and I have set the constant  $C$  equal to zero. This means that the potential energy of the system is always *negative*, which is, on the face of it, a strange result. However, there is no way to choose the constant  $C$  in (10.1.5) that will prevent that: no matter how big and positive  $C$  might be, the first term in (10.1.5) can always become larger (in magnitude) and negative, if the particles are very close together. So we might as well choose  $C = 0$ , which, at least, gives us the somewhat comforting result that the potential energy of the system is zero when the particles are “infinitely” distant from each other—that is to say, so far apart that they do not feel a force from each other any more.

But Equation (10.1.5) also makes sense in a different way: namely, it shows that the system’s potential energy *increases* as the particles are moved farther apart. Indeed, we expect, physically, that if you separate the particles by a great distance and then release them, they will pick up a lot of speed as they approach each other; or, put differently, that the force doing work over a large distance will give them a large amount of kinetic energy—which must come from the system’s potential energy. But, in fact, mathematically, Equation (10.1.6) agrees with this expectation: for any finite distance,  $U^G$  is negative, and it gets smaller *in magnitude* as the distance increases, which means algebraically it *increases* (since a number like, say,  $-0.1$  is, in fact, greater than a number like  $-10$ ). So as the particles are moved farther and farther apart, the potential energy of the system *does* increase—all the way up to a maximum value of zero!

Still, even if it makes sense mathematically, the notion of a “negative energy” is hard to wrap your mind around. I can only offer you two possible ways to look at it. One is to simply not think of potential energy as being anything like a “substance,” but just an accounting device that we use to keep track of the potential that a system has to do work for us—or (more or less equivalently) to give us kinetic energy, which is always positive and hence may be thought of as the “real” energy. From this point of view, whether  $U$  is positive or negative does not matter: all that matters is the *change*  $\Delta U$ , and whether this change has a sign that makes sense. This, at least, is the case here, as I have argued in the paragraph above.

The other perspective is almost opposite, and based on Einstein’s theory of relativity: in this theory, the total energy of a system is indeed “something like a substance,” in that it is directly related to the system’s total inertia,  $m$ , through the famous equation  $E = mc^2$ . From this point of view, the total energy of a system of two particles, interacting gravitationally, at rest, and separated by a distance  $r_{12}$ , would be the sum of the gravitational potential energy (negative), and the two particles’ “rest energies,”  $m_1c^2$  and  $m_2c^2$ :

$$E_{\text{total}} = m_1c^2 + m_2c^2 - \frac{Gm_1m_2}{r_{12}} \quad (10.1.7)$$

and this quantity *will* always be positive, unless one of the “particles” is a black hole and the other one is inside it<sup>3</sup>!

Please note that we will *not* use equation (10.1.7) this semester at all, since we are concerned only with nonrelativistic mechanics here. In other words, we will not include the “rest energy” in our calculations of a system’s total energy at all. However, if we did, we would find that a system whose rest energy is given by Equation (10.1.7) does, in fact, have an inertia that is *less* than the sum  $m_1 + m_2$ . This strongly suggests that the negativity of the potential energy is not just a mathematical convenience, but rather it reflects a fundamental physical fact.

For a system of more than two particles, the total gravitational potential energy would be obtained by adding expressions like (10.1.6) over all the pairs of particles. Thus, for instance, for three particles one would have

$$U^G(\vec{r}_1, \vec{r}_2, \vec{r}_3) = -\frac{Gm_1m_2}{r_{12}} - \frac{Gm_1m_3}{r_{13}} - \frac{Gm_2m_3}{r_{23}}. \quad (10.1.8)$$

A large mass such as the earth, or a star, has an intrinsic amount of gravitational potential energy that can be calculated by breaking it up into small parts and performing a sum like (10.1.8) over all the possible pairs of “parts.” (As usual, this sum is usually

evaluated as an integral, by taking the limit of an infinite number of infinitesimally small parts.) This “self-energy” does not change with time, and hence does not need to be included in most energy calculations involving gravitational forces between extended objects.

One thing that you may be wondering about, regarding Equation (10.1.6) for the potential energy of a pair of particles (or, for that matter, Equation (10.1.1) for the force), is what happens when the distance  $r_{12}$  goes to zero, since the mathematical expression appears to become infinite. This is technically true, but, in practice, it would only be a problem for a pair of true point particles—objects that would literally be mathematical points, with no dimensions at all. Such things may exist in some sense—electrons may well be an example—but they need to be described by quantum mechanics, which is an altogether different mathematical theory.

For finite-sized objects, you cannot continue to use an equation like (10.1.6) (or (10.1.1), for the force) when you are *under the surface* of the object. If you could dig a tunnel all the way down to the center of a hypothetical “earth” that had a constant density, the potential energy of the system formed by this “earth” and a particle of mass  $m$ , a distance  $r$  from the center, would look as shown in Figure 10.1.2 Notice how  $U^G$  becomes “flat,” indicating an equilibrium position (zero force), as  $r \rightarrow 0$ . It stands to reason that the net gravitational force at the center of this model “earth” should be zero, since one would be pulled equally strongly in all directions by all the mass around.

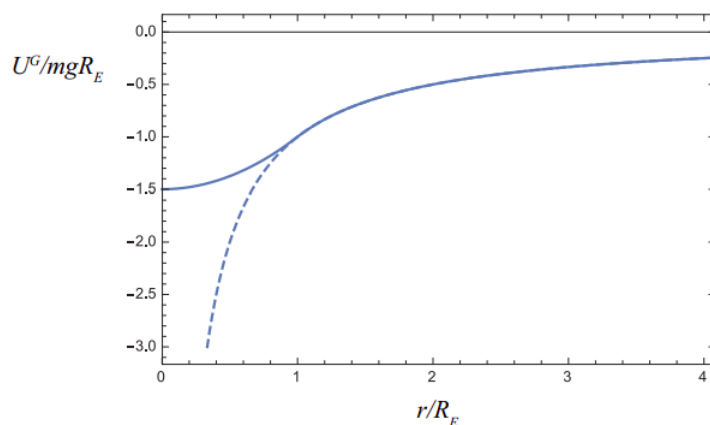


Figure 10.1.2 Gravitational potential energy of a system formed by a particle of mass  $m$  and a hypothetical earth with uniform density, a mass  $M$ , and a radius  $R_E$ , as a function of the distance  $r$  between the particle and the center of the “earth” (solid line). The dashed line shows the result for a system of two (point-like) particles. The energy  $U^G$  is expressed in units of  $mgR_E$ , where  $g = GM/R_E^2$ .

Finally, let me show you that the result (10.1.6) is fully consistent with the approximation  $U^G = mgy$  that we have been using up till now near the surface of the earth. (If you are not interested in mathematical derivations, feel free to skip this next bit.) Consider a particle of mass  $m$  that is initially on the surface of the earth, and then we move it to a height  $h$  above the earth. The change in potential energy, according to (10.1.6), is

$$U_f^G - U_i^G = -\frac{GM_E m}{R_E + h} + \frac{GM_E m}{R_E}. \quad (10.1.9)$$

If we write both terms with a common denominator, we get

$$U_f^G - U_i^G = \frac{GM_E m}{(R_E + h) R_E} h \simeq \frac{GM_E m}{R_E^2} h = mgh. \quad (10.1.10)$$

The only approximation here has been to set  $R_E + h \simeq R_E$  in the denominator of this expression. Since  $R_E$  is of the order of thousands of kilometers, this is an excellent approximation, as long as  $h$  is less than, say, a few hundred meters.

<sup>3</sup>For a justification of this statement, please see the definition of the Schwarzschild radius, later on in this chapter.

## Types of Orbits Under an Inverse-Square Force

Consider a system formed by two particles (or two perfect, rigid spheres) interacting only with each other, through their gravitational attraction. Conservation of the total momentum tells us that the center of mass of the system is either at rest or moving with constant velocity. Let us assume that one of the objects has a much greater mass,  $M$ , than the other, so that, for practical purposes, its center coincides with the center of mass of the whole system. This is not a bad approximation if what we are interested in is, for instance, the orbit of a planet around the sun. The most massive planet, Jupiter, has only about 0.001 times the mass of the sun.

Accordingly, we will assume that the more massive object does not move at all (by working in its center of mass reference frame, if necessary—note that, by our assumptions, this will be an inertial reference frame to a good approximation), and we will be concerned only with the motion of the less massive object under the force  $F = GMm/r^2$ , where  $r$  is the distance between the centers of the two objects. Since this force is always pulling towards the center of the more massive object (it is what is often called a *central force*), its torque around that point is zero, and therefore the angular momentum,  $\vec{L}$ , of the less massive body around the center of mass of the system is constant. This is an interesting result: it tells us, for instance, that the motion is *confined to a plane*, the same plane that the vectors  $\vec{r}$  and  $\vec{v}$  defined initially, since their cross-product cannot change.

In spite of this simplification, the calculation of the object's trajectory, or *orbit*, requires some fairly advanced mathematical techniques, except for the simplest case, which is that of a circular orbit of radius  $R$ . Note that this case requires a very precise relationship to hold between the object's velocity and the orbit's radius, which we can get by setting the force of gravity equal to the centripetal force:

$$\frac{GMm}{R^2} = \frac{mv^2}{R}. \quad (10.1.11)$$

So, if we want to, say, put a satellite in a circular orbit around a central body of mass  $M$  and at a distance  $R$  from the center of that body, we can do it, but only provided we give the satellite an initial velocity  $v = \sqrt{GM/R}$  in a direction perpendicular to the radius. But what if we were to release the satellite at the same distance  $R$ , but with a different velocity, either in magnitude or direction? Too much speed would pull it away from the circle, so the distance to the center,  $r$ , would temporarily increase; this would increase the system's potential energy and accordingly reduce the satellite's velocity, so eventually it would get pulled back; then it would speed up again, and so on.

You may experiment with this kind of thing yourself using the PhET demo at this link:

<https://phet.colorado.edu/en/simulation/gravity-and-orbits>

You will find that, as long as you do not give the satellite—or planet, in the simulation—too much speed (more on this later!) the orbit you get is, in fact, a closed curve, the kind of curve we call an *ellipse*. I have drawn one such ellipse for you in Figure 10.1.3

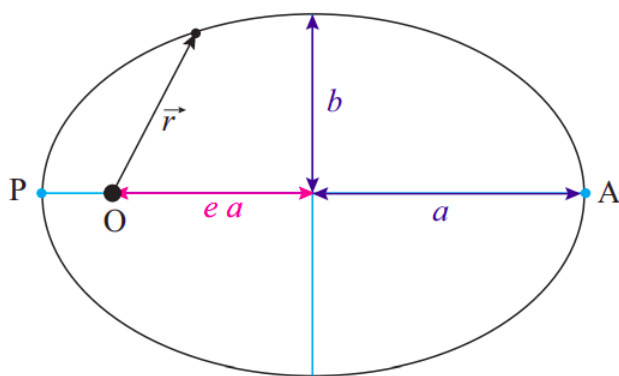


Figure 10.1.3 An elliptical orbit. The semimajor axis is  $a$ , the semiminor axis is  $b$ , and the eccentricity  $e = \sqrt{1 - b^2/a^2} = 0.745$  in this case.. The “center of attraction” (the sun, for instance, in the case of a planet's or comet's orbit) is at the point O.

As a geometrical curve, any ellipse can be characterized by a couple of numbers,  $a$  and  $b$ , which are the lengths of the *semimajor* and *semiminor* axes, respectively. These lengths are shown in the figure. Alternatively, one could specify  $a$  and a parameter known

as the *eccentricity*, denoted by  $e$  (do not mistake this “ $e$ ” for the coefficient of restitution of Chapter 4!), which is equal to  $e = \sqrt{1 - b^2/a^2}$ . If  $a = b$ , or  $e = 0$ , the ellipse becomes a circle.

The most striking feature of the elliptical orbits under the influence of the  $1/r^2$  gravitational force is that the “central object” (the sun, for instance, if we are interested in the orbit of a planet, asteroid or comet) is *not* at the geometric center of the ellipse. Rather, it is at a special point called the *focus* of the ellipse (labeled “O” in the figure, since that is the origin for the position vector of the orbiting body). There are actually two foci, symmetrically placed on the horizontal (major) axis, and the distance of each focus to the center of the ellipse is given by the product  $ea$ , that is, the product of the eccentricity and the semimajor axis. (This explains why the “eccentricity” is called that: it is a measure of how “off-center” the focus is.)

For an object moving in an elliptical orbit around the sun, the distance to the sun is minimal at a point called the perihelion, and maximal at a point called the aphelion. Those points are shown in the figure and labeled “P” and “A”, respectively. For an object in orbit around the earth, the corresponding terms are perigee and apogee; for an orbit around some unspecified central body, the terms periapsis and apoapsis are used. There is some confusion as to whether the distances are to be measured from the surface or from the center of the central body; here I will assume they are all measured from the center, in which case the following relationships follow directly from Figure 10.1.3

$$\begin{aligned} r_{\max} &= (1 + e)a \\ r_{\min} &= (1 - e)a \\ r_{\min} + r_{\max} &= 2a \\ e &= \frac{r_{\max} - r_{\min}}{2a}. \end{aligned} \quad (10.1.12)$$

The ellipse I have drawn in Figure 10.1.3 is actually way too eccentric to represent the orbit of any planet in the solar system (although it could well be the orbit of a comet). The planet with the most eccentric orbit is Mercury, and that is only  $e = 0.21$ . This means that  $b = 0.978a$ , an almost imperceptible deviation from a circle. I have drawn the orbit to scale in Figure 10.1.4 and as you can see the only way you can tell it is an ellipse is, precisely, because the sun is not at the center.

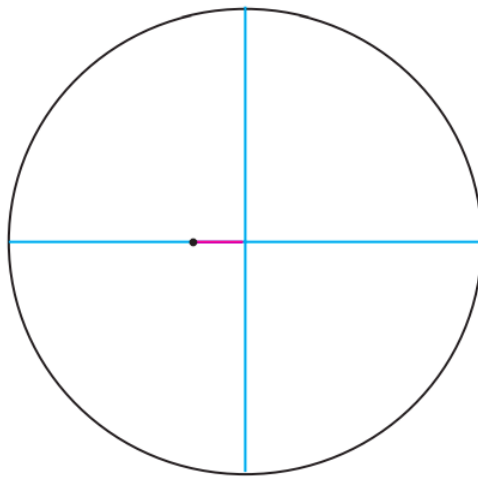


Figure 10.1.4 Orbit of Mercury, with the sun approximately to scale

Since an ellipse has only two parameters, and we have two constants of the motion (the total energy,  $E$ , and the angular momentum,  $L$ ), we should be able to determine what the orbit will look like based on just those two quantities. Under the assumption we are making here, that the very massive object does not move at all, the total energy of the system is just

$$E = \frac{1}{2}mv^2 - \frac{GMm}{r}. \quad (10.1.13)$$

For a circular orbit, the radius  $R$  determines the speed (as per Equation (10.1.11)), and hence the total energy, which is easily seen to be  $E = -\frac{GMm}{2R}$ . It turns out that this formula holds also for elliptical orbits, if one substitutes the semimajor axis  $a$  for  $R$ :

$$E = -\frac{GMm}{2a}. \quad (10.1.14)$$

Note that the total energy (10.1.14) is *negative*. This means that we have a *bound* orbit, by which I mean, a situation where the orbiting object does not have enough kinetic energy to fly arbitrarily far away from the center of attraction. Indeed, since  $U^G \rightarrow 0$  as  $r \rightarrow \infty$ , you can see from Equation (10.1.13) that if the two objects could be infinitely far apart, the total energy would eventually have to be positive, for any nonzero speed of the lighter object. So, if  $E < 0$ , we have bound orbits, which are ellipses (of which a circle is a special case), and conversely, if  $E > 0$  we have “unbound” trajectories, which turn out to be hyperbolas<sup>4</sup>. These trajectories just pass near the center of attraction once, and never return.

The special borderline case when  $E = 0$  corresponds to a *parabolic* trajectory. In this case, the particle also never comes back: it has just enough kinetic energy to make it “to infinity,” slowing down all the while, so  $v \rightarrow 0$  as  $r \rightarrow \infty$ . The initial speed necessary to accomplish this, starting from an initial distance  $r_i$ , is usually called the “escape velocity” (although it really should be called the escape speed), and it is found by simply setting Equation (10.1.13) equal to zero, with  $r = r_i$ , and solving for  $v$ :

$$v_{esc} = \sqrt{\frac{2GM}{r_i}}. \quad (10.1.15)$$

In general, you can calculate the escape speed from any initial distance  $r_i$  to the central object, but most often it is calculated from its surface. Note that  $v_{esc}$  does not depend on the mass of the lighter object (always assuming that the heavier object does not move at all). The escape velocity from the surface of the earth is about 11 km/s, or  $1.1 \times 10^4$  m/s; but this alone would not be enough to let you leave the attraction of the sun behind. The escape speed from the sun starting from a point on the earth’s orbit is 42 km/s.

To summarize all of the above, suppose you are trying to put something in orbit around a much more massive body, and you start out a distance  $r$  away from the center of that body. If you give the object a speed smaller than the escape speed at that point, the result will be  $E < 0$  and an elliptical orbit (of which a circle is a special case, if you give it the precise speed  $v = \sqrt{GM/r}$  in the right direction). If you give it precisely the escape speed (10.1.15), the total energy of the system will be zero and the trajectory of the object will be a parabola; and if you give it more speed than  $v_{esc}$ , the total energy will be positive and the trajectory will be a hyperbola. This is illustrated in Figure 10.1.5 below.

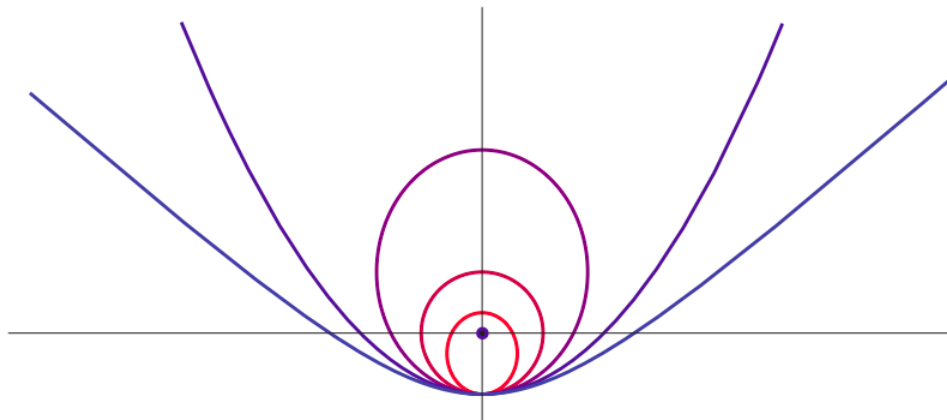


Figure 10.1.5 Possible trajectories for an object that is “released” with a sideways velocity at the lowest point in the figure, under the gravitational attraction of a large mass represented by the black circle. Each trajectory corresponds to a different value of the object’s initial kinetic energy: if  $K_{circ}$  is the kinetic energy needed to have a circular orbit through the point of release, the figure shows the cases  $K_i = 0.5K_{circ}$  (small ellipse),  $K_i = K_{circ}$  (circle),  $K_i = 1.5K_{circ}$  (large ellipse),  $K_i = 2K_{circ}$  (escape velocity, parabola), and  $K_i = 2.5K_{circ}$  (hyperbola).

Note that all the trajectories shown in Figure 10.1.5 have the same potential energy at the “point of release” (since the distance from that point to the center of attraction is the same for all), so increasing the kinetic energy at that point also means increasing the total energy (10.1.13) (which is constant throughout). So the picture shows different orbits in order of increasing total energy.

For a given total energy, the total angular momentum does not change the fundamental nature of the orbit (bound or unbound), but it can make a big difference on the orbit’s shape. Generally speaking, for a given energy the orbits with less angular momentum will be “narrower,” or “more squished” than the ones with more angular momentum, since a smaller initial angular momentum at the point of insertion means a smaller sideways velocity component. In the extreme case of zero initial angular momentum (no



sideways velocity at all), the trajectory, regardless of the total energy, reduces to a straight line, either straight towards or straight away from the center of attraction.

For elliptical orbits, one can prove the result

$$e = \sqrt{1 - \frac{L^2}{aGMm^2}} \quad (10.1.16)$$

which shows how the eccentricity increases as  $L$  decreases, for a given value of  $a$  (which is to say, for a given total energy). I should at least sketch how to obtain this result, since it is a variant of a procedure that you may have to use for some homework problems this semester. You start by writing the angular momentum as  $L = mr_P v_P$  (or  $mr_A v_A$ ), where  $A$  and  $P$  are the special points shown in Figure 10.1.3 where  $\vec{v}$  and  $\vec{r}$  are perpendicular. Then, you note that  $r_P = r_{\min} = (1 - e)a$  (or, alternatively,  $r_A = r_{\max} = (1 + e)a$ ), so  $v_P = L/[m(1 - e)a]$ . Then substitute these expressions for  $r_P$  and  $v_P$  in Equation (10.1.13), set the result equal to the total energy (10.1.14), and solve for  $e$ .

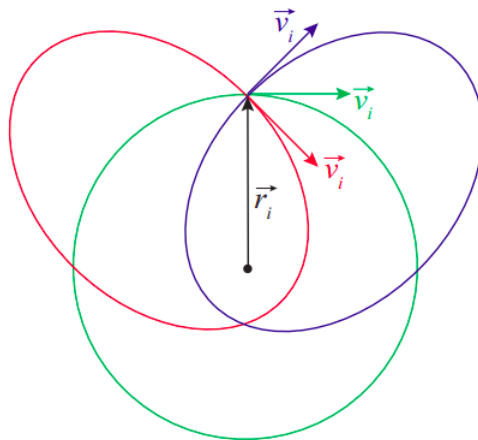


Figure 10.1.6 Effect of the "angle of insertion" on the orbit.

Figure 10.1.6 illustrates the effect of varying the angular momentum, for a given energy. All the initial velocity vectors in the figure have the same magnitude, and the release point (with position vector  $\vec{r}_i$ ) is the same for all the orbits, so they all have the same energy; indeed, you can check that the semimajor axis of the two ellipses is the same as the radius of the circle, as required by Equation (10.1.14). The difference between the orbits is their total angular momentum. The green orbit has the maximum angular momentum possible at the given energy, since the green velocity vector is perpendicular to  $\vec{r}_i$ . Note that this (maximizing  $L$  for a given  $E < 0$ ) always results in a circle, in agreement with Equation (10.1.15): the eccentricity is zero when  $L = L_{\text{circ}} \equiv \sqrt{aGMm^2}$ , which is the largest value of  $L$  allowed in Equation (10.1.15).

For the other two orbits,  $\vec{v}_i$  and  $\vec{r}_i$  make angles of  $45^\circ$  and  $135^\circ$ , and so the angular momentum  $L$  has magnitude  $L = L_{\text{circ}} \sin 45^\circ = L_{\text{circ}}/\sqrt{2}$ . The result are the red and blue ellipses, with eccentricities  $e = \sqrt{1 - \sin^2(45^\circ)} = 0.707$ .

<sup>4</sup>There is apparently a way to describe a hyperbola as an ellipse with eccentricity  $e > 1$ , but I'm definitely not going to go there.

## Kepler's Laws

The first great success of Newton's theory was to account for the results that Johannes Kepler had extracted from astronomical data on the motion of the planets around the sun. Kepler had managed to find a number of regularities in a mountain of data (most of which were observations by his mentor, the Danish astronomer Tycho Brahe), and expressed them in a succinct way in mathematical form. These results have come to be known as *Kepler's laws*, and they are as follows:

1. The planets move around the sun in elliptical orbits, with the sun at one focus of the ellipse
2. (Law of areas) A line that connects the planet to the sun (the planet's position vector) sweeps equal areas in equal times.
3. The square of the orbital period of any planet is proportional to the cube of the semimajor axis of its orbit (the same proportionality constant holds for all the planets).



I have discussed the first “law” at length in the previous section, and also pointed out that the math necessary to prove it is far from trivial. The second law, on the other hand, while it sounds complicated, turns out to be a straightforward consequence of the conservation of angular momentum. To see what it means, consider Figure 10.1.7.

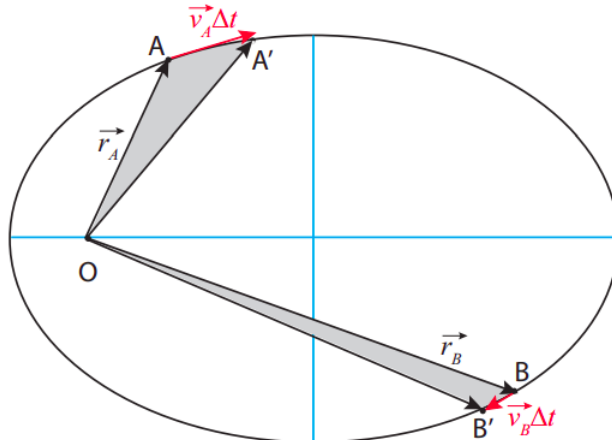


Figure 10.1.7: Illustrating Kepler’s law of areas. The two gray “curved triangles” have the same area, so the particle must take the same time to go from A to A’ as it does to go from B to B’.

Suppose that, at some time  $t_A$ , the particle is at point A, and a time  $\Delta t$  later it has moved to A’. The area “swept” by its position vector is shown in grey in the figure, and Kepler’s second law states that it must be the same, for the same time interval, at any point in the trajectory; so, for instance, if the particle starts out at B instead, then in the same time interval  $\Delta t$  it will move to a point B’ such that the area of the “curved triangle” OBB’ equals the area of OAA’.

Qualitatively, this means that the particle needs to move more slowly when it is farther from the center of attraction, and faster when it is closer. Quantitatively, this actually just means that its angular momentum is constant! To see this, note that the straight distance from A to A’ is the displacement vector  $\Delta \vec{r}_A$ , which, for a sufficiently short interval  $\Delta t$ , will be approximately equal to  $\vec{v}_A \Delta t$ . Again, for small  $\Delta t$ , the area of the curved triangle will be approximately the same as that of the straight triangle OAA’. It is a well-known result in trigonometry that the area of a triangle is equal to 1/2 the product of the lengths of any two of its sides times the sine of the angle they make. So, if the two triangles in the figures have the same areas, we must have

$$|\vec{r}_A| |\vec{v}_A| \Delta t \sin \theta_A = |\vec{r}_B| |\vec{v}_B| \Delta t \sin \theta_B \quad (10.1.17)$$

and we recognize here the condition  $|\vec{L}_A| = |\vec{L}_B|$ , which is to say, conservation of angular momentum. (Once the result is established for infinitesimally small  $\Delta t$ , we can establish it for finite-size areas by using integral calculus, which is to say, in essence, by breaking up large triangles into sums of many small ones.)

As for Kepler’s third result, it is easy to establish for a circular orbit, and definitely not easy for an elliptical one. Let us call  $T$  the orbital period, that is, the time it takes for the less massive object to go around the orbit once. For a circular orbit, the angular velocity  $\omega$  can be written in terms of  $T$  as  $\omega = 2\pi/T$ , and hence the regular speed  $v = R\omega = 2\pi R/T$ . Substituting this in Equation (10.1.11), we get  $GM/R^2 = 4\pi^2 R/T^2$ , which can be simplified further to read

$$T^2 = \frac{4\pi^2}{GM} R^3. \quad (10.1.18)$$

Again, this turns out to work for an elliptical orbit if we replace  $R$  by  $a$ .

Note that the proportionality constant in Equation (10.1.18) depends only on the mass of the central body. For the solar system, that would be the sun, of course, and then the formula would apply to any planet, asteroid, or comet, with the same proportionality constant. This gives you a quick way to calculate the orbital period of anything orbiting the sun, if you know its distance (or vice-versa), based on the fact that you know what these quantities are for the Earth.

More generally, suppose you have two planets, 1 and 2, both orbiting the same star, at distances  $R_1$  and  $R_2$ , respectively. Then their orbital periods  $T_1$  and  $T_2$  must satisfy  $T_1^2 = (4\pi^2/GM) R_1^3$  and  $T_2^2 = (4\pi^2/GM) R_2^3$ . Divide one equation by the other, and the proportionality constant cancels, so you get

$$\left(\frac{T_2}{T_1}\right)^2 = \left(\frac{R_2}{R_1}\right)^3. \quad (10.1.19)$$

From this, some simple manipulation gives you

$$T_2 = T_1 \left(\frac{R_2}{R_1}\right)^{3/2}. \quad (10.1.20)$$

Note you can express  $R_1$  and  $R_2$  in any units you like, as long as you use the same units for both, and similarly  $T_1$  and  $T_2$ . For instance, if you use the Earth as your reference “planet 1,” then you know that  $T_1 = 1$  (in years), and  $R_1 = 1$ , in AU (an AU, or “astronomical unit,” is the distance from the Earth to the sun). A hypothetical planet at a distance of 4 AU from the sun should then have an orbital period of 8 Earth-years, since  $4^{3/2} = \sqrt{4^3} = \sqrt{64} = 8$ .

A formula just like (10.1.18), but with a different proportionality constant, would apply to the satellites of any given planet; for instance, the myriad of artificial satellites that orbit the Earth. Again, you could introduce a “reference satellite” labeled 1, with known period and distance to the Earth (the moon, for instance?), and derive again the result (10.1.20), which would allow you to get the period of any other satellite, if you knew how its distance to the earth compares to the moon’s (or, conversely, the distance at which you would need to place it in order to get a desired orbital period).

For instance, suppose I want to place a satellite on a “geosynchronous” orbit, meaning that it takes 1 day for it to orbit the Earth. I know the moon takes 29 days, so I can write Equation (10.1.20) as  $1 = 29(R_2/R_1)^{3/2}$ , or, solving it,  $R_2/R_1 = (1/29)^{2/3} = 0.106$ , meaning the satellite would have to be approximately 1/10 of the Earth-moon distance from (the center of) the Earth.

In hindsight, it is somewhat remarkable that Kepler’s laws are as accurate, for the solar system, as they turned out to be, since they can only be mathematically derived from Newton’s theory by making a number of simplifying approximations: that the sun does not move, that the gravitational force of the other planets has no effect on each planet’s orbit, and that the planets (and the sun) are perfect spheres, for instance. The first two of these approximations work as well as they do because the sun is so massive; the third one works because the sizes of all the objects involved (including the sun) are much smaller than the corresponding orbits. Nevertheless, Newton’s work made it clear that Kepler’s laws could only be approximately valid, and scientists soon set to work on developing ways to calculate the corrections necessary to deal with, for instance, the trajectories of comets or the orbit of the moon.

Of the main approximations I have listed above, the easiest one to get rid of (mathematically) is the first one, namely, that the sun does not move. Instead, what one finds is that, as long as the sun and the planet are still treated as an isolated system, they will both revolve around the system’s center of mass. Of course, the sun’s motion (a slight “wobble”) is very small, but not completely negligible. You can even see it in the simulation I mentioned earlier, at

[phet.colorado.edu/en/simulation/gravity-and-orbits](http://phet.colorado.edu/en/simulation/gravity-and-orbits).

What is much harder to deal with, mathematically, is the fact that none of the planets in the solar system actually forms an isolated system with the sun, since all the planets are really pulling gravitationally on each other all the time. Particularly, Jupiter and Saturn have a non-negligible influence on each other’s orbits, and on the orbits of every other planet, which can only be perceived over centuries. Basically, the orbits still look like ellipses to a very good degree, but the ellipses rotate very, very slowly (so they fail to exactly close in on themselves). This effect, known as orbital precession, is most dramatic for Mercury, where the ellipse’s axes rotate by more than one degree per century.

Nevertheless, the Newtonian theory is so accurate, and the calculation techniques developed over the centuries so sophisticated, that by the early 20th century the precession of the orbits of all planets *except* Mercury had been calculated to near exact agreement with the best observational data. The unexplained discrepancy for Mercury amounted only to 43 seconds of arc per century, out of 5600 (an error of only 0.8%). It was eventually resolved by Einstein’s general theory of relativity.

---

This page titled 10.1: The Inverse-Square Law is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Julio Gea-Banacloche (University of Arkansas Libraries) via source content that was edited to the style and standards of the LibreTexts platform.