PHYSICS 9HA: CLASSICAL MECHANICS

Tom Weideman
University of California, Davis

LibreTexts™

UCD: Physics 9HA – Classical Mechanics

The LibreTexts libraries are Powered by NICE CXOne and are supported by the Department of Education Open Textbook Pilot Project, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org. More information on our activities can be found via Facebook (https://facebook.com/Libretexts), Twitter (https://twitter.com/libretexts), or our blog (http://Blog.Libretexts.org).

This text was compiled on 04/15/2025

TABLE OF CONTENTS

Licensing

A detailed breakdown of this resource's licensing can be found in ***Back Matter/Detailed Licensing****.*

CHAPTER OVERVIEW

1: Motion

1.1: Vectors

Definition of a Vector

Just being able to put numbers on physical quantities is not sufficient for describing nature. Very often physical quantities have directions. For example, a description of something's motion is incomplete if you merely state how fast it is going. [*Okay, so an asteroid is moving at 35,000 miles per hour, but is it headed for Earth?!*] We therefore have the following definition for physical quantities that exhibit both these properties:

Definition: Vector

A vector is a quantity with both magnitude and direction.

We will frequently represent a vector quantity with an arrow, where the direction of the vector is the direction that the arrow points, and the magnitude of the vector is represented by the length of the arrow. This is not to say that vectors *are* arrows – arrows just make a handy geometric representation. So while an arrow representing a vector might be 6*cm* long, that doesn't mean that the vector has a magnitude of 6*cm*. The vector might represent the speed and direction of a moving object, for example, and then the vector's magnitude isn't even in units of *cm*. However, if we draw two arrow representations of the same sort of quantity, and one is twice as long as the other, the implication is that the longer arrow represents a vector with twice the magnitude of the vector represented by the shorter arrow.

Alert

There is no way to compare magnitudes of different physical quantities. If a distance vector is drawn as an arrow on the same page as a velocity vector's arrow, the relative sizes of the two arrows are meaningless.

There are a few other things that we should say about vectors and the arrows that represent them:

- *Where* the arrow representing a vector is positioned is not a distinct feature of the vector. That is, an arrow representing a vector can be moved at will, and so long as it isn't stretched, shrunk, or rotated, it will represent the same vector. Just changing an arrow's location does not change its magnitude or its direction if it is moved carefully.
- Vector directions (and therefore the directions of their representative arrows) can be reversed mathematically through multiplication by –1.
- Vector lengths can be expanded or shrunk (scaled) through multiplication by a regular number (called a scalar). If the number is greater than 1, the vector expands in length, and if it is less than 1, it contracts.

One other thing... When we write a symbol for a vector quantity, we will do so with a small arrow above the letter, like this: $\overrightarrow{A}$. Variables with the same letter as a defined vector that do not include an arrow, are assumed to represent the *magnitude* of that vector. So for example, when used in the same context, the variable A represents the magnitude of $\overrightarrow{A}$.

Vector Addition/Subtraction

For these mathematical quantities we call vectors to have any value to us, they have to allow for simple mathematical operations, such as addition. The directional nature of vectors makes addition much trickier than simply summing the two magnitudes. It turns out that a well-defined vector addition involves simple geometry. It goes like this: Transport one of the vectors (in a parallel fashion, so as not to change its direction) so that its tail is in contact with the head of the other vector. Then fashion a new vector such that its tail is at the open tail and its head is at the open head.

Figure 1.1.1 – Graphical Vector Addition

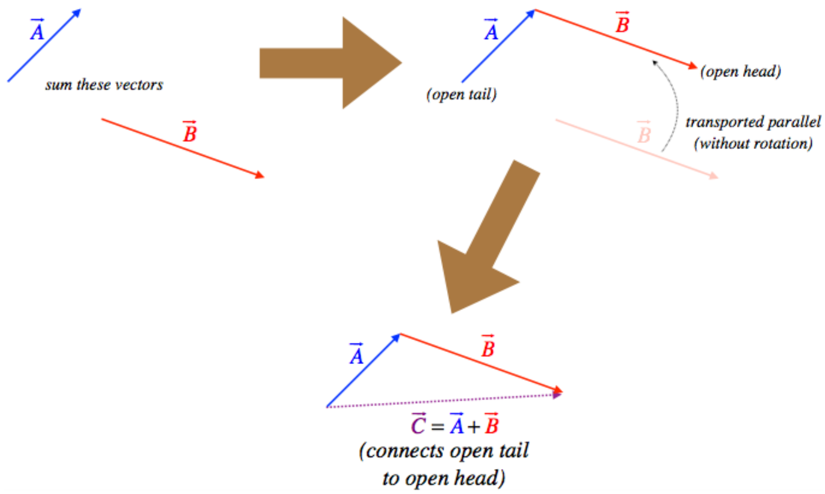
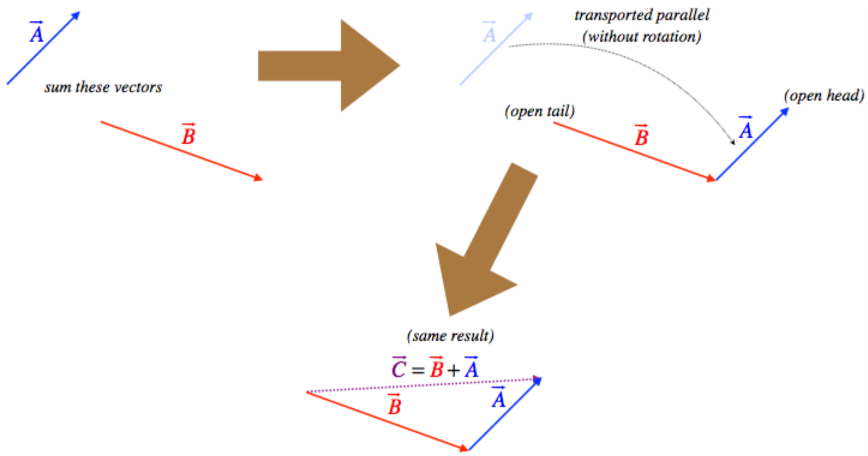


Figure 1.1.2 – Vector Addition is Commutative



What about subtracting two vectors? Well, we can do this by following the same method as for regular numbers: Whichever vector we wish to subtract we multiply by –1, and then add the result to the other vector, which we do in the manner described above. We already know that multiplying a vector by –1 reverses its direction (and leaves its magnitude unchanged), so this is a well-defined operation for us.

Vector Components

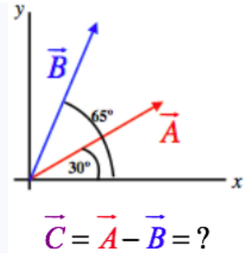
The graphical method of adding vectors are not always convenient. For example, we shouldn't have to actually measure the length of the new vector, we should be able to calculate it. Well, of course we can do this using some sophisticated knowledge of triangles. For example, given we know the lengths and directions of the two vectors we are adding, we can determine the length of the third leg of the triangle using the Law of Cosines:

$$C^2 = A^2 + B^2 - 2AB\cos\theta \tag{1.1.1}$$

With all of the lengths of the triangle legs and one of the angles (the one between A and B), we can get the other angles using the Law of Sines.

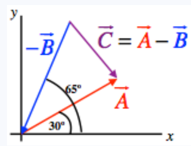
Example 1.1.1

The magnitudes of the two vectors shown in the diagram below are: $A = 132$ and $B = 145$. Find the magnitude and direction (angle made with the x-axis) of the vector that is the difference of these two vectors.



Solution

Using the fact that the negative of a vector is the same vector pointing in the opposite direction along with using tail-to-head vector addition, we get the following diagram for the three vectors:



The angle between $\vec{A}$ and $\vec{B}$ is obviously 65° – 30° = 35°, so for this triangle we have the lengths of two sides and the angle between them. We can therefore find the length of the third side ($\vec{C}$) from the law of cosines:

$$C^2 = A^2 + B^2 - 2AB\cos\theta \quad \Rightarrow \quad C = \sqrt{(132)^2 + (145)^2 - 2\,(132)\,(145)\cos(35^o)} = 84.2$$

Next we can determine the angle between $\vec{A}$ and $\vec{C}$ using the law of sines:

$$\frac{\sin 35^o}{C} = \frac{\sin\theta_{AC}}{B} \quad \Rightarrow \quad \theta_{AC} = \sin^{-1}\left[\frac{145}{84.2}\sin 35^o\right] = 81^o$$

If we rotate $\vec{C}$ counterclockwise through this angle, it will be parallel to $\vec{A}$, if we then rotated it back clockwise by 30° (the angle $\vec{A}$ makes with the x-axis), then it will be parallel to the x-axis. Therefore the angle $\vec{C}$ makes with the x-axis is: –81° + 30° = –51° (below the x-axis). This answer certainly conforms with the diagram above, which shows $\vec{C}$ with a smaller magnitude than $\vec{A}$ and $\vec{B}$ and pointing down to the right.

While we can use these tools to mathematically solve for the sum of two vectors, it turns out that there is another way we can do it that doesn't require quite as much geometrical reasoning. This method exploits three simple facts:

- We can replace any single vector as a sum of two (or more) vectors.
- It is easy to add two vectors that are parallel.
- If we use right triangles, trigonometry is easier to work with than with general triangles and the law of cosines/sines.

The trick is to select two (or three, if necessary) perpendicular axes (they do not have to be horizontal and vertical, they only need to be perpendicular to each other), and break up every vector involved into a sum of two perpendicular vectors parallel to these axes. The lengths of these perpendicular vectors are called the *components of the vector along those axes*. Going back to the list of advantages above, remember that we can add similar components like numbers, and we can determine these components easily using trigonometry.

Figure 1.1.3 – Vector Components

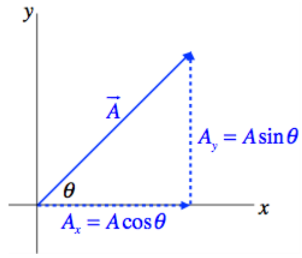
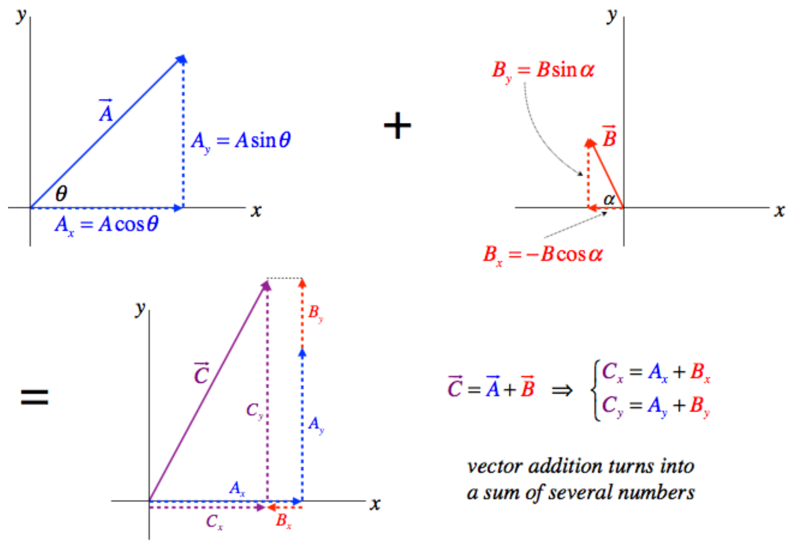


Figure 1.1.4 – Summing Vectors Using Components



The sums of components are like summing numbers, but only components along the same axes can be added. The results are then more components, which then have to be reconstructed into a vector.

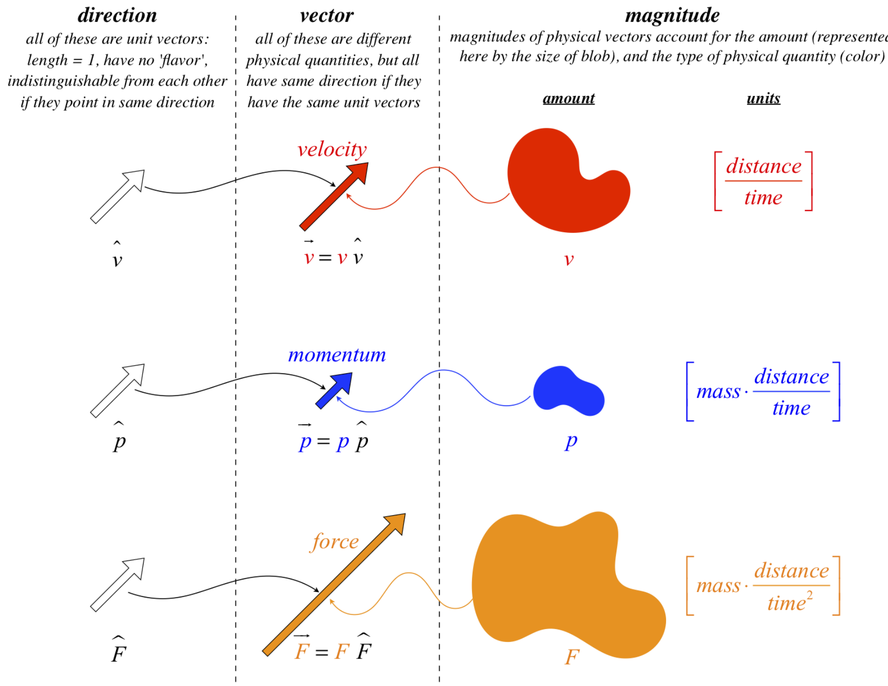
Unit Vectors

So we can use perpendicular coordinate systems to describe vectors in terms of their components. Essentially this means that to describe a vector in terms of a set of three axes, we need to know three numbers. But it might be useful to actually express these vectors as a single mathematical entity, and that's where the notion of the *unit vector* comes in. Vectors have magnitude and direction, and with unit vectors we can mathematically break up the vector into those two parts. The magnitude is just a number (with physical units) without direction, and a unit vector is a vector (without units) that has a length of 1, so that it can be scaled to any length without contributing anything to the magnitude. Therefore we can write a vector as a simple product:

$$\overrightarrow{A} = A\widehat{A} \tag{1.1.2}$$

where $\widehat{A}$ is the unit vector (usually pronounced "A-hat"). It is a unitless vector of length 1 that points in the direction of the vector $\overrightarrow{A}$. The value A is a number with physical units that equals the magnitude. The diagram below gives a graphic description of how this construction works for a few common physical vectors. The unit vectors provide a very basic template by defining the direction, and the magnitude fills in the template by contributing the girth and 'flavor' (physical units) of the vector.

Figure 1.1.4 – Unit Vectors and Magnitudes



If we combine this notion with components, we can write any vector as a sum of components multiplying unit vectors in the directions of the three spatial dimensions. By convention, we give these unit vectors the names $\hat{i}$, $\hat{j}$, and $\hat{k}$ for the axes x, y, and z, respectively. So specifically, we have:

$$\overrightarrow{A} = A_x\hat{i} + A_y\hat{j} + A_z\hat{k} \tag{1.1.3}$$

Now we can just use this as a mathematical representation of vectors, and we do not have to appeal to geometry at all. For example,

$$\begin{aligned}\overrightarrow{C} &= \overrightarrow{A} + \overrightarrow{B} \\ &= (A_x\hat{i} + A_y\hat{j} + A_z\hat{k}) + (B_x\hat{i} + B_y\hat{j} + B_z\hat{k}) \\ &= C_x\hat{i} + C_y\hat{j} + C_z\hat{k} \\ &= (A_x + B_x)\hat{i} + (A_y + B_y)\hat{j} + (A_z + B_z)\hat{k}\end{aligned} \tag{1.1.4}$$

Giving us the same result as we got before for the components of the sum of two vectors.

Example 1.1.2

Repeat the calculation of Example 1.1.1, this time using components.

Solution

Breaking the two vectors into their x and y components gives:

$$\overrightarrow{A} = A_x\hat{i} + A_y\hat{j} = A\cos\theta\ \hat{i} + A\sin\theta\ \hat{j} = 132\cos 30^o\hat{i} + 132\sin 30^o\hat{j} = 114.3\ \hat{i} + 66.0\ \hat{j}$$
$$\overrightarrow{B} = A_x\hat{i} + A_y\hat{j} = B\cos\theta\ \hat{i} + B\sin\theta\ \hat{j} = 145\cos 65^o\hat{i} + 145\sin 65^o\hat{j} = 61.3\ \hat{i} + 131.4\ \hat{j}$$

Next we subtract $\overrightarrow{B}$ from $\overrightarrow{A}$ to get $\overrightarrow{C}$, then compute its magnitude (using the Pythagorean theorem) and direction (using trigonometry):

$$\overrightarrow{C} = \overrightarrow{A} - \overrightarrow{B} = 53.0\ \hat{i} - 65.4\ \hat{j}$$

$$\left|\overrightarrow{C}\right| = \sqrt{53.0^2 + 65.4^2} = 84.2$$

$$angle = \tan^{-1}\left(\frac{-65.4}{53.0}\right) = -51^o$$

This matches the answer found in example 1.1.1.

Column Matrices

As you read this introduction to unit vectors, it may have occurred to you that the $\hat{i}$, $\hat{j}$, and $\hat{k}$ are essentially placeholders that provide the "that-a-way" to accompany the amount in each direction (provided by the scalar multiplying each unit vector). Of course, there is nothing special about this "hat" notation, and in fact there is another notation that is commonly used – column matrices, with three rows; one for each direction. In other words, we can make these replacements:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \leftrightarrow \hat{i} \qquad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \leftrightarrow \hat{j} \qquad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \leftrightarrow \hat{k} \tag{1.1.5}$$

This means that a vector can be expressed in these two ways, both of them using the cartesian system:

$$\overrightarrow{A} = A_x\hat{i} + A_y\hat{j} + A_z\hat{k} \leftrightarrow \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \tag{1.1.6}$$

Addition of vectors works in the usual method of matrix addition – simply add the values in the same slots and place them in the same slot:

$$\overrightarrow{A} + \overrightarrow{B} \leftrightarrow \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} + \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} = \begin{pmatrix} A_x + B_x \\ A_y + B_y \\ A_z + B_z \end{pmatrix} \tag{1.1.7}$$

Both of these notations have their practical advantages. In this text, we will stick with the "hat" notation, but you should be aware that the matrix notation is very common, so you should be comfortable using it and expect to see it frequently in classes to come.

1.2: Vector Multiplication

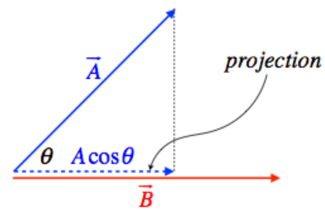
Now we know how to do some math with vectors, and the question arises, “If we can add and subtract vectors, can we also multiply them?” The answer is yes and no. It turns out that there is not one unique way to define a product of two vectors, but instead there are two…

Scalar Product

Soon we will be looking at how we can describe the effect that a force pushing on an object has on its speed as it moves from one position to another. The force is a vector, because it has a magnitude (the amount of the push) and a direction (the way the push is directed). And the movement of the object is also a vector (tail is at the object’s starting point, and head is at its ending point). It will turn out that this effect is describable mathematically as the product of the amount of force and the amount of movement. This is simple to compute if the push is along the direction of movement, but what if it is not? It turns out that only the amount of push that *acts in the direction of the movement* will affect the object's speed.

We therefore would like to introduce the notion of the *projection* of one vector onto another. The best description of this is, "the amount of a given vector that points along the other vector." This could be imagined as the “shadow” one vector casts upon another vector:

Figure 1.2.1 – Projecting One Vector Onto Another



$$\left(\textit{amount of vector } \overrightarrow{A} \textit{ that points along the direction of vector } \overrightarrow{B}\right) = A\cos\theta$$

So if we want to multiply the length of a vector by the amount of a second vector that is projected onto it we get:

$$(\text{projection of } \overrightarrow{A} \text{ onto } \overrightarrow{B})(\text{magnitude of } \overrightarrow{B}) = (A\cos\theta)(B) = AB\cos\theta \tag{1.2.1}$$

This is the first of the two types of vector multiplication, and it is called a *scalar product*, because the result of the product is a scalar. We usually write the product with a dot (giving its alternative name of *dot product*):

$$\overrightarrow{A}\cdot\overrightarrow{B} \equiv AB\cos\theta, \qquad \theta = \textit{angle between } \overrightarrow{A} \textit{ and } \overrightarrow{B} \tag{1.2.2}$$

Example 1.2.1

The vector $\overrightarrow{A}$ has a magnitude of 120 units, and when projected onto $\overrightarrow{B}$, the projected portion has a value of 105 units. Suppose that $\overrightarrow{B}$ is now projected onto $\overrightarrow{A}$, and the projected length is 49 units. Find the magnitude of $\overrightarrow{B}$.

Solution

The factor that determines the length of the projection is $\cos\theta$. The angle between the two vectors is the same regardless of which vector is projected, so the factor is the same in both directions. The projection of $\overrightarrow{A}$ onto $\overrightarrow{B}$ is 7/8 of the magnitude of $\overrightarrow{A}$, so the magnitude of $\overrightarrow{B}$ must be 8/7 of its projection, which is 56 units. Note that when the projection of one vector is multiplied by the magnitude of the other, the same product results regardless of which way the projection occurs. That is, the scalar product is the same in either order (i.e. it is commutative).

Note that a scalar product of a vector with itself is the square of the magnitude of that vector:

$$\overrightarrow{A}\cdot\overrightarrow{A} = A^2\cos 0 = A^2 \tag{1.2.3}$$

It should be immediately clear what the scalar products of the unit vectors are. They have unit length, so a scalar product of a unit vector with itself is just 1.

$$\hat{i}\cdot\hat{i}=\hat{j}\cdot\hat{j}=\hat{k}\cdot\hat{k}=1 \tag{1.2.4}$$

They are also *mutually orthogonal*, so the scalar products with each other are zero:

$$\hat{i}\cdot\hat{j}=\hat{j}\cdot\hat{k}=\hat{k}\cdot\hat{i}=0 \tag{1.2.5}$$

This gives us an alternative way to look at components, which are projections of a vector onto the coordinate axes. Since the unit vectors point along the x, y, and z directions, the components of a vector can be expressed as a dot product. For example:

$$\begin{aligned}\overrightarrow{A}\cdot\hat{i} &= (A_x\hat{i}+A_y\hat{j}+A_z\hat{k})\cdot\hat{i}\\ &= A_x\,\cancelto{1}{\hat{i}\cdot\hat{i}}+A_y\,\cancelto{0}{\hat{j}\cdot\hat{i}}+A_z\,\cancelto{0}{\hat{k}\cdot\hat{i}}\end{aligned} \tag{1.2.6}$$

Unit vectors also show us an easy way to take the scalar product of two vectors whose components we know. Start with two vectors written in component form:

$$\begin{aligned}\overrightarrow{A} &= A_x\hat{i}+A_y\hat{j}\\ \overrightarrow{B} &= B_x\hat{i}+B_y\hat{j}\end{aligned}$$

then just do "normal algebra," distributing the dot product as you would with normal multiplication:

$$\begin{aligned}\overrightarrow{A}\cdot\overrightarrow{B} &= (A_x\hat{i}+A_y\hat{j})\cdot(B_x\hat{i}+B_y\hat{j})\\ &= (A_xB_x)\,\cancelto{1}{\hat{i}\cdot\hat{i}}+(A_yB_x)\,\cancelto{0}{\hat{j}\cdot\hat{i}}+(A_xB_y)\,\cancelto{0}{\hat{i}\cdot\hat{j}}+(A_yB_y)\,\cancelto{1}{\hat{j}\cdot\hat{j}}\\ &= A_xB_x+A_yB_y\end{aligned} \tag{1.2.7}$$

If we didn't have this simple result, think about what we would have to do: We would need to calculate the angles each vector makes with (say) the x-axis. Then from those two angles, we need to figure out the angles between the two vectors. Then we would need to compute the magnitudes of the two vectors. Finally, with the magnitudes of the vectors and the angle between the vectors, we could finally plug into our scalar product equation.

Alert

With two different ways to compute a scalar product, it should be clear that the simplest method to use will depend upon what information is provided. If you are given (or can easily ascertain) the magnitudes of the vectors and the angle between them, then use Equation 1.2.2, but if you are given (or can easily ascertain) the components of the vectors, use Equation 1.2.7.

Example 1.2.2

The two vectors given below are perpendicular to each other. Find the unknown z-component.

$$\overrightarrow{A}=+5\hat{i}-4\hat{j}-\hat{k}\qquad\overrightarrow{B}=+2\hat{i}+3\hat{j}+B_z\hat{k}$$

Solution

The scalar product of two vectors is proportional to the cosine of the angle between them. This means that if they are orthogonal, the scalar product is zero. The dot product is easy to compute when given the components, so we do so and solve for B_z:

$$0=\overrightarrow{A}\cdot\overrightarrow{B}=(+5)(+2)+(-4)(+3)\;+(-1)(B_z)\quad\Rightarrow\quad B_z=-2$$

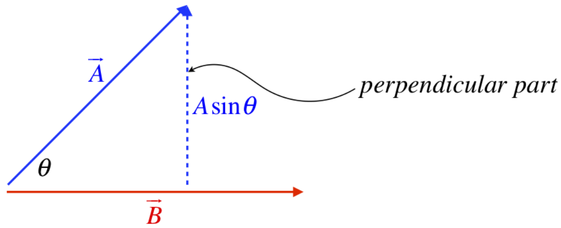
The scalar product of two vectors in terms of column vectors works exactly how you would expect – simply multiply the similar components and sum all the products. Or, if you are familiar with matrix multiplication, multiply the transpose (row matrix) of one vector by the column matrix of the other:

$$\vec{A} \cdot \vec{B} = \begin{pmatrix} A_x & A_y & A_z \end{pmatrix} \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} = A_x B_x + A_y B_y + A_z B_z \tag{1.2.8}$$

Vector Product

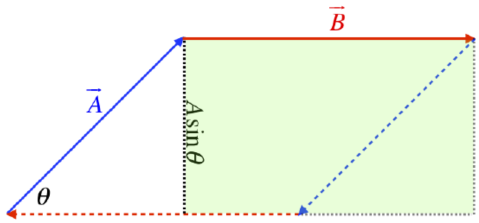
As mentioned earlier, there are actually two ways to define products of vectors. If the scalar product involves the amount of one vector that is *parallel* to the other vector, then it should not be surprising that our other product involves the amount of a vector that is *perpendicular* to the other vector.

Figure 1.2.2 – Portion of One Vector Perpendicular to Another



If we take a product like before, multiplying this perpendicular piece by the magnitude of the other vector, we get an expression similar to what we got for the scalar product, this time with a sine function rather than a cosine. For reasons that will be clear soon, this type of product is referred to as a *vector product*. Because this is distinct from the scalar product, we use a different mathematical notation as well – a cross rather than a dot (giving it an alternative name of *cross product*). This has a simple (though not entirely useful, at least not in physics) geometric interpretation in terms of the parallelogram defined by the two vectors:

Figure 1.2.3 – Constructing a Vector Product (Magnitude)



$$magnitude\ of\ \vec{A} \times \vec{B} = \left|\vec{A} \times \vec{B}\right| = \left|\vec{A}\right|\left|\vec{B}\right| \sin\theta = AB\sin\theta \tag{1.2.9}$$

But there is another even bigger difference between the vector and scalar products. While the projection always lands parallel to the second vector, the perpendicular part implies an orientation, since the perpendicular part can point in multiple directions. Any quantity that has an orientation has the potential to be a vector, and in fact we will define a vector that results from this type of product as follows: If we follow the perimeter of the parallelogram above in the direction given by the two vectors, we get a clockwise orientation [*Would we get the same orientation if the product was in the opposite order:* $\vec{B} \times \vec{A}$?]. We turn this circulation direction into a vector direction (which points in a specific direction in space) using a convention called the *right-hand-rule*:

Convention: Right-Hand-Rule

Point the fingers of your right hand in the direction of the first vector in the product, then orient your hand such that those fingers curl naturally into the direction of the second vector in the product. As your fingers curl, your extended thumb points in a direction that is perpendicular to both vectors in the product. This is the direction of the vector that results from the cross product.

If we perform the cross product with the vectors in the opposite order, our fingers curl in the opposite direction, which makes our thumb point in the opposite direction in space. This means that the cross product has an *anticommutative* property:

$$\vec{A} \times \vec{B} = -\vec{B} \times \vec{A} \tag{1.2.10}$$

A cross product of any vector with itself gives zero, since the part of the first vector that is perpendicular to the second vector is zero:

$$|\overrightarrow{A} \times \overrightarrow{A}| = A^2 \sin 0 = 0 \tag{1.2.11}$$

As with the scalar product, the vector product can be easily expressed with components and unit vectors. The vector products of the unit vectors with themselves are zero. Each of the unit vectors is at right angles with the other two unit vectors, so the magnitude of the cross product of two unit vectors is also a unit vector (since the sine of the angle between them is 1).

Convention: Right-handed Coordinate Systems

We will always choose a right-handed coordinate system, which means that using the right-hand-rule on the +x to +y axis yields the +z axis.

In terms of the unit vectors, we therefore have:

$$\hat{i} \times \hat{i} = \hat{j} \times \hat{j} = \hat{k} \times \hat{k} = 0 \tag{1.2.12}$$

and

$$\hat{i} \times \hat{j} = -\hat{j} \times \hat{i} = \hat{k}, \qquad \hat{j} \times \hat{k} = -\hat{k} \times \hat{j} = \hat{i}, \qquad \hat{k} \times \hat{i} = -\hat{i} \times \hat{k} = \hat{j} \tag{1.2.13}$$

This allows us to do cross products purely mathematically (without resorting to the right-hand-rule) when we know the components, as we did for the scalar product. Again start with two vectors in component form:

$$\overrightarrow{A} = A_x \hat{i} + A_y \hat{j}$$

$$\overrightarrow{B} = B_x \hat{i} + B_y \hat{j}$$

then, as in the case of the scalar product, just do "normal algebra," distributing the cross product, and applying the unit vector cross products above:

$$\overrightarrow{A} \times \overrightarrow{B} = (A_x \hat{i} + A_y \hat{j}) \times (B_x \hat{i} + B_y \hat{j}) \tag{1.2.14}$$

$$= (A_x B_x)\ \cancelto{0}{\hat{i} \times \hat{i}} + (A_y B_x)\ \cancelto{-\hat{k}}{\hat{j} \times \hat{i}} + (A_x B_y)\ \cancelto{+\hat{k}}{\hat{i} \times \hat{j}} + (A_y B_y)\ \cancelto{0}{\hat{j} \times \hat{j}} \tag{1.2.15}$$

$$= (A_x B_y - A_y B_x)\hat{k} \tag{1.2.16}$$

It is not obvious right now how we will use the dot and cross product in physics, but it is coming, so it's a good idea to get a firm grasp on these important tools now.

1.3: Straight-Line Motion

There is nothing more fundamental in the study of physics than motion. We will bring a lot of mathematics to bear on this subject (including the vectors we just learned about), but we are going to start as easy as possible – with motion that remains on a straight line. This simplifies our task in several ways, the primary of which is the reduction of vectors to simple positive and negative quantities (one direction is arbitrarily chosen as the positive direction, and the opposite direction is negative).

Displacement

In order for motion to occur for an object, obviously its *position* must change from one instant in time to another. We will refer to the coordinate position of the straight line on which the object moves as $x(t)$. A change in this position we call the *displacement*, and refer to it as a change in position:

$$\text{displacement} = \Delta x \equiv x_f - x_o \tag{1.3.1}$$

Alert

It's a good idea to get used to this now, as you will use it throughout the Physics 9-series: When referring to a time-dependent quantity, the "delta" (Δ) means "after minus before," or "final minus initial."

Notice that if the final position is a smaller number than the initial position, then the object has a negative displacement. Eventually we will treat displacement as a vector, but for our straight-line motion, the sign of the value provides all the information we need about the direction. In this text you will receive several warnings about the precise use of physics language, which is frequently at odds with how the same words are used in casual conversation. Here is the first example:

Alert

*"Displacement" sounds a lot like "distance covered." Walking a mile to the store and back again is a two mile walk, but the displacement in this case is **not** two miles. Displacement is a vector whose magnitude is the distance between the starting and ending points, and whose direction points from the starting point to the ending point.*

Average Velocity

Of course, there is more to motion than just displacement. We will generally also be interested in how fast that displacement occurs. We therefore define a rate called the *average velocity* thus:

$$\text{average velocity} = v_{ave} \equiv \frac{\Delta x}{\Delta t} = \frac{x_f - x_o}{t_f - t_o} \tag{1.3.2}$$

Since we know displacement is a vector (of course in our current simple 1-dimensional model it can only have two directions), then average velocity must be a vector as well.

Instantaneous Velocity

Just talking about the before and after gets pretty boring, so what do we do about during? That is, how do we define a velocity at a single moment in time – the instantaneous velocity? Well, we know the answer to this from calculus. We start with our idea of average velocity, and just shrink the time span down very small, until it vanishes:

$$\text{instantaneous velocity} = v = \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} = \frac{dx}{dt} \tag{1.3.3}$$

Average and Instantaneous Acceleration

We take our discussion of motion to one level more – we consider that things might speed up or slow down. Just as we defined average velocity in terms of before and after positions, we also define *average acceleration* in terms of before and after (instantaneous) velocities:

$$\text{average acceleration} = a_{ave} = \frac{\Delta v}{\Delta t} = \frac{v_f - v_o}{t_f - t_o} \tag{1.3.4}$$

And, as before, we use calculus to extend this notion of average acceleration to instantaneous acceleration, which we describe as the amount that our object is speeding up or slowing down at a single moment in time:

$$\text{instantaneous acceleration} = a = \lim_{\Delta t \to 0} \frac{\Delta v}{\Delta t} = \frac{dv}{dt} \tag{1.3.5}$$

$$= \frac{d^2x}{dt^2} \tag{1.3.6}$$

Alert

Another language warning – In standard English parlance, we are used to reserving the word "acceleration" to mean only "speeding up." In physics it means specifically the rate of change of velocity, which for straight line motion includes both speeding up and slowing down (for multi-dimensional motion it gets even trickier).

Example 1.3.1

If a moving object is slowing down, is it possible that the magnitude of its acceleration is increasing? If an object is speeding up, is it possible that the magnitude of its acceleration is decreasing? In either of these cases, can the magnitude of the acceleration be zero? Explain.

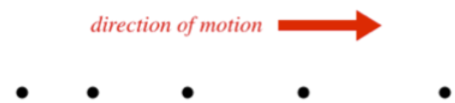
Solution

If an object is slowing down, then it is experiencing an acceleration in the direction opposite to its motion. If this acceleration increases in magnitude, then it slows down faster. So naturally it can be slowing down as the acceleration magnitude increases. Similarly, as an object is speeding up, it is experiencing an acceleration in the direction of its motion. If the magnitude of this acceleration decreases, then the rate at which it speeds up decreases, but it is still speeding up. If the object is either slowing down or speeding up, then its velocity must be changing, and the acceleration cannot be zero.

Motion Diagrams

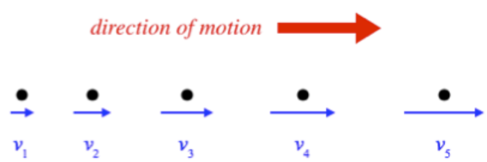
A motion diagram starts as merely a series of collinear dots that represent the position of an object at different equally-spaced intervals of time. You can think of it as a time-lapse photograph using a strobe light. There is one other piece of information that goes with this starting diagram: the direction that the object is moving. An example of this starting point might be this:

Figure 1.3.1a – Creating a Motion Diagram



From this we need to somehow extract the instantaneous velocity (magnitude and direction, which may be changing) at each position, and the acceleration (magnitude and direction, assumed to be constant throughout) of the object. At this point we are only working qualitatively, so our goal is to sketch onto the diagram velocity vectors at each dot that have magnitudes and directions that approximately represent the velocities of the object at those points (v_1, v_2, etc.), keeping in mind that the time intervals between dots are all the same, and the acceleration is constant throughout. You can do this intuitively (it must be going faster if it covers more distance in the same time), or you can figure it out from Equation 1.3.2. Adding the instantaneous velocity vectors to the above diagram makes it look like this:

Figure 1.3.1b – Creating a Motion Diagram

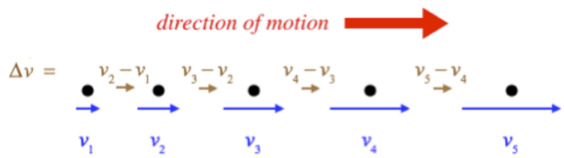


Now for acceleration. Since we are assuming constant acceleration (at least for the five-data-point interval we are considering), the average acceleration equals the instantaneous acceleration. With each dot being separated by the same time interval, the acceleration between dots is proportional to the velocity changes (magnitude and direction), and in this case of constant acceleration, is the same between every pair of dots:

$$a = constant = a_{ave} = \frac{\Delta v}{\Delta t} = \frac{v_2 - v_1}{t_2 - t_1} = \frac{v_3 - v_2}{t_3 - t_2} = \dots \tag{1.3.7}$$

Putting this into the diagram gives:

Figure 1.3.1c – Creating a Motion Diagram



Note that Δv is determined using the usual tail-to-head vector addition, which in one dimension just consists of keeping the signs straight.

If we didn't know whether or not the acceleration was constant, we could make a good guess by comparing the Δv's. Notice that we need two dots to determine the average velocity for a single time interval, since two dots gives us a displacement. But if we want to know how the speed is changing (i.e. the acceleration), we need three dots. If dots #1 and #2 are closer together than dots #2 and #3, we know the object has sped up, and if the first two dots are father apart, then the object is slowing down. So when we label our motion diagram, we can arbitrarily draw-in the first velocity vector on the first dot, but we can't add the velocity vector to the second dot if there is no third dot present to show us if the object is object is going faster, slower, or the same speed at the second dot. The more changes we want to consider (like if we want to know about a changing acceleration), the more dots we need.

This is in fact the nature of calculus – the change of a change of a change, etc, requires an additional measurement of position for each additional change computed. So the motion diagram only needs three dots if the acceleration is known (or assumed) to be constant, but to confirm that it is constant requires four dots.

1.4: Kinematics

Equations of Motion

Okay, enough of the definitions. Let's see how these things all fit together, and how they can be used. What we will be looking at are called the *equations of motion*, and this topic is often referred to as *kinematics*. It is important to note that we are not yet dealing with causes for these motions, but only the motions themselves.

We will mostly only deal with constant accelerations (unless otherwise specified), and since instantaneous acceleration is the derivative of velocity, it is not difficult in the case of constant acceleration to integrate it to get the instantaneous velocity as a function of time:

$$\left.\begin{array}{l} a=\dfrac{dv}{dt} \quad \Rightarrow \quad v(t)=\int a\;dt=at+const \\ const=v(t=0)\equiv v_o \end{array}\right\} \quad v(t)=at+v_o \tag{1.4.1}$$

The constant of integration is found by plugging $t=0$ into Equation 1.4.1, which results in the velocity of the object at the starting time, which is typically designated as v_o.

We can play exactly the same game to obtain the equation of motion for position as a function of time, since we know how it relates to the instantaneous velocity:

$$\left.\begin{array}{l} v=\dfrac{dx}{dt} \quad \Rightarrow \quad x(t)=\int v dt=\int (at+v_o)\,dt=\frac{1}{2}at^2+v_ot+const \\ const=x(t=0)\equiv x_o \end{array}\right\} \quad x(t)=\frac{1}{2}at^2+v_ot+x_o \tag{1.4.2}$$

Notice that if we have all the details of this last equation, we can obtain the velocity equation above simply by taking a derivative. We cannot go in the opposite direction without also obtaining the starting position.

Example 1.4.1

*The acceleration of a particle moving along the x-axis is given by the equation [note that it is **not** constant!]:*

$$\alpha(t)=\beta+\lambda t, \qquad \beta=2.40\frac{m}{s^2}, \qquad \lambda=0.300\frac{m}{s^2}$$

The particle is at position $x=+4.60m$ and is moving in the $-x$ direction at a speed of $12.0\frac{m}{s}$ at time $t=0s$.

a. *Find the time at which the particle (briefly) comes to rest.*
b. *Find the position where the particle (briefly) comes to rest.*

Solution

a. We start by finding the equation for the velocity, as we are interested in the time at which this value goes to zero. The velocity function is the integral of the acceleration function:

$$v(t)=\int a(t)\,dt=\int[\beta+\lambda t]\,dt=\beta t+\frac{1}{2}\lambda t^2+const$$

We can determine the integration constant by plugging in what we know about the velocity at time $t=0$:

$$v(t=0)=\beta\cdot 0+\frac{1}{2}\lambda(0)^2+const \quad \Rightarrow \quad const=v_o=-12.0\frac{m}{s}$$

Plugging this back into our equation for velocity and setting the velocity equal to zero, we can calculate the time at which this occurs:

$$v(t=0)=\frac{1}{2}\lambda t^2+\beta t+v_o \quad \Rightarrow \quad t=\frac{-\beta\pm\sqrt{\beta^2-4\left(\frac{1}{2}\lambda\right)v_o}}{2\left(\frac{1}{2}\lambda\right)}=\boxed{4.00s}$$

b. We know the time at which it comes to rest, so we need the equation for the position as a function of time. We get this by integrating the velocity function:

$$x(t)=\int v(t)\,dt=\int\left[\frac{1}{2}\lambda t^2+\beta t+v_o\right]dt=\frac{1}{6}\lambda t^3+\frac{1}{2}\beta t^2+v_ot+const$$

As we did above, we find the constant of integration by putting in what we know about the position at $t=0$:

$$x\left(t=0\right)=\frac{1}{6}\lambda(0)^3+\frac{1}{2}\beta(0)^2+v_o\cdot 0+const \quad\Rightarrow\quad const=4.60m$$

And finally, plug this result back into the equation for $x\left(t\right)$ and put in the time we found above ($t=4s$) to get:

$$x\left(4\right)=\boxed{-21.0m}$$

Let's make an accounting of all the numbers we can encounter in a constant-acceleration situation:

- independent variable: t
- dependent variables: x, v
- constants of the motion: x_o, v_o, a (acceleration is constant by assumption)

With six numbers to work with, you can imagine there are many ways to set up a problem to solve for something unknown. Everything you need to solve any such problem is provided in the above equations. However, it is often easier to put those equations together to form a new equation, to cut down on the algebra needs for certain types of problems. The most common useful re-combining of these variables involves eliminating time from the two equations, since you may be given velocities and positions. The algebra is straightforward:

$$\left.\begin{array}{l} v_f=at+v_o \;\Rightarrow\; t=\dfrac{v_f-v_o}{a} \\ x_f-x_o=\frac{1}{2}at^2+v_ot \end{array}\right\}\quad x_f-x_o=\frac{1}{2}a\left(\frac{v_f-v_o}{a}\right)^2+v_o\left(\frac{v_f-v_o}{a}\right)\quad\Rightarrow\quad 2a\left(x_f-x_o\right)=v_f{}^2-v_o{}^2 \tag{1.4.3}$$

You can think of this equation as the "before/after" equation, because it deals only with starting and ending positions and velocities, and has eliminated time as an input variable.

While we are accumulating useful (though unnecessary) equations for motion with constant acceleration, we should also include the two equations that involve average velocity. The first is just a rewriting of the definition of average velocity, with the "final" position occurring at time t:

$$v_{ave}=\frac{x_f-x_o}{t}=\frac{x\left(t\right)-x_o}{t}\quad\Rightarrow\quad x\left(t\right)=v_{ave}t+x_o \tag{1.4.4}$$

The second equation is quite useful, though it applies *only* to motion involving constant acceleration:

$$v_{ave}=\frac{x_f-x_o}{t}=\frac{\frac{1}{2}at^2+v_ot}{t}=\frac{1}{2}at+v_o=\frac{1}{2}\left(v_f-v_o\right)+v_o\quad\Rightarrow\quad v_{ave}=\frac{v_o+v_f}{2} \tag{1.4.5}$$

For constant acceleration, the average velocity simply equals the arithmetic average of the starting and ending velocities. We will better see why it comes out this way when we start discussing graphing shortly.

Free-Fall

There is one type of straight-line motion that involves constant acceleration that we are all familiar with: free-fall.

We will look more closely at how to explain this in terms of forces in a future section, but assuming air resistance has a small effect (remember, we are devising a simplified model here), then it turns out (as shown by Galileo dropping stones from the Tower of Pisa, and more dramatically in the demonstration) that objects all accelerate at the same constant rate as they fall to Earth. This rate of acceleration is commonly given the symbol g, and it has the value:

$$\text{acceleration due to gravity near the surface of the earth} = g = 9.8\,\frac{m}{s^2}$$

Note the units of distance-per-time-squared are the units of acceleration. This acceleration is of course always directed downward, and depending on our choice of coordinate system, this can be either positive or negative. Once the coordinate system is selected, the sign for g stays the same no matter which way the object is moving. If the positive direction is chosen to be upward, and the object is moving upward, then its velocity is positive and the negative value of g leads to a slowing of the object's motion. If it is moving down, then its velocity is negative, and the negative acceleration leads to the velocity becoming more negative (i.e. it is speeding up).

Example 1.4.2

A ball is thrown vertically upward at the same instant that a second ball is dropped from rest directly above it. The two balls are $12.0m$ apart when they start their motion. Find the maximum speed at which the first ball can be thrown such that it doesn't collide with the second ball before it returns to its starting height. Treat the balls as being very small (i.e. ignore their diameters).

Solution

Both balls are under the influence of the earth's gravity, and therefore both accelerate at a rate g downward. Treating up as the positive direction and the starting height of the ball thrown up as the origin, we have the following equations of motion for the two balls:

$$y_1(t) = -\frac{1}{2}gt^2 + v_o t + 0 \qquad y_2(t) = -\frac{1}{2}gt^2 + 0\cdot t + y_o$$

We can use the equation for the dropped ball and its starting height to determine the time it takes to reach the origin, then plug that result into the equation for the thrown ball, to determine how high it will be when the dropped ball gets to the origin:

$$t = \sqrt{\frac{2y_o}{g}} \quad\Rightarrow\quad y_1 = -\frac{1}{2}g\left(\sqrt{\frac{2y_o}{g}}\right)^2 + v_o\left(\sqrt{\frac{2y_o}{g}}\right) = -y_o + \sqrt{\frac{2y_o}{g}}v_o$$

We insist that the the thrown ball has fallen back to at least its starting height by the time the dropped ball gets there, so we want y_1 to be no greater than zero, which gives us an inequality for v_o:

$$v_o \le \sqrt{\frac{gy_o}{2}} \quad\Rightarrow\quad v_o \le 7.67\,\frac{m}{s}$$

1.5: Graphing

Interpreting Graphs

We conclude our discussion of straight-line motion by taking on the topic of representing motion with graphs. These graphs represent what is happening to the various dependent variables (x, v, and a) over time. There are three goals here:

1. To interpret a graph in terms of the physical motion of the object it represents.
2. To sketch a graph that represents the physical motion of an object, given a description of that motion.
3. To sketch a graph of one or two dependent variables based on the graph of another dependent variable.

Alert

These are not always easy tasks to perform, for two main reasons: First, our first instinct when we see a graph is to interpret it as a picture, rather than a plot of a quantity vs. time. The second problem (and this persists throughout the study of physics) is the tendency to confuse the change of a quantity for the value of that quantity. More precisely, we tend to lose sight of the fact that a variable's value at an instant and its rate of change are quite independent of each other.

For task #1, here are some of the questions we should be able to answer:

Q1: Where is the object (which side of the origin is it on)?

This would seem to be quite trivial (and it is): The position at any given time is the value on the vertical axis for the x vs. t graph. Where we run into trouble is thinking that we might have some idea of how to answer this question for the velocity and acceleration graphs. Those graphs only give us information about the object's changing position and changing speed, respectively, not where the object is at any given time. If we are separately given where the object is at some point in time (say at $t=0$), then we can determine its position at other times. One way to think of this is that the velocity graph gives us the *shape* of the position graph, but that shape could be located anywhere up-and-down the vertical axis. All of this is just repeating what we found in Section 1.4 – that when we integrate $v\left(t\right)$, we get an unknown constant x_o that must be provided separately.

Q2: Is the object at rest, or is it moving?

Another seemingly obvious question to answer, but again there are things to keep in mind. Although this is a property of velocity we *can* answer it using the position graph (we only get unknown constants when we integrate, not when we take derivatives). Mathematically, we know that the velocity is the slope of the position graph, so since "at rest" means zero velocity, the object is at rest when the tangent line to the x vs. t graph has zero slope. But we should strive to look at this *physically* as well. Obviously an object that is moving is one whose position is changing, so if the x value is changing, the object is moving. If we are given a v vs. t graph, we have to be careful not to use the same criterion as we did for the x vs. t graph. Instead, whether the object is moving or not is a simple matter of whether or not the value of v is zero. If we have the acceleration graph, then integrating it to get the velocity graph leaves an unknown constant (v_o). We know the shape of the v vs. t graph, but not where it is located up-and-down the vertical axis. This means that with just the acceleration graph we cannot know where the velocity graph crosses the horizontal axis, and therefore have no idea where the object is coming to rest.

Q3: Which way is the object moving?

The direction of motion of the object can also be obtained from both the position and velocity graphs. From the position graph, we know that the sign of the slope is the sign of the velocity (which is the direction of motion). On the velocity graph, we simply need to determine if the value of the velocity is positive or negative (i.e. is the graph below or above the horizontal axis). A common mistake is to confuse these two things. For example, the position graph being below the horizontal axis does not mean the object is moving in the $-x$ direction, and a positive slope of the velocity graph does not mean that the object is moving in the $+x$ direction. Once again, the acceleration graph does not – by itself – provide information about the direction of the object's motion, because the question of above-or-below the horizontal axis for the velocity graph cannot be answered when the acceleration graph only gives the v vs. t graph's shape.

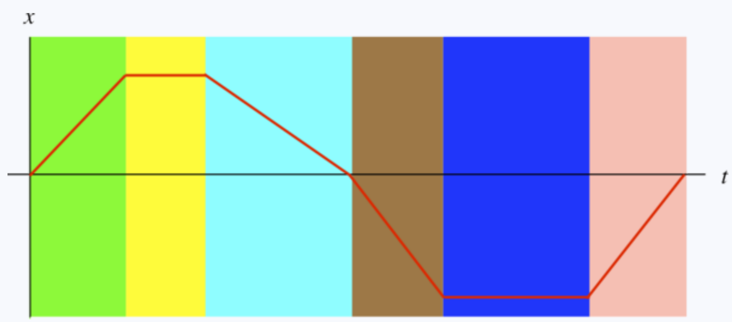
Q4: Is the object speeding-up or slowing down?

This is probably the trickiest question of all, because it doesn't have a direct correlation to the value or slope of any of the graphs. To make this determination, you actually need *two* pieces of information – the directions of both the velocity and the acceleration. This is because if the object is accelerating in the same direction that it is moving, then it is speeding up, and if it is accelerating in the opposite direction as the direction of motion, then it is slowing down. We therefore cannot determine the answer to this question from the acceleration graph alone, because that graph by itself does not provide the direction of motion (the function $v(t)$ associated with this acceleration could be above or below the horizontal axis anywhere). We can determine speeding-up/slowing-down from the v vs. t graph, by comparing the slope of the graph with the value of the graph at the same point. If they have the same sign, then the acceleration is in the same direction as the velocity, and it is speeding up. If they are opposite, then it is slowing down. But there is a simpler, physical way to make this determination: If the v vs. t graph at the point in question is heading closer to the horizontal axis, then its velocity is heading toward zero, and it is slowing down, while if it is heading away, it is speeding up. Naturally horizontal parts of the the v vs. t graph represent motion in which the object is neither speeding up nor slowing down.

Making this determination from the x vs. t graph is even more challenging. Clearly if a section of the x vs. t graph is a straight line, then the velocity is constant, and the object is neither speeding up nor slowing down. So what about when $x(t)$ is curved? The trick to use here is to determine if continuing this curve will eventually cause the graph to go horizontal (i.e. reach a max or a min), or vertical. If it is the former, then the object is slowing (a horizontal slope is stationary), and the latter is speeding up. Note that both of these can occur for either concave or convex curves, for positive or negative slopes, and above or below the horizontal axis.

Example 1.5.1

For the position vs. time graph of an object moving in one dimension given below answer each of the four questions given above for every segment of time indicated by the different colors.



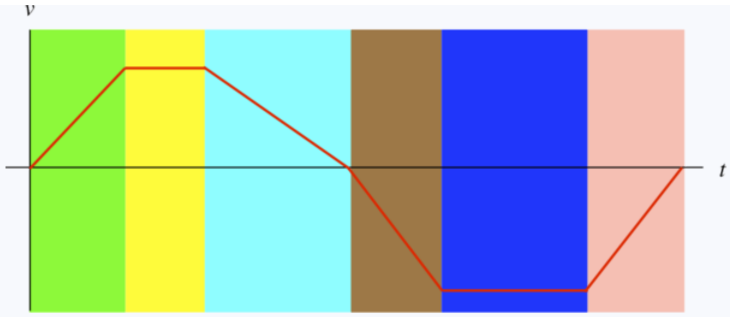
Solution

Where is the object?	*At rest or moving?*	*Moving which way?*	*Speeding up or slowing down?*
+ side of origin	***moving***	***+x direction***	***neither… constant speed***
+ side of origin	***at rest***	***not moving***	***neither… constant speed***
+ side of origin	***moving***	***–x direction***	***neither… constant speed***
– side of origin	***moving***	***–x direction***	***neither… constant speed***
– side of origin	***at rest***	***not moving***	***neither… constant speed***
– side of origin	***moving***	***+x direction***	***neither… constant speed***

Example 1.5.2

For the velocity vs. time graph of an object moving in one dimension given below:

a. *Answer each of the four questions given above for every segment of time indicated by the different colors.*
b. *Sketch the position vs. time and acceleration vs. time graphs associated with this same motion. Assume that the object was at the origin at time $t = 0$.*

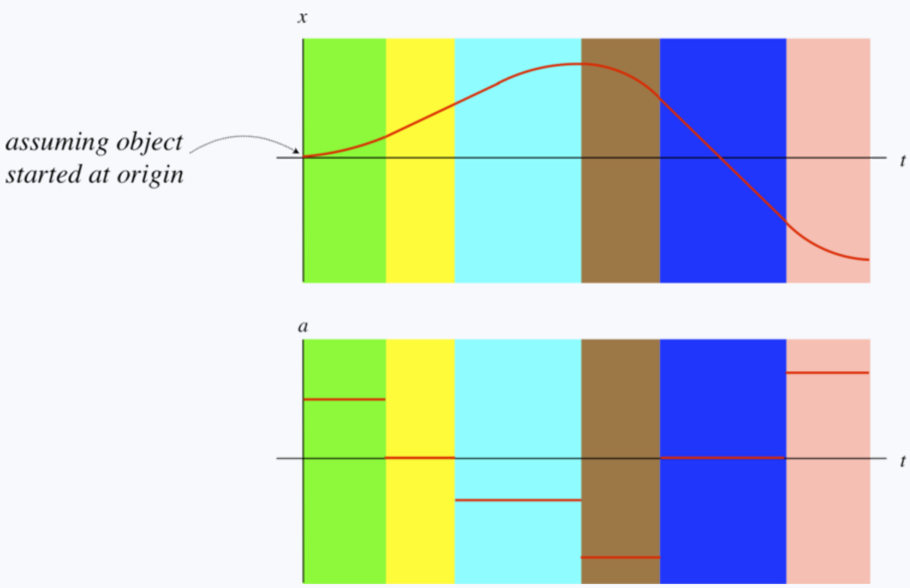


Solution

a.

Where is the object?	*At rest or moving?*	*Moving which way?*	*Speeding up or slowing down?*
don't know	***moving***	***+x direction***	***speeding up***
don't know	***moving***	***+x direction***	***neither… constant speed***
don't know	***moving***	***+x direction***	***slowing down***
don't know	***moving***	***–x direction***	***speeding up***
don't know	***moving***	***–x direction***	***neither… constant speed***
don't know	***moving***	***–x direction***	***slowing down***

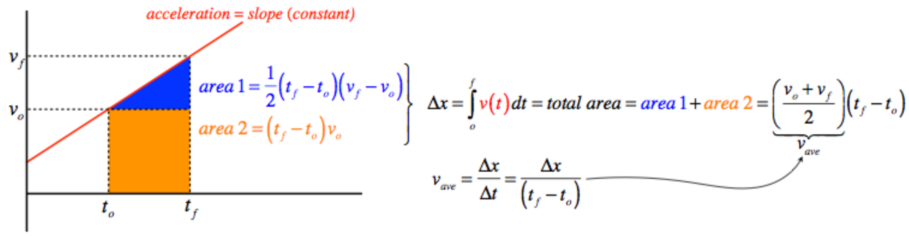
b.



Integrating Using Graphs

We have already seen that we can derive equations of motion for $v\left(t\right)$ and $x\left(t\right)$ by integrating their derivatives, and we know that integrals of functions equal the areas under the curves those functions represent, so we can use this knowledge to tie together these two facts. If we are given the graph of a motion, we can compute the area under the curve between the starting and ending points to get a definite integral, and therefore the change between the starting and ending values. So for example, if we again assume constant acceleration, a velocity-vs-time graph is a straight line whose slope is the acceleration. The area under this line from the starting time to the ending time will be the displacement between these two times (note: we still don't know the positions for these times, only the change in positions). This actually demonstrates the average velocity relation we found earlier:

Figure 1.5.1 – Area Under Velocity Curve Is Displacement



Notice that it is vital that the acceleration is constant for this formula for average velocity to come out, because the area under the curve involves the area of a triangle that requires a straight line on top. Of course, the average velocity could *accidentally* come out to equal the arithmetic average of the starting and ending velocities when the acceleration is not constant (if the area under the curved graph happens to equal the area under the straight line graph between the same two points), but t we cannot rely on such coincidences when solving problems. Moreover, this means that we cannot assume the converse – if the arithmetic mean of a starting and ending velocity equals the average velocity, we cannot conclude that the acceleration was constant over that time interval.

1.6: Motion in Multiple Dimensions

Now that we have mastered the simplest form of motion, it's time to branch out to more general cases. No longer will the motion of objects be constrained to move along a straight line! Of course, this means that we can no longer allow simple positive and negative values to tell us about directions – we need to introduce vectors into the story. Fortunately, we have built our vector mathematics tools to the point where we can make use of them here.

Position and Displacement

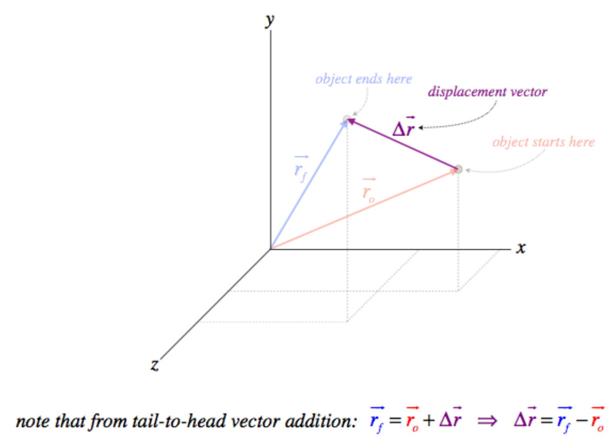
Without the luxury of being able to describe the position of an object with a single (positive or negative) value, we now have to do so in terms of something called a *position vector*. If we assume a coordinate system is in place, the position of the object can be described by its coordinates, x, y, and z. These also happen to be the components of the position vector, which we define as the vector that points from the origin to the point in space:

$$\overrightarrow{r} = x\hat{i} + y\hat{j} + z\hat{k} \tag{1.6.1}$$

If an object moves from one position to another, then clearly it is displaced, and we can describe this displacement in terms of the change of the position, as we did for straight-line motion. The only difference is that here we create a *displacement vector*:

$$\Delta\overrightarrow{r} = \overrightarrow{r}_f - \overrightarrow{r}_o = (x_f - x_o)\,\hat{i} + (y_f - y_o)\,\hat{j} + (z_f - z_o)\,\hat{k} \tag{1.6.2}$$

Figure 1.6.1 – Displacement Vector



displacement vector points from point where the object started to the point where it ends

Velocity

We follow the same process as we did with straight-line motion to determine average and instantaneous velocity vectors. Namely, we define the average and instantaneous velocities in terms of the rate of change of the displacement:

$$\overrightarrow{v}_{ave} = \frac{\Delta\overrightarrow{r}}{\Delta t} = \frac{\Delta x}{\Delta t}\hat{i} + \frac{\Delta y}{\Delta t}\hat{j} + \frac{\Delta z}{\Delta t}\hat{k} \tag{1.6.3}$$

$$\overrightarrow{v} = \frac{d\overrightarrow{r}}{dt} = \frac{dx}{dt}\hat{i} + \frac{dy}{dt}\hat{j} + \frac{dz}{dt}\hat{k} \tag{1.6.4}$$

While this vector formula might appear to imply that the direction of the velocity vector is the same as the direction of the position vector, it's important to understand that in fact the direction of the velocity vector is the same as the direction of *the infinitesimal change* of the position vector. Let's look at an example that makes this clear...

Consider a particle moving in the $x-y$ plane. Its time-dependent position vector can be expressed in terms of its time-dependent components as:

$$\overrightarrow{r}\,(t) = x\,(t)\,\hat{i} + y\,(t)\,\hat{j}$$

For the sake of this example, let's suppose that the particle's position components have the following time dependences:

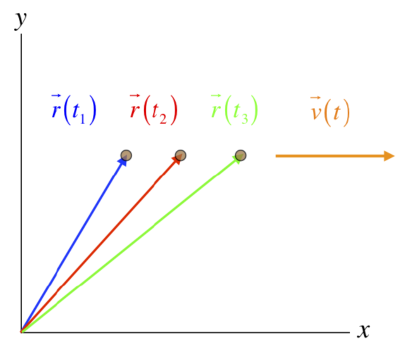
$$x(t) = \left(3.0\frac{m}{s}\right)t \qquad y(t) = 5.0m$$

Now we can calculate the velocity vector by taking the time derivative of the position vector. The unit vectors $\hat{i}$ and $\hat{j}$ don't change with time, so the derivative is simply:

$$\overrightarrow{v}(t) = \frac{d}{dt}\overrightarrow{r} = \left[\frac{d}{dt}x(t)\right]\hat{i} + \left[\frac{d}{dt}y(t)\right]\hat{j} = \left(3.0\frac{m}{s}\right)\hat{i}$$

The position vector changes in both magnitude and direction, while the velocity vector does neither. This demonstrates that the formula that relates the position and velocity vectors might appear to imply some kinship between these vectors, but in fact the presence of the derivative removes the possibility of generalizations like them pointing in the same direction. A diagram of this example for three different times should help visualize this difference:

Figure 1.6.2 – Comparison of a Position Vector and the Related Velocity Vector



Acceleration

Naturally acceleration works the same way as velocity in terms of the calculus:

$$\overrightarrow{a}_{ave} = \frac{\Delta\overrightarrow{v}}{\Delta t} = \frac{\Delta v_x}{\Delta t}\hat{i} + \frac{\Delta v_y}{\Delta t}\hat{j} + \frac{\Delta v_z}{\Delta t}\hat{k} \tag{1.6.5}$$

$$\overrightarrow{a} = \frac{d\overrightarrow{v}}{dt} = \frac{dv_x}{dt}\hat{i} + \frac{dv_y}{dt}\hat{j} + \frac{dv_z}{dt}\hat{k} \tag{1.6.6}$$

Alert

Notice that if we confine ourselves to motion in just one dimension (say the x-axis), then we get exactly the equations we obtained in Section 1.3. So what motion in three dimensions amounts to is additional bookkeeping – we have three separate sets of kinematic relations to keep track of, instead of only one.

Splitting Direction and Magnitude – Velocity

Alert

Students occasionally struggle with what follows, perhaps because the idea of a unit vector is still a bit abstract to them. If you find yourself in this situation, you should spend a little extra time to become comfortable with these ideas, as they are central to everything that follows in this course.

We know that whenever we take the derivative of a vector like position (to obtain velocity) or velocity (to obtain acceleration), a non-zero result comes about when that vector is changing in some way (magnitude, direction, or both). Let's see what happens if we split these two vector properties up and treat them separately...

The unit vectors we have encountered to this point have been exclusively the *cartesian unit vectors* – those that point in the x, y, and z directions: $\hat{i}$, $\hat{j}$, and $\hat{k}$. When we first encountered unit vectors, we saw that a vector can be written as a product of its magnitude and the unit vector that points in its direction (Equation 1.1.2). If this vector happens to be changing direction over time, then unlike the cartesian unit vectors, this unit vector changes over time. As a first example, let's look at what this means for the position vector (the derivative of which is the velocity vector). We know how to express the position vector in terms of cartesian unit vectors (Equation 1.6.1), but in terms of its magnitude and directional unit vector, it is written in the same manner as Equation 1.1.2}:

$$\overrightarrow{r} = r\,\hat{r} \tag{1.6.7}$$

Recall that when a vector's variable name (in this case, r) is written without the arrow over it, it refers to the magnitude of the vector. In this case, this magnitude is, in terms of the cartesian components:

$$r = \sqrt{x^2 + y^2 + z^2} \tag{1.6.8}$$

Referring back to Figure 1.6.2, we see a case where both the magnitude and direction of the position vector are changing. Therefore when we compute the velocity vector, the derivative will act on both the magnitude and on the unit vector, and it turns out that the usual product rule works perfectly well here:

$$\overrightarrow{v} = \frac{d}{dt}\overrightarrow{r} = \frac{d}{dt}(r\hat{r}) = \frac{dr}{dt}\hat{r} + r\frac{d\hat{r}}{dt} \tag{1.6.9}$$

Alert

This was just stated above, but it bears repeating... The cartesian unit vectors $\hat{i}$, $\hat{j}$, and $\hat{k}$ don't change with time (they always point in the x, y, and z directions), but other unit vectors (like $\hat{r}$) can and do change with time, so their derivatives don't automatically vanish. It is the derivative of this unit vector that determines how that vector's direction is changing.

Okay, so let's consider the following questions:

Example 1.6.1

How is Equation 1.6.9 affected when the object happens to be moving either directly toward or directly away from the origin?

Solution

If the object is moving directly toward or away from the origin, then the position vector (whose tail is at the origin and head is at the object) is always pointing in the same direction, but never changes direction. Therefore the second term in that equation vanishes, leaving only the first term.

Example 1.6.2

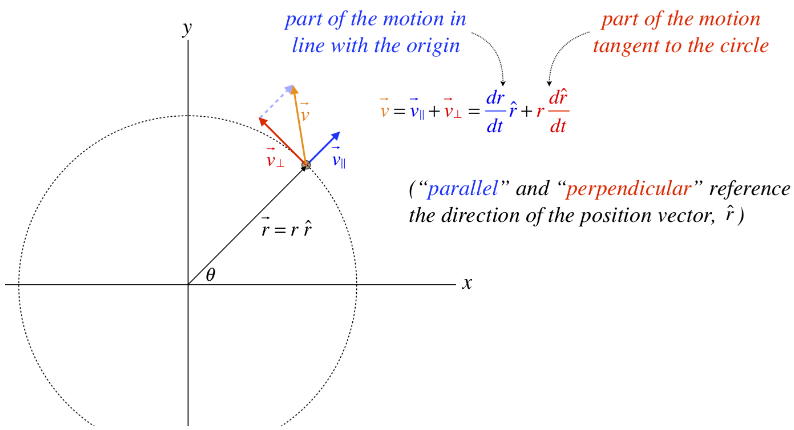
How is Equation 1.6.9 affected when the object happens to be moving such that its distance from the origin never changes?

Solution

If the object's distance from the origin never changes, then the magnitude of the position vector is not changing, which means that the first term vanishes. Clearly in order to move while staying the same distance from the origin, the direction of motion must be changing, so the second term is not zero.

But wait, an object moving such that its distance from a single point never changes must be traveling in a circle (assuming its motion remains in a plane). So this velocity vector is that of circular motion around the origin. A general velocity vector (one in which neither term from the product rule vanishes) can therefore be thought of as a vector sum of a velocity vector that points radially outward from an origin, and one that points tangent to a circle centered at that origin. Geometrically, it looks like this:

Figure 1.6.3 – Parallel and Perpendicular Components



Clearly the derivative of the position unit vector is a new vector that is perpendicular to the position unit vector. We can check to see if this is true, as well as make sense of all this by returning to the easier-to-work-with cartesian unit vector approach. We do this by writing the position vector in polar coordinates. Using θ as the angle in the diagram above, we use trigonometry to break the position vector into x and y components, written in terms of r and θ:

$$\overrightarrow{r} = r\cos\theta\;\hat{i} \;+ r\sin\theta\;\hat{j} \tag{1.6.10}$$

Combining this with Equation 1.6.1 and Equation 1.6.7 gives us the position unit vector in terms of the cartesian unit vectors:

$$\hat{r} = \cos\theta\;\hat{i} \;+ \sin\theta\;\hat{j} \tag{1.6.11}$$

It's easy to confirm that this unit vector has a length of 1, as it should. The claim above is that the time derivative of $\hat{r}$ is perpendicular to $\hat{r}$ itself, which we can now check directly, using our clever tool from Section 1.2 – the scalar product of these two vectors should vanish. Start by computing the derivative of the position unit vector. The cartesian unit vectors have zero derivative, but of course θ can be changing as the object moves, so:

$$\frac{d\hat{r}}{dt} = \frac{d}{dt}\left(\cos\theta\;\hat{i} \;+ \sin\theta\;\hat{j}\right) = -\sin\theta\left(\frac{d\theta}{dt}\right)\;\hat{i} + \cos\theta\left(\frac{d\theta}{dt}\right)\;\hat{j} = \frac{d\theta}{dt}\left(-\sin\theta\;\hat{i} + \cos\theta\;\hat{j}\right) \tag{1.6.12}$$

Now perform the dot product:

$$\hat{r}\cdot\frac{d\hat{r}}{dt} = \left(\cos\theta\;\hat{i} \;+ \sin\theta\;\hat{j}\right)\cdot\left[\frac{d\theta}{dt}\left(-\sin\theta\;\hat{i} + \cos\theta\;\hat{j}\right)\right] = \frac{d\theta}{dt}(-\cos\theta\sin\theta + \sin\theta\cos\theta) = 0 \tag{1.6.13}$$

Example 1.6.3

Show that the time derivative of any unit vector is either zero (as in the case of $\hat{i}$) or is perpendicular to the unit vector itself (as in the case of $\hat{r}$). [Hint: The product rule for the derivative works for the scalar product.]

Solution

Naturally the derivative of the number 1 is zero, and this happens to be the result of a scalar product of a unit vector with itself, so applying the product rule:

$$0 = \frac{d}{dt}(1) = \frac{d}{dt}\left(\widehat{A}\cdot\widehat{A}\right) = \frac{d\widehat{A}}{dt}\cdot\widehat{A} + \widehat{A}\cdot\frac{d\widehat{A}}{dt} = 2\widehat{A}\cdot\frac{d\widehat{A}}{dt}$$

This can only equal zero if the vector resulting from the derivative is zero, or it is perpendicular to $\widehat{A}$.

Splitting Direction and Magnitude – Acceleration

Above we found that a velocity vector can be broken into two components – one parallel to the position vector and one perpendicular to it. The first accounts for velocity in line with the origin, and the second for velocity tangent to a circle around the origin. This has few applications in physics, because typically the choice of origin is arbitrary. But when we extend the use of this vector calculus machinery to acceleration, it gets more interesting and far more useful, as we will see.

If we replace the position vector with the velocity vector and follow the same procedure as above, we get for the acceleration:

$$\overrightarrow{a}=\frac{d}{dt}\overrightarrow{v}=\frac{d}{dt}(v\,\hat{v})=\frac{dv}{dt}\,\hat{v}+v\,\frac{d\hat{v}}{dt} \tag{1.6.14}$$

Once again we see that the product rule separates the derivative into a sum of two vectors – one that is parallel to the original velocity, and one that is perpendicular to it. For future reference, we'll write the two terms this way:

$$\overrightarrow{a}_{\|}\equiv\frac{dv}{dt}\,\hat{v}\qquad\overrightarrow{a}_{\perp}\equiv v\,\frac{d\hat{v}}{dt} \tag{1.6.15}$$

We already know that the acceleration vector is the rate of change of the velocity vector, and that the velocity vector includes the speed and direction of motion, so here we see that the acceleration breaks into two vectors: $\overrightarrow{a}_{||}$, which handles only the change of speed, and $\overrightarrow{a}_{\perp}$, which handles only the change of direction. If only the first term is non-zero, then the object is speeding up or slowing down in a straight line. If only the second term is non-zero, then the object is neither speeding-up nor slowing down, but its direction of motion is changing. We will get a lot of mileage out of this division of labor in the chapters to come.

Example 1.6.4

A particle moves through space with a velocity vector that varies with time according to:

$$\overrightarrow{v}(t)=\alpha\,\hat{i}-\beta t\,\hat{j}\,,$$

where α and β are positive constants. Find the rate at which the ***speed*** *of this particle is changing at time $t=0$. Does this rate remain the same for all later times?*

Solution

Acceleration is the rate of velocity change, which can come in the form of speed change or direction change (or both). This problem asks specifically for the rate of change of the speed. Let's start by computing the acceleration vector:

$$\overrightarrow{a}(t)=\frac{d}{dt}\overrightarrow{v}(t)=-\beta\,\hat{j}$$

One's first inclination might be to state that the magnitude of the acceleration is β, and there is no time dependence, so the speed is changing at this rate at all times. But we must be careful to only consider the part of the acceleration vector that describes speed change, and not direction change. We know that the parallel part of the acceleration deals with speed change, so we need a way to extract this piece from the full acceleration vector. But we know how to extract the part of one vector parallel to another, using the scalar product. Namely:

$$a_{\|}=\overrightarrow{a}\cdot\hat{v}=\overrightarrow{a}\cdot\frac{\overrightarrow{v}}{v}=\left(-\beta\,\hat{j}\right)\cdot\left(\frac{\alpha\,\hat{i}-\beta t\,\hat{j}}{\sqrt{\alpha^2+\beta^2t^2}}\right)=\frac{\beta^2t}{\sqrt{\alpha^2+\beta^2t^2}}$$

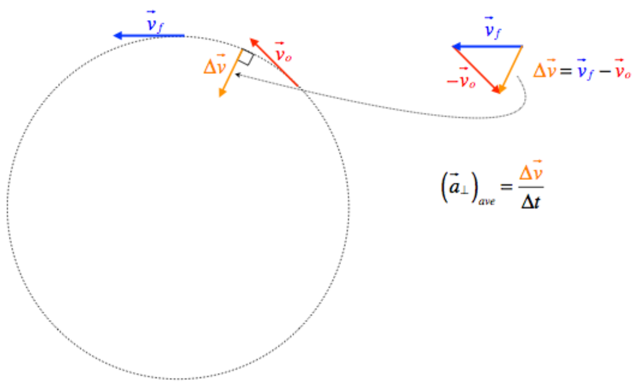
Plugging in $t=0$ reveals that the particle is not speeding up at all at that time. We also see this does not remain true for all values of t. The reason is that the acceleration vector is a constant, and is initially perpendicular to the velocity vector, so at that moment it can only change its direction. When the velocity vector changes direction, this new direction has a component parallel to the unchanging acceleration vector, which means the particle speeds up.

1.7: Examples of 2-Dimensional Motion

Circular Motion

Using what we just derived regarding the parallel and perpendicular components of acceleration, we turn now to the special case of an object traveling in a circle. The parallel part of the acceleration obviously always points tangent to the circle, which narrows it down to two directions at any given point on the circle. If the object is speeding up, then of course the tangent points in the direction of motion, and if it is slowing down, the tangent vector points in the direction opposite to the motion. The perpendicular part must be at a right angle to this tangent, which means it must be toward or away from the center of the circle. Unlike the tangent case, however, both directions are *not* possible. We can see this by considering the average perpendicular acceleration vector over two nearby moments in time:

Figure 1.7.1 – Direction of Centripetal Acceleration



this acceleration is called centripetal, which means "center seeking"

What is the magnitude of this *centripetal acceleration*? Well, it depends upon how fast the object is going (the faster it is moving, the more acceleration is required to turn in the same circle), and the radius of the circle (the acceleration is greater when the radius is smaller). Deriving the magnitude is left as an exercise, but the answer comes out to be:

$$|\overrightarrow{a}_c| = \frac{v^2}{r} \tag{1.7.1}$$

Sometimes circular motion is the result of something rotating. For example, a bug on the outer rim of a rotating turntable travels in a circle, and therefore experiences centripetal acceleration. Well, when we deal with rotating objects we often know only the rate of its rotation (say in rpms), and we have to translate into linear motion to know the speed. There is a simple and standard way to do this.

Digression: Radians

If we are talking about rotational motion, we need to discuss how we measure such motion. We clearly don't measure the speed of a spinning top or turntable in terms of meters per second, but rather how much it turns in a period of time. How does one measure the angle through which something turns? One way is to divide the full circle up into 360 equally-sized pie slices. The magnitude of one of these angles we call a 'degree.' But really this division is arbitrary. So why was 360º selected? Well, given that we often have to deal with slices of pie, we can avoid having to use fractional degrees if we select a number with lots of factors, and 360 certainly fits the bill – it is divisible by 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 18, 20, 24, 30, 36, 40, 45, 60, 120, and 180.

But there is no reason at all that divisibility needs to be our only criterion. In fact, we don't even need to divide the circle into an integral number of pieces. For example, we could divide it into 7.5 pieces and call the size of each piece "1 flibber." We can even translate between different systems of units: $1^f = 48^o$.

But there is another criterion that leads to a definition of a measurement of angles that is different from degrees. Suppose we want a simple translation from angle to arclength (the distance traveled along the circular curve subtended by that angle). We know that traveling around an entire circle requires a journey of a distance equal to 2π *times the radius of the circle. So going around some fraction of a circle requires a journey equal to that fraction times* $2\pi r$*. So if we divide the circle into an uneven* *number of pieces such that 2π of these pieces fit into the circle, then in these units you can calculate the arclength by just multiplying the angle measured in those units (called radians) by the radius of the circle:*

$$s=r\theta,\ \theta\,\text{measured in radians} \tag{1.7.2}$$

If we want to translate between the speed that something is going around the circle to the angular speed at which the slice of the pie is being swept-out by this motion, we need only take a derivative:

$$v=\frac{ds}{dt}=\frac{d}{dt}(r\theta)=\cancelto{0}{\frac{dr}{dt}}\theta+r\frac{d\theta}{dt}=r\omega, \qquad \omega\equiv\frac{d\theta}{dt}=rate\ of\ rotation\ in\ \frac{rad}{s} \tag{1.7.3}$$

This gives us an alternative way of expressing the centripetal acceleration:

$$|\vec{a}_c|=\frac{v^2}{r}=\frac{(r\omega)^2}{r}=r\omega^2 \tag{1.7.4}$$

Example 1.7.1

A bead is threaded onto a circular hoop of wire which lies in a vertical plane. The bead starts at the bottom of the hoop from rest, and is pushed around the hoop such that it speeds up at a steady rate. Find the angle that the bead's acceleration vector makes with the horizontal when it gets back to the bottom of the hoop.

Solution

The tangential motion of the bead is in one dimension, so we can use the usual kinematics equations to describe its motion along the circle. Let's call the radius of the circle R and the final velocity v. The tangential acceleration is constant, the bead starts from rest, and the bead travels one circumference, so using Equation 1.4.3, we get:

$$2a\Delta x={v_f}^2-{v_o}^2 \quad\Rightarrow\quad a_{\parallel}=\frac{\left(v^2-0^2\right)}{2\left(2\pi R\right)}=\frac{v^2}{4\pi R}$$

The centripetal acceleration is toward the center of the circle, so it points upward and its magnitude is simply:

$$a_{\perp}=\frac{v^2}{R}$$

The tangent of the angle that the full acceleration vector makes with the horizontal is the vertical component divided by the horizontal component, so:

$$\theta=\tan^{-1}\left(\frac{\frac{v^2}{R}}{\frac{v^2}{4\pi R}}\right)=\tan^{-1}(4\pi)=85^o$$

Projectile Motion

For circular motion, we have the components of velocity changing in tandem in a specific manner to keep the path circular. Another – actually simpler – form of motion involves only a single fixed coordinate component of velocity changing, while the other components involve no change in velocity. What I am alluding to here is *projectile motion*, which comes about because the vertical component of motion is subject to constant acceleration (as we discussed when we talked about free-fall), while the horizontal component is unaffected by gravity's influence. This kind of motion is only one small step from the free-fall we are already familiar with, in that it includes a *totally independent* horizontal component of motion that incorporates (to the extent that air resistance can be ignored) no acceleration. As simple as this sounds, a couple of examples muddy the waters a bit, and sorting them out is very instructive:

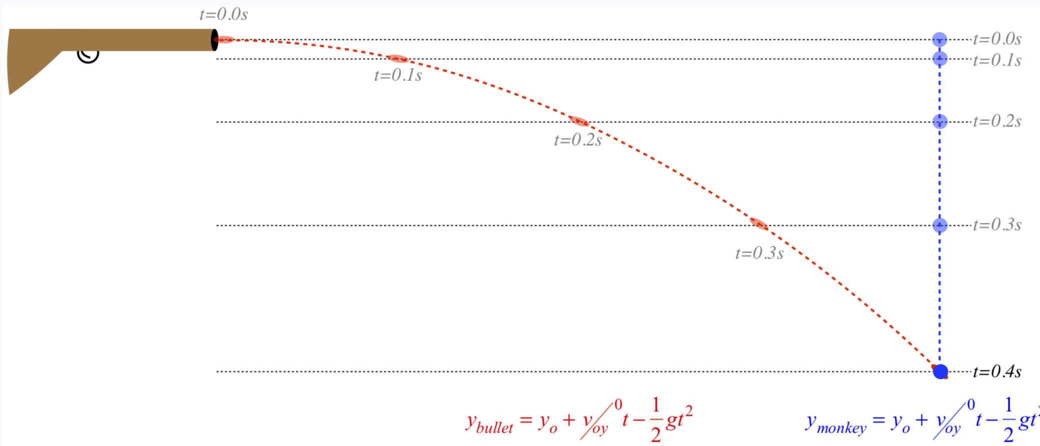
Example 1.7.2

A hunter climbs a tree and fires a bullet directly at a monkey that is hanging from a branch of another tree at precisely the same height as the barrel of the hunter's gun. The instant the bullet leaves the gun, the monkey lets go of the branch. Ignoring air resistance (and the size of the monkey - assume it is very small), what is the fate of the monkey?

A. *The monkey will be hit by the bullet.*
B. *The bullet will pass beneath the monkey.*
C. *The bullet will fly over the monkey's head.*
D. *Whether the bullet flies over the monkey's head or passes beneath it depends upon how fast the bullet is moving when it leaves the barrel of the gun.*
E. *What kind of jerk shoots a monkey?*

Solution

A (and E). The vertical motion of the monkey is independent of the horizontal motion, so in equal time spans, the bullet falls the same distance as the monkey. Since they started at the same level, they remain at the same level at all times, including when the horizontal position of the bullet equals the horizontal position of the monkey.



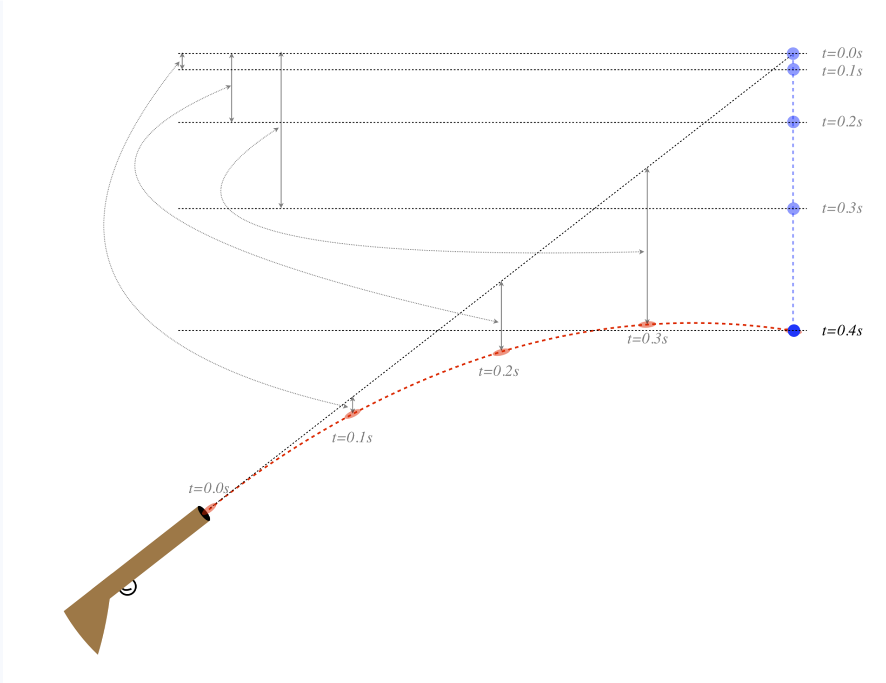
Example 1.7.3

After falling out of the tree the last time he tried to shoot a monkey (his gun misfired), the hunter now decides to shoot a monkey from the ground. He is aiming upward at an angle, and is assuming the monkey will again let go of the branch just as the bullet is on its way. How should the hunter aim this time, if he is to bag his simian prize?

A. *He should aim right at the monkey.*
B. *He should aim above the monkey.*
C. *He should aim below the monkey.*
D. *Unlike the level-shot case, where he should aim this time does depend upon how fast the bullet is coming out of the gun.*
E. *The jerk should just aim at himself.*

Solution

A (and E). If there was no gravity, the bullet would follow a straight line to the monkey. With gravity acting straight down, the amount that the bullet drops below that straight line is the same that an object starting from rest on that straight line falls in an equal time. So the bullet and the monkey remain the same distance from the straight line at all times. When the bullet's horizontal position equals the monkey's horizontal position, they will coincide.



The only difference between this example and the previous one is that in the previous case, the line joining the barrel of the gun and the target was horizontal. Still, not everyone may be as convinced in this case as in the previous one, so let's do the math...

Suppose there is no gravity. The path that the bullet takes (y as a function of x) can be written down pretty easily. If the point where the bullet exits the barrel is chosen to be the origin, then the straight line to the monkey has a slope equal to the ratio of the vertical and horizontal components of velocity:

$$y = mx + b = \left(\frac{v_{oy}}{v_{ox}}\right) x$$

Now suppose there is gravity. We have separate horizontal and vertical equations of motion. Again, with the bullet starting at the origin, we have:

$$x = v_{ox} t$$
$$y = v_{oy} t - \frac{1}{2} g t^2$$

Now solve for t in the first equation and plug it into the first term of the second equation to get:

$$y = \frac{v_{oy}}{v_{ox}} x - \frac{1}{2} g t^2$$

Comparing this with the first equation above, we see that the y value would follow the same straight line if not for the second term, and the amount that the height of the bullet y is decreased from that line after a time t is exactly the same distance that the monkey falls from that line in the same time.

With the number of variables and constants involved in projectile motion problems, there are countless ways to construct problems. There is no substitute for independent thinking and creativity, but the steps given below provide a good starting point for solving these kinds of problems.

- Draw a picture, labeling it as completely as you can, using information you have been given. Then spend some time thinking about what is happening – put yourself into the situation.

Alert

*While this is given as a step for projectile problems, this is actually how you should start **every** physics problem!*

- Pick an x, y origin as well as $+x$ and $+y$ directions. Often for projectile problems up is chosen as the positive directions (making the acceleration due to gravity a negative value), but this is by no means required. What is important is that you use the positive direction consistently throughout the constants and variables in the equations.
- Break any initial velocities into components along the x and y directions.
- Write down the equations of motion, circling quantities that you know, and underlining the number you are looking for. If you have too many un-circled quantities for the number of equations available, you cannot yet do the algebra, so you'll need to review the statement of the problem for any values concealed in the language of the problem (if you just scan a problem for numbers without carefully reading it, you will miss these).
- Solve the algebra and reconstitute components back into vectors, if necessary.
- Briefly check to see if the answer makes sense.

Alert

*This is actually how you should end **every** physics problem!*

One thing in projectile motion that is a useful tool is known as the range equation. This was of particular use to military firing cannonballs or (farther back in history than that), catapults. This equation relates the distance that a projectile will fly assuming it lands at the same vertical position that it started, given the starting speed of the projectile and the starting angle. Let's treat finding this equation as if it was a projectile motion problem given to us, and follow the procedure outlined above

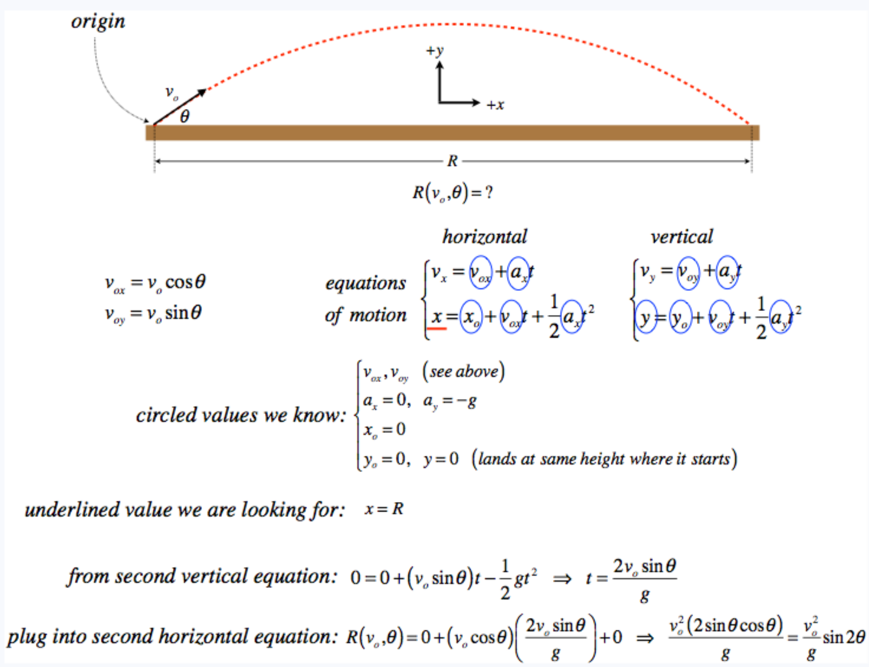
Problem: A cannonball is fired at an angle θ up from the horizontal at a speed of vo. Ignoring air resistance, how far does the cannonball travel before landing back on ground level with where it was fired?

Example 1.7.4

A cannonball is fired at an angle θ up from the horizontal at a speed of v_o. Ignoring air resistance, how far does the cannonball travel before landing back on ground level with where it was fired?

Solution

The diagram below labels the origin and the + directions. We are looking for R (referred to as the "range") in terms of the initial speed v_o of the projectile and the launch angle θ, which we treat as known values. The rest of the procedure is given below.



We can do other things with this result. For example, suppose we wanted to fire the cannonball as far as possible. Obviously we want it to start at the fastest possible speed, but given that, what angle should we aim at? Let's solve it! We want to maximize the range R in terms of our choice of angle. This is a straightforward calculus problem. Take the derivative and set it equal to zero to find the maximum:

$$0=\frac{dR}{d\theta}=\frac{v_0^2}{g}\frac{d}{d\theta}\left(\cos 2\theta\right)(2) \;\Rightarrow\; \boxed{\theta=45^\circ} \tag{1.7.5}$$

This really shouldn't be too surprising, based on a simple symmetry argument: If we shoot the projectile with a really steep angle, it goes nearly straight up and doesn't travel very far. We can send the projectile the same distance by choosing a sufficiently *shallow* angle. That means that every landing point in range of the launcher has two possible distinct values for launch angle, *except for the one corresponding to the longest range*. It makes sense that complimentary launch angles reach the same landing point: 89^o and 1^o, 52^o and 38^o, and so on, which means that the self-complimentary angle of 45^o hits the maximum range.

The following example is a challenging application of this principle, and gives an idea of the rich diversity of problems that can be conceived for projectile motion:

Example 1.7.5

Two warlords aim identical catapults (i.e. they both release rocks at the same speed) at each other, with both of them being at the same altitude. The warlords have made the necessary computations to crush the other, and fire their catapults simultaneously. Amazingly, the two stones do not collide with each other in mid-air, but instead the stone Alexander fired passes well below the stone that Genghis shot. Genghis is annihilated 8.0s after the catapults are fired, and Alexander only got to celebrate his victory for 4.0s before he too was destroyed.

a. *Find the maximum height reached by each of the rocks.*
b. *Find the amount of time that elapses from the launch to moment that the rocks pass each other in the air.*
c. *Find the angles at which each warlord fires his rock.*

Solution

a. The time it takes a rock to travel to its peak height and back down again is equal to twice the time it takes to travel down from its peak height. Traveling down from its peak height, it starts with zero initial velocity, so we can calculate the height immediately for each rock:

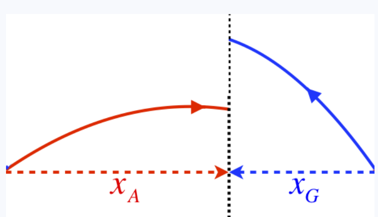
$$h_A=\frac{1}{2}g\left(\frac{t_A}{2}\right)^2=\frac{1}{2}\left(9.8\frac{m}{s^2}\right)\left(\frac{8.0s}{2}\right)^2=\boxed{78.4m}$$

$$h_G=\frac{1}{2}g\left(\frac{t_G}{2}\right)^2=\frac{1}{2}\left(9.8\frac{m}{s^2}\right)\left(\frac{8.0s+4.0s}{2}\right)^2=\boxed{176.4m}$$

b. The x–components of the velocities of the rocks never change, and since it takes 12s for Genghis's rock to travel the same horizontal distance as Alexander's rock traveled in 8s, Alexander's rock is traveling in the x-direction at a rate 1.5 times as great as Genghis's rock is traveling in the x–direction. When they are at the same x–position (passing each other), the distance each has traveled is each one's velocity times the time we are looking for, and we can express both of these distances in terms of the x–component of Genghis's rock using the ratio described above:

$$\begin{aligned}&x_A=v_{Ax}t, \quad v_{Ax}=1.5v_{Gx} \quad\Rightarrow\quad x_A=1.5v_{Gx}t\\&x_G=v_{Gx}t\end{aligned}$$

Since the rocks travel from both ends and are now at the same horizontal position, the sum of the distances they travel equals the total separation of the two warlords. This allows us to calculate the time:



$$x_A+x_G=x_{tot}=v_Gt_{tot} \quad\Rightarrow\quad (1.5v_G)\,t+v_Gt=v_G\,(12.0s) \quad\Rightarrow\quad t=\frac{12.0s}{2.5}=\boxed{4.8s}$$

c. Clearly there are two different angles that will result in the rock traveling the same distance. One can see this from the range equation, but from a physical standpoint, this happens because one rock spends less time in the air but has a greater x–velocity, while the other spends more time in the air with a smaller x–velocity. To spend 1.5 times as long in the air, Genghis's rock needs to start with 1.5 times as much vertical component of velocity as Alexander's rock. This means that the ratios of the x and y components of the two rock velocities are inverses of one another, which means that the two angles are complimentary (i.e. $\theta_A = 90^o - \theta_G$ *). But the total speeds of the rocks are the same, so:*

$$\left.\begin{array}{l} v_{Ax} = v_o\cos\theta_A = v_o\cos(90^o-\theta_G) = v_o\sin\theta_G \\ v_{Gx} = v_o\cos\theta_G \end{array}\right\} \quad\Rightarrow\quad \frac{v_{Ax}}{v_{Gx}} = 1.5 = \frac{\sin\theta_G}{\cos\theta_G} \quad\Rightarrow\quad \left\{\begin{array}{l} \theta_G = \tan^{-1}1.5 = \boxed{56.3^o} \\ \theta_A = 90^o - \theta_G = \boxed{33.7^o} \end{array}\right.$$

1.8: Relative Motion

Reference Frames

Our last topic for motion in multiple dimensions relates what different observers of the same motion measure for velocities. Let's start with the following simple example:

Ann and Bob are traveling on a train together. The train is traveling north at 60 mph, and while Bob remains in place on the train, Ann runs south through the dining car at 10 mph. Bob sees Ann traveling south at 10 mph, but Ann & Bob's mutual friend Chu, who is off the train and looking through the windows, sees Ann moving *north* at 50 mph. Both Bob and Chu are witnessing the same event, but they are doing so from do distinctly different perspectives, which we call *reference frames*. As a result of being in different reference frames, Bob and Chu make different measurements of Ann's velocity vector (they disagree on both the magnitude and direction of her motion).

If we can clearly describe how the two reference frames are related to one another, we can translate between Bob's measurements and Chu's by doing the proper mathematics. In the example above the mathematics is intuitive, but we will want a systematic way of doing it for more complicated situations, such as when the motions are not along the same line. It shouldn't be surprising that the way to do this is to bring vectors into the conversation.

Relative Velocity Vectors

We begin by introducing some language. When an observer – who we will call "A" – in a given reference frame measures the velocity vector of an object (or another frame) – which we will call "B" – we express this vector in words and symbols in this way:

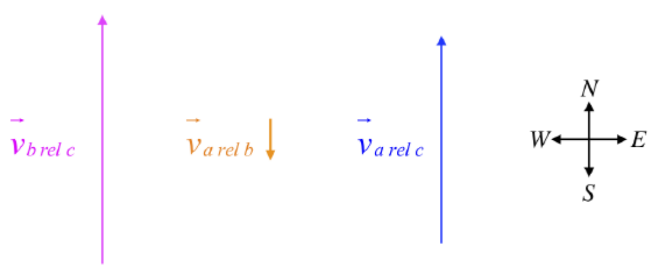
$$\text{" velocity of } B \text{ relative to } A\text{ "} \quad \Longleftrightarrow \quad \overrightarrow{v}_{B\;rel\;A} \tag{1.8.1}$$

Let's see if we can put the above example into this language. There are three entities here. Two are frames and on is a moving object. The moving object is Ann, and she is being observed by Bob, in the reference frame of the train, and Chu, in the reference frame of the earth. In the example, we expressed three different relative velocity vectors:

$$\begin{aligned} \text{" velocity of Bob relative to Chu "} \Longleftrightarrow \overrightarrow{v}_{b\;rel\;c} &= (60mph)\,\widehat{north} \\ \text{" velocity of Ann relative to Bob "} \Longleftrightarrow \overrightarrow{v}_{a\;rel\;b} &= (10mph)\,\widehat{south} \\ \text{" velocity of Ann relative to Chu "} \Longleftrightarrow \overrightarrow{v}_{a\;rel\;c} &= (50mph)\,\widehat{north} \end{aligned} \tag{1.8.2}$$

Let's represent these three vectors as arrows beside each other in a diagram:

Figure 1.8.1 – Relative Velocity Vectors



The first thing we notice when we look closely at these is that our intuitive understanding of the original statement of the situation can be represented as a vector addition. Placing the the tail of the first vector at the head of the second vector, we find that the third vector can connect the open tail to the open head. In other words, we can express the result of the above example as a vector addition:

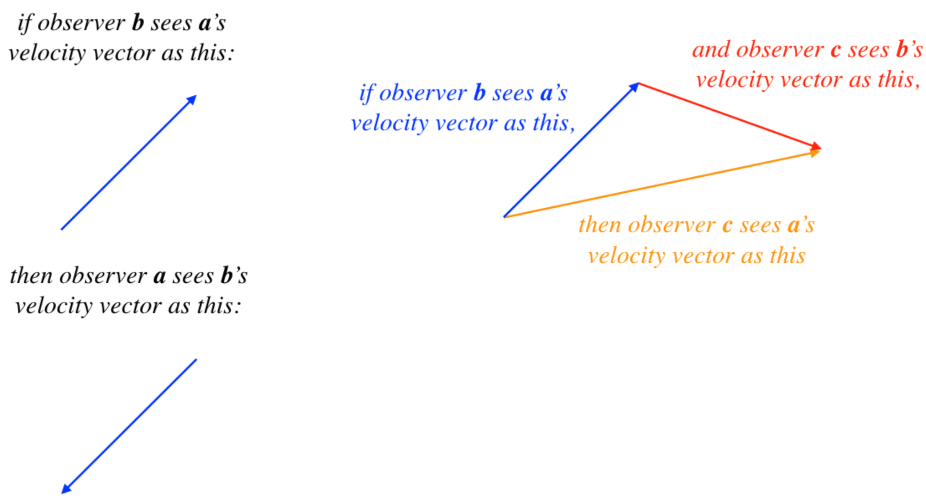
$$\overrightarrow{v}_{a\;rel\;b} + \overrightarrow{v}_{b\;rel\;c} = \overrightarrow{v}_{a\;rel\;c} \tag{1.8.3}$$

Note the ordering of the frames here is like a chain connecting Ann to Chu through Bob: Ann relative to Bob, then Bob relative to Chu, gives Ann relative to Chu. It turns out that this vector equation works not only when the velocities lie along a line, but also when they do not. For example, we can use the same vector equation if Ann were walking *across* the train (perpendicular to its motion.

There is one other feature of these relative velocity vectors that we will need, and that is reversing the perspective. In the case above, we have that Ann is moving 10 mph south relative to the Bob, but we can also talk about how Ann sees Bob moving *relative to her*. Bob starts off south of her, and as she runs by him, he ends up north of her. Therefore from Ann's perspective, Bob is moving north at 10 mph. So there is a simple way to alter a relative vector to reverse the perspective of reference frames: Switch the two frames in the subscript, and reverse the direction of the vector (i.e. multiply the original vector by –1). Here is a summary of these two rules:

Figure 1.8.2 – Summary of Relative Velocity Rules

$\vec{v}_{a\ relative\ to\ b} = -\vec{v}_{b\ relative\ to\ a}$ $\qquad$ $\vec{v}_{a\ relative\ to\ b} + \vec{v}_{b\ relative\ to\ c} = \vec{v}_{a\ relative\ to\ c}$



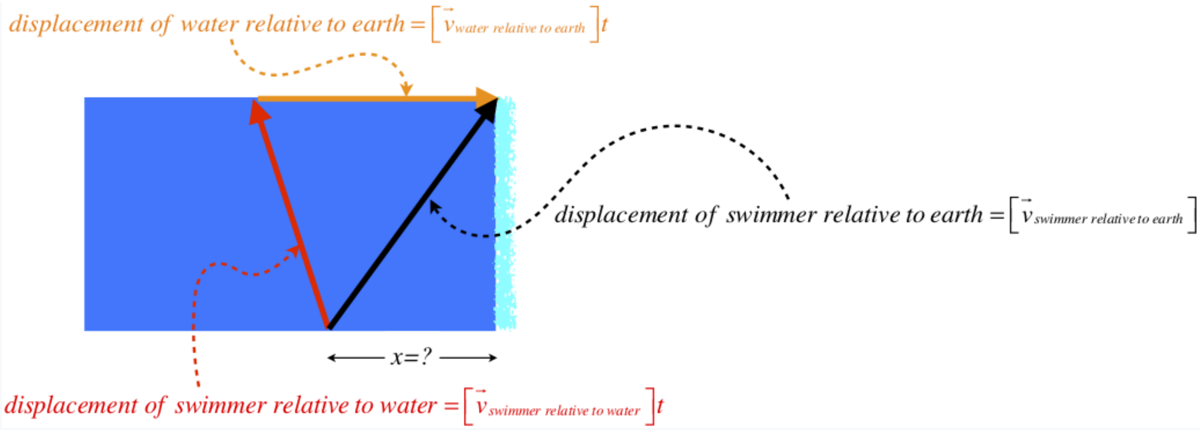
Example 1.8.1

You stand on the bank of a river, contemplating swimming across, but the place where you hope to cross is close to a dangerous waterfall. When you look at the speed of the river, you estimate that it is about the same speed as you are able to swim. You realize that you can only swim so far in the cold water at this speed before your muscles shut down, and in still water you estimate that this distance is about $100m$. The width of the river is about $80m$.

a. *Find the minimum distance that you must start upstream of the waterfall in order to not be swept over it.*
b. *If the river flows west-to-east and you start on its south shore, compute the direction in which you must swim in order to get safely across if you leave from the starting point computed in part (a).*

Solution

a. Clearly to minimize the distance upstream that you need to start, you must swim with a component of your velocity relative to the river being upstream. The more you are able to turn yourself upstream, the less you will float downstream, and the closer you can start to the waterfall. But there is a limit to how far you can swim relative to the water, so your angle with the river must be such that when you reach your limit relative to the river, you reach the other side. The velocities are all constant and the time spans are all equal, so they are proportional to the displacements, which we can draw:



We are given that the speed of the river relative to the earth is the same as the speed of the swimmer relative to the water, so we'll call that quantity v, and the width of the river (which we know), we'll call w. From the pythagorean theorem we can get the distance swum upstream against the current:

$$distance\ swum\ against\ water = \sqrt{(vt)^2 - (w)^2}$$

The distance the water moves downstream relative to the earth is clearly vt, so the total distance the swimmer moves downstream is:

$$x = vt - \sqrt{(vt)^2 - (w)^2}$$

But we actually know the value of vt, because it is the maximum distance that the swimmer can go in the water. Plugging in all the values therefore gives our answer:

$$x = (100m) - \sqrt{(100m)^2 - (80m)^2} = \boxed{40m}$$

b. The angle is easy to determine, since we know the length of the displacement vector of the swimmer relative to the water and the width of the river:

$$\cos\theta = \frac{w}{vt} = \frac{80m}{100m} \quad \Rightarrow \quad \theta = \boxed{37^o west\ of\ north}$$

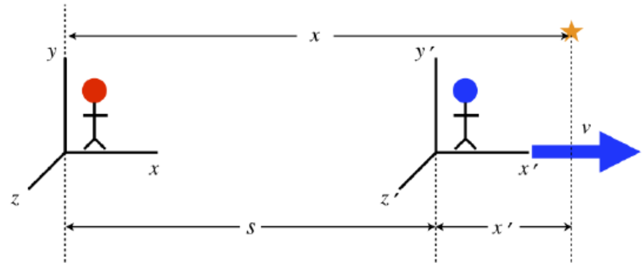
Galilean Transformation

Let's now consider two observers in difference reference frames that are moving at a constant speed relative to one another, which we will call v. We'll define the coordinate systems of these two observers such that their origins coincide at time $t = 0$, and both observers agree on this starting time. Since the frames are moving relative to each other, this common origin only lasts for that one instant in time. We'll also define the coordinate systems such that they have common x, y, and z axes when their origins coincide, and have their relative motion be along their common x-axis. We will label position coordinates and time measured by the frame moving in the $+x$-direction with a prime, to distinguish it from the other frame.

Suppose both observers record the motion of the same object. One observer gets equations of motion of this object for its three spatial coordinates (x, y, z) as a function of time t, while the other observer gets equations of motion of the object for (x', y', z') as a function of time t'. The question we want to answer is, "Given what we know about how these frames are related to each other, what are the relations between the primed and unprimed coordinates?"

Let's start by noting that when the primed observer's origin has moved a distance s relative to the unprimed observer's origin, the x-component of an object's position measured in the unprimed frame will be greater than the same component measured in the primed frame by that amount:

Figure 1.8.3 – Relating Coordinates of Reference Frames



We defined the frames so that their origins coincided when each of them measured the time to be zero, so the distance s is simply equal to vt. The only difference in the two frames is in the x-direction, and the clocks are synchronized, so we have a complete translation of the two frames:

$$\begin{aligned} t' &= t \\ x' &= x - vt \\ y' &= y \\ z' &= z \end{aligned} \tag{1.8.4}$$

These are referred to as the *Galilean transformation equations*. They translate the coordinates of one frame into another that is moving relative to the first, with the restrictions indicated above regarding coinciding origins and so on. While this may not seem particularly interesting, keep in mind that these coordinates (when combined with cartesian unit vectors) compose the position vector, whose first derivative with respect to time is the velocity vector, etc. That is, every element of 3-dimensional kinematics – all the equations of motion of observed objects – can be translated into what they would be in another frame of reference through this transformation.

Example 1.8.2

Ann and Bob are observers from different reference frames in relative motion, with all of the conditions necessary for their coordinate systems to be related by the Galilean transformation given above (Bob is in the primed frame, moving in the x-direction relative to Ann at a speed v). Ann observes a toy rocket moving in the y-direction with a speed u. Show that the velocity vector of this same rocket as measured by Bob is the same as would be obtained using the method of relative velocity vectors described in the previous section.

Solution

Let's start by computing the velocity vector of the ball according to Bob using the Galilean transformation. Taking the derivative of the position components with respect to time gives the components of the velocity vector seen by Bob, so substituting for t' and x' in the derivative gives:

$$\left.\begin{aligned} \frac{dx'}{dt'} &= \frac{d\left(x-vt\right)}{dt} = \frac{dx}{dt} - v = -v \\ \frac{dy'}{dt'} &= \frac{dy}{dt} = u \\ \frac{dz'}{dt'} &= \frac{dz}{dt} = 0 \end{aligned}\right\} \;\Rightarrow\; \overrightarrow{u}' = -v\hat{i} + u\hat{j}$$

Now let's use the tail-to-head relative velocity vector method from the previous section. The velocity of the rocket relative to Ann is $u\hat{j}$, and the velocity of Bob relative to Ann is $+v\hat{i}$. To get the velocity of the rocket relative to Bob, we need to form the "vector chain," which means we first need to get the velocity of Ann relative to Bob. Swapping the relative order requires only a minus sign, so doing this and putting together the vector chain gives:

$$\left.\begin{aligned} \text{velocity of rocket relative to Bob} &= \overrightarrow{u}' \\ \text{velocity of rocket relative to Ann} &= u\hat{j} \\ \text{velocity of Ann relative to Bob} &= -v\hat{i} \end{aligned}\right\} \;\Rightarrow\; \overrightarrow{u}' = u\hat{j} - v\hat{i}$$

CHAPTER OVERVIEW

2: Force

Thumbnail: A line drawing of two ice skaters demonstrating Newton's third law. Image used wtih permission (CC BY-SA 3.0 Unported; Benjamin Crowell).

2.1: Properties of Force

Newton's First Law

We now understand how to handle motion in all its forms, but really we haven't done much in the way of physics, because we haven't explained what causes these different motions. In ancient times, Aristotle made the observation that eventually all things seem to come to rest, which led him to conclude that a stationary condition was "natural" for everything (well, everything on Earth – heavenly bodies never seemed to stop moving). He stated that keeping things moving requires constant pushing or pulling, or it would eventually settle into a state of rest.

This is a very intuitive way of describing the nature of things, and most people even today see the world this way. It wasn't until nearly 2000 years after Aristotle that a genius born on Christmas day would overturn that long-held belief. His name was Isaac Newton, and he claimed that in fact nature behaved in precisely the opposite manner. Newton claimed that it was not natural for objects to be at rest unless they were already at rest. If they were already moving, then it was natural for them to continue moving. He claimed that it was the fact that objects on Earth could not escape pushes and pulls that accounted for them always coming to rest.

But Newton was more specific about this "natural state of motion." He stated that the only type of motion that would continue indefinitely if undisturbed by pushes or pulls was constant velocity (speed and direction) motion. That is, any motion that involved changes of speed or direction requires a push or pull.

Newton's 1st Law of Motion

Objects at rest or in motion at a constant speed in a straight line will remain in that state unless acted upon by an external influence.

Definition of Force

What we have been calling "pushes and pulls" or "external influences" is called *force* in physics. Most people have an intuitive idea of what force is, and like so many other physics concepts, this intuition is likely *wrong*. We'll start by saying what force is *not*, then move on to its definition.

Alert

Force is not a quantity stored in, or possessed by, an object. Force cannot be transferred from one object to another, nor can one claim that one object "has" more force than another. This can be a hard notion to shake.

Definition: Force

Force is an interaction between two objects, which comes in the form of a push or a pull.

When this interaction occurs, each object is affected by the other, and there are always two objects involved. As you begin your exploration into the concepts related to forces, it is a good idea – at least for awhile – to adhere to a very strict structure of wording when describing a force. Doing so will help you avoid common pitfalls in understanding. After awhile it is okay to abbreviate the description some, but if confusion ever returns when talking about a force, returning to the phrasing should help:

The "Force Phrase"

"...$\langle type\ of\ force\rangle$ **on** $\langle object\ experiencing\ force\rangle$ **by** $\langle object\ exerting\ force\rangle$ ..."

The only part we are not yet ready to fill in is the type of force. So let's take a moment to catalog a few types of forces that we will be working with. We will add more as the quarter goes on, but we need something to get started with for the sake of discussion.

Types of Forces

According to Newton, forces are interactions that cause objects to vary from their natural state of motion, but we can best describe them as pushes or pulls. We have already talked about one way to get objects to speed up, slow down, and change direction, and that is gravity. Gravity is a type of force, but what is more, it is what is referred to as a *fundamental force*. Essentially a fundamental force is one that acts rather mysteriously at a distance. Other examples of fundamental forces you are familiar with from your experience are the forces due to magnetism and static electricity (there are also nuclear forces that are fundamental with which you do not have direct experience). Every other force we will discuss in this class requires some kind of contact between macroscopic objects. This goes back to our notion of physics being about constructing usable models. Ultimately every push or pull breaks down to one or more fundamental forces, but (other than gravity), these forces are between microscopic particles, and we want our model to give results in the macroscopic realm.

So let's name some of the other (non-fundamental) forces we will be studying, so that we have some common language. We will not discuss details of the properties of these forces just yet – we are just establishing some language so that we can describe forces in physical situations.

- *tension* – a pulling (attractive) force transmitted via a rope or string
- *contact or normal* – a pushing (repulsive) force resulting from direct contact of two surfaces ["normal" means "perpendicular", and is appropriate because this force always acts at right angles to the surfaces in contact]
- *friction* – a force that occurs due to rubbing or attempted rubbing of two surfaces, which opposes the relative motion or attempted relative motion
- *air resistance or drag* – a force similar to friction that comes about because the atmosphere impacts with and rubs across the surface of an object moving relative to it

There are one or two more modeled forces we can (and will) add, but these are sufficient for our purposes right now.

Newton's Third Law

As is implied by the name "first law," Newton was not finished – he posited two other laws of motion as well. We'll return to the second law shortly, but first we will discuss the third law. You have almost certainly heard it before:

Newton's Third Law

For every action there is an equal and opposite reaction.

This is an extremely unfortunate use of language, and this law has been misinterpreted for hundreds of years as a result. It is often heard quoted in movies to essentially express how natural it is to seek retribution. Something like, if someone hits you, you will hit them back afterward.

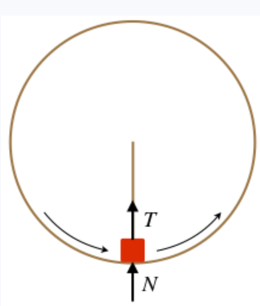
ALERT

The idea of a "reaction" as we understand it in common parlance is that it is a consequence of a previous action, but this is not the way that Newton means it.

Okay, then, so how does Newton mean it? Forces are *interactions*, and just as it is impossible for a single hand to clap, it is equally impossible for a single object to be the sole participant in a force interaction. So if one object experiences a force from another, there must be a reciprocal force also felt in the other direction at exactly the same moment, with precisely the same magnitude and in precisely the opposite direction (yes, of course forces are vectors). So for every force you can name, there exists an evil twin that acts in the opposite direction with equal magnitude. These "twins" are often called *Newton's third law force pairs*. The easiest way to identify these pairs is to use the force phrase for one force and then reverse the "on" and "by" to describe its third law pair.

Example 2.1.1

A block weighing 12lb travels in a circular path in a vertical plane. As the block does this, it slides along a frictionless circular track, and it is also attached to a string, the other end of which is attached to a fixed point at the center of the circle. When the block is at the bottom of its circular path, the contact force exerted on it by the track equals the tension force exerted on it by the string, and both are equal to 12lb. Which of the following forces is the Newton's 3rd Law pair corresponding to the gravity force on the block?



a. *the normal force on the block*
b. *the tension force on the block*
c. *Either (a) or (b) can be considered a third law pair for the gravity force.*
d. *the sum of (a) and (b)*
e. *None of the above is a third law pair to the gravity force on the block.*

Solution

(e) Don't let all the special information provided and coincidental numbers fool you! Just reverse the "on" and the "by" in the force phrase. The gravity force interaction is between the block and the earth, so the third law pair of the gravity force on the block by the earth is the gravity force on the earth by the block.

Now for a puzzler that gets to the heart of understanding the third law:

> ***A Puzzling Question:*** *When I push against the wall neither of us "wins" – I don't get the wall to move, and it doesn't get me to move. So it makes sense that the forces we exert on each other have to be equal-and-opposite. But if Newton's Third Law applies to every interaction, how does anyone ever win, in say, a tug-o-war?*

We'll answer this question in a moment, but before we do, we'll need an extremely powerful tool for analyzing physical situations involving forces.

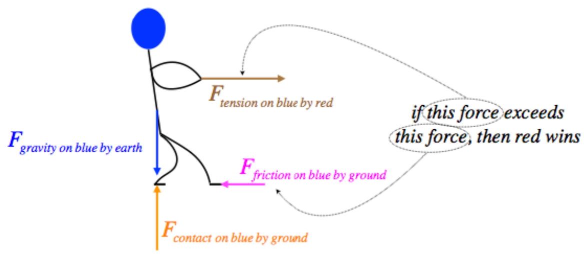
Free-Body Diagrams

Possibly our most powerful tool for analyzing forces and their effects on the motions of objects is the *free-body diagram* (or *FBD* for short). This is a diagram that consists of a single system (the "free-body," which can be a single object or a collection of objects with a collective fate), with arrows representing force vectors drawn on it. There are a few rules to drawing an accurate FBD:

- It must include only an isolated "system." This system can consist of one object or many, but the analysis that follows applies to the system as a whole, and nothing outside this system – whatever its role in the physics – is included in the diagram.
- The force vectors must be "real" forces. If you can't name the force with one of the forces mentioned earlier, then you are probably trying to fix something that isn't broken by inventing a force. Also, no forces calculated from aggregates of other forces should be included – just separate, physically-describable forces.
- Only forces *on* the system can be included – never forces *by* the system. If every vector is labeled using the force phrase, there is no way to go wrong here.
- For now, where the force vectors are located on the system is not important, so the entire system can be reduced to a single dot for simplicity. But later this quarter the location where the force acts will become important, so it might be a good idea to try to place the force vectors properly right away. Since the type of force and its basic nature are related to where it acts on an system, this will also help confirm that you are dealing with the right forces, and are not trying to invent a force that doesn't exist.

Okay, let's return to the puzzling question... If two entities interact through (say) a tension force (this is fancy physics-speak for "two people have a tug-o-war with a rope"), and if they always experience the same force as the other person, how does one side ever "win?" Let's draw a FBD to see if we can see why.

Figure 2.1.1 – Analyzing a Tug-o-War Using a FBD



The reason the question is confusing is that we think that the two forces that are equal-and-opposite must always cancel out, but how exactly do forces “cancel?” They *have to act on the same system*. By drawing a careful force diagram in which we only include the forces on the system in question (in this case, the blue-headed stick figure), we see that in fact the force pairs that are equal and opposite are split ("tension on blue by red" is split from "tension on red by blue", and so on), and therefore cannot cancel each other. The real determining factor of whether an individual wins the tug-o-war is whether that individual receives a friction force from the ground that is greater than or less than the tension force, unbalancing the total horizontal force on them.

Does this mean the person is at the whim of the ground, that either decides to provide a big or small friction force? Of course not! The friction force on our feet by the ground is equal-and-opposite to the friction force our feet exert on the ground, and we do this by leaning back and sliding (or push our foot forward as if to slide it) across the ground. We can see that this is the case, because even the strongest human in the world cannot win a tug-o-war against a small child if the strong person is on ice or on some rolling device that doesn’t allow them to push horizontally (and thereby be pushed back the opposite way).

2.2: Effects of Force

Newton's Second Law

We have built some tools for analyzing situations where forces act on objects (force phrase, FBDs), and we know that there can only be accelerations when forces are present (first law). But we still are not yet able to describe the motion of an object under the influence of one or more forces. That's because the first law only tells us qualitatively what is happening. In physics we seek to develop quantitative models, and that's where the second law comes in. It is really just a more detailed description of the first law, or alternatively, the first law is just a special case of the second law.

We know that force is related not to velocity (because the first law says that constant velocity exists in the absence of force), but rather the *change* of velocity. More specifically, the rate of change of the velocity – the acceleration. Newton defined force in the simplest possible fashion in terms of acceleration – with a linear relationship. He knew that pushing equal amounts on two objects of different masses resulted in different changes of motion, so he stated that the relationship between force and acceleration as a simple proportionality:

$$\text{acceleration of object} = \frac{\text{force acting on object}}{\text{mass of object}} \tag{2.2.1}$$

The idea is that for a given force, the reaction of the object (in the form of an acceleration) is inversely-proportional to the amount of mass the object possesses. Let's take a moment to mention units:

$$[F] = \frac{kg \cdot m}{s^2} = \text{"Newtons"} \;\; (N)$$

There is much more detail lurking in here. First of all, acceleration and force are both vectors, while mass is a scalar, so the second law is actually a vector equation:

$$\overrightarrow{a} = \frac{\overrightarrow{F}}{m} \tag{2.2.2}$$

This means that the acceleration experienced by an object is just a scaled vector of the force exerted on the object. That is, the acceleration and the force always point in the same direction (mass is never negative). Of course, the scaling also changes the units.

ALERT

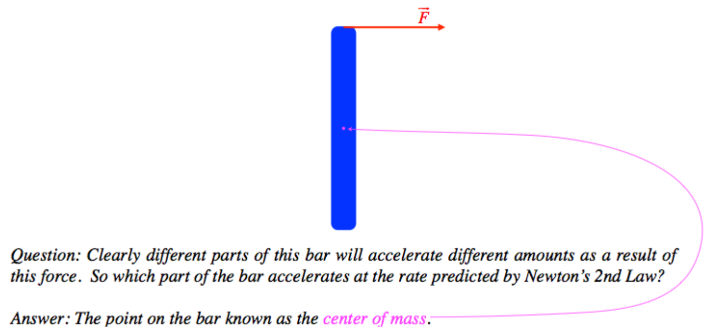
Most people first encounter Newton's second law expressed as $\overrightarrow{F} = m\overrightarrow{a}$. While this is mathematically equivalent to what is above, it is very dangerous to write this way, as it encourages a very common misconception. We write it as we do above to emphasize the interpretation: "the ***effect on the motion*** *(the acceleration) results from the* ***cause*** *(the force), moderated by a* ***property*** *of the object experiencing the effect (the mass)." The danger of using the other expression is that it reads like, "the force* ***of*** *the object equals the mass* ***of*** *the object multiplied by the acceleration* ***of*** *the object." This turns the quantities of force and acceleration into properties of the object, rather than cause and effect, and this leads to subtle but important misconceptions.*

We aren't done modifying the second law to its proper form yet! We know from our discussion of force diagrams that a large number of forces can be on an object at the same time. Which force is the one that causes the acceleration? All of them. Do we figure out the accelerations of each force and then add them up? That makes no sense physically – objects do not experience lots of accelerations at once. Instead, we take all of the forces together and add them as vectors to create a single composite force we call the *net force*, and that is what goes into the equation:

$$\overrightarrow{a} = \frac{\overrightarrow{F}_{net}}{m} \tag{2.2.3}$$

Still not done! Suppose only a single force is acting on a specific point of an extended object? Remarkably, it doesn't matter where the force acts on an object, if you measure the motion of the *center of mass* of the object (a quantity we will examine in more detail later on), it is the acceleration of *that point* that is determined by the force and mass.

Figure 2.2.1 – What Part of an Extended Object Accelerates According to the Second Law?



So finally we have arrived at the full-blown version of *Newton's second law of motion*:

$$\vec{a}_{cm} = \frac{\vec{F}_{net}}{m} \tag{2.2.4}$$

Now perhaps the main point of the free-body diagram is clear: The diagram facilitates our ability to add force vectors (all of which are on the object, not by it), giving us the net force acting on it. We use the usual tools for vector addition to obtain this net force, perhaps expressing it in terms of its components in some coordinate system. Then we divide this vector by the mass of the object, and we have the acceleration of its center of mass, which could also be expressed in terms of its components. Then the acceleration is used to describe the object's motion as we have used it in kinematics.

Example 2.2.1

A driver steps on the brake pedal of her car, slowing the car down, and her body experiences an acceleration as a result. Which of the following forces does Newton's 2nd Law include when determining her acceleration?

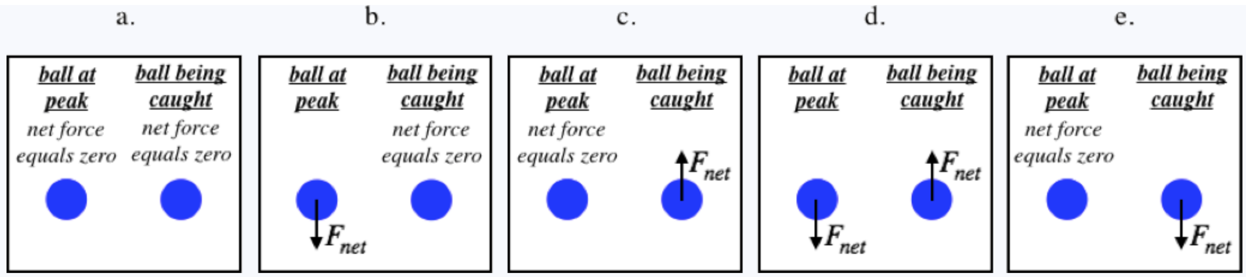
a. *normal force by driver's foot on the brake pedal*
b. *friction force by the car tires on the road*
c. *friction force by the road on the car tires*
d. *all of these*
e. *none of these*

Solution

(e) One must be very precise when it comes to identifying forces, as ultimately they must be plugged into a mathematical formula. It is not enough that a force sets off a chain of events that leads to an acceleration, it must be the direct effect of that acceleration in order to be the force that is featured in the second law. In addition to being a direct force, it isn't even sufficient to isolate the correct interaction – the specific "twin" from the third law force pair must be identified. That is, the force must be ***on*** *the object in order to accelerate it. The problem asks what force slows down her body. The normal force on the brake pedal affects the motion of the brake pedal. The friction force on the road affects the motion of the road. The friction force on the tires (which are part of the car's system) affects the motion of the car system. While parts of the car system (namely tension by the seatbelt, friction by the car seat, and normal force by the steering wheel do slow her down, and the friction force on the tires slows down the car, this chain of events does not mean that the friction force slows her down. If you plug the friction force on the tires and her body's mass into the second law, the acceleration you calculate for her will not be correct.*

Example 2.2.2

A boy throws a ball straight up, and catches it when it returns. Which pair of diagrams best represents the directions of the net forces experienced by the ball when it hits the peak of its flight (i.e. when it isn't moving), and while the boy is catching it (i.e. not after he has caught it)?



Solution

(d) Gravity is always acting on the ball, no matter where it is. When it reaches its peak, there are no other forces on it (the boy's hand is no longer in contact with it), so the gravity force is the net force, and it points down. When the ball is in the process of being caught, there is a force up on it by the boy's hand, and since the ball is moving downward and is slowing, it is accelerating upward, which means the force from the boy's hand exceeds the force of gravity and the net force is upward.

Combining Newton's Second and Third Laws

We already saw in the case of the tug-o-war how the second and third laws work together: Each participant exerts an equal tension force on the other. These forces don't cancel because according to the second law, it is only the forces *on* the system that cause it to accelerate, and these two forces act on different systems. One individual feels this tension force on them, and if it is the only force, it will accelerate them forward (i.e. they will lose). To counter this, this person pushes the ground forward with their feet. But this is a force *on the ground*, which can only affect the acceleration of the ground, not the person, according to the second law. But now the third law comes to the rescue – when the person pushes on the ground, the ground pushes back equally, and this force *is* on the person. This opposes the pull of the rope, and if it is stronger than the rope's pull, the net force on the person is backward, and that person wins.

Another question that can be very confusing at first about the third law is this:

Another Puzzling Question: *If the gravity force on a rock by the earth is equal and opposite to the gravity force on the earth by the rock, then why doesn't the earth accelerate upward when a rock is dropped to the ground?*

As you might have guessed from the title of this section, it is the second law that comes to the rescue here. The answer is that the earth *does* accelerate up toward the rock! It feels the same net force that the rock feels, thanks to the third law. But when it comes to the acceleration, we need to divide this net force by the mass of the accelerated object, and since the mass of the earth is so much greater than that of the rock, it experiences a much, much smaller acceleration – so much smaller that it is imperceptible.

Finally, for a complete combination of all three of Newton's laws of motion, consider this: You are floating by yourself, untethered and weightless in outer space, mere feet from the hatch to the space station. To get there, you think of the idea of grabbing your own space suit and pulling yourself toward the hatch, and naturally it doesn't work (how you made the space program with such nutty ideas is a mystery). Why doesn't this work? Clearly you can exert a force on your space suit, which should result in an acceleration of that object, right? Well, in this case the relevant "system" is your entire body (you can remove your space suit and accelerate it to the hatch if you like, but that is not advisable). The system of your entire body includes your hand, which is doing the pulling, and experiences the third-law-pair force of the space suit pulling on it. If you draw the free-body diagram of your body and include the force on your suit as well as the force on your hand (both are acting on the system), then they clearly cancel. This had better be the case, because if it wasn't, it would mean that objects can just accelerate themselves, which violates Newton's *first* law of motion!

A Summary of Concepts Related to Newton's Laws

Much of what we have discussed in this section and the one before it will be repeated below, but putting all of these idea in one place may help the reader consolidate the ideas into a cogent "big picture."

1. Force is not a quantity contained within an object.
2. Forces are push or pull interactions between two objects. If one looks at the two individual forces that make up the interaction, then those two forces are always equal in magnitude and opposite in direction (Newton's 3rd Law).
3. To avoid confusion, we learned the all-important "force phrase," which reminds us that the individual forces that make up the interaction force pairs always act *on* one object and *by* another.

4. Forces are the cause of accelerations. It is impossible to have one of these without the other. This means that forces (if the vectors don't all cancel each other out) speed up, slow down, or change the direction of an object’s motion. And conversely, if an object’s motion slows down, speeds up, or changes direction, then it must be experiencing a (net) force. (Newton’s 1st Law)
5. Forces are vectors, which is to say that they have magnitude and direction.
6. The force vector that causes an object to accelerate is the *net* force on that object, that is, the vector sum of all of the individual forces exerted on the object. A net force is a combination of one or more real forces, but is not itself a type of force.
7. Only the forces *on* an object can contribute to its acceleration (i.e. added together to give the net force), never the forces *by* it. Forces by an object only affect the motions of the *other* objects that they act on.
8. The amount of net force on an object is proportional to the amount of acceleration it experiences, and the constant of proportionality is the mass, a measure of how much stuff is present in the object. (Newton’s 2nd Law)
9. The fact that net force and acceleration are proportional means that as vectors, they must point in the same direction, since mass is never negative.
10. Mass is sometimes called “inertia,” which can be loosely thought of as resistance to acceleration. But this must not be confused with resistance to motion – the smallest net force will cause an acceleration of the largest mass. If a mass at rest doesn’t start to move when a small individual force acts on it, it is because there is another force balancing it out, causing zero net force, not because the inertia of the object cannot be overcome.
11. The part of the object that experiences the acceleration described in Newton’s 2nd Law is the center of mass of the object, not the point on the object where the force is acting.
12. A useful tool for analyzing forces is the force diagram, which consists of isolating an object (which is why it is called a “free body diagram”), followed by drawing in all the force vectors acting on it. Careful use of the force phrase helps us avoid putting incorrect forces on this diagram.

2.3: Types of Forces

Gravity

We already know something about gravity from our study of free-fall and projectile motion. We know that the acceleration is the same for objects of different masses. While we have used this as a model, it is a big step to claim that gravity fundamentally follows this rule. We know that a feather will experience the same acceleration due to gravity as a stone, if air resistance is removed. Now how do we put air resistance back into our model so that the reduced acceleration of the feather makes sense?

The effect of reduced acceleration is easy to show with a FBD of two objects that are identical except for mass and are falling through the air at the same speed. For these two objects the air resistance forces are equal, and the gravity force is greater on the heavier object. The net forces on the two objects are therefore different, giving the following accelerations:

$$\left.\begin{array}{l} a_{(heavy)}=\dfrac{F_{gravity\;(heavy)}-F_{air}}{m_{(heavy)}} \\ a_{(light)}=\dfrac{F_{gravity\;(light)}-F_{air}}{m_{(light)}} \\ g=\dfrac{F_{gravity}}{m} \end{array}\right\} \Rightarrow \begin{array}{l} a_{(heavy)}=g-\dfrac{F_{air}}{m_{(heavy)}} \\ a_{(light)}=g-\dfrac{F_{air}}{m_{(light)}} \end{array} \tag{2.3.1}$$

So the reason the heavier mass accelerates more is simply that the effect that the air resistance force has on it is smaller. Care must be taken not to draw quick conclusions, because it is possible to draw an incorrect conclusion about a single force by not paying attention to a second force that is present. One simple result comes out of the free-fall case without air resistance:

$$F_{gravity}=ma,\; a=g \Rightarrow F_{gravity}=mg \tag{2.3.2}$$

Alert

It is important to understand that here g has a different meaning than it had when we were discussing motion involving gravity-caused acceleration. Here the g is a physical constant, which we use to determine the gravity force on an object with mass m. It does ***not*** *mean that the object is accelerating at* $9.8\frac{m}{s^2}$*! When an object experiences no other force than gravity, the object's acceleration just happens to equal this constant, but the constant is present regardless of the state of acceleration of the object.*

Tension

When we model problems that involve tension forces exerted by strings or ropes, we usually assume that the string or rope has no mass. Not to do so brings in some challenging (but interesting!) complications that are generally not the focus of that problem. For example, imagine two people having a tug-o-war, where one person is winning, and is therefore accelerating the other. A FBD of just the rope shows two forces on it (if we ignore gravity), one in each direction. With one of the people winning, not only do both people accelerate, the rope does as well. This means that there is a net force on the rope, and therefore one person is pulling on the rope harder than the other person.

By assuming that the rope is massless, the mass-times-acceleration for the rope is just zero, which means there is no net force on it, despite the fact that it is accelerating. With no net force on it, the two combatants are pulling on it with equal force, which means that essentially the force one person exerts on it is transmitted through the rope to act on the other person. That is, the rope becomes no different than a case of the two people clasping hands and pulling on each other. This allows us to use ropes as an idealized means of allowing the "by" object to exert a force "on" another object from a distance. Without this device, we would either have to devise examples in an awkward manner, or overly-complicate problems by including accelerating massive ropes.

Alert

Occasionally ropes will be used as a conduit for expressing third-law pairs ("tension force on A by B," where neither A nor B is the rope), but this is dangerous, because ropes can also transmit the force around a pulley, in which case the Newton's third law pair of forces are clearly not in opposite directions.

Pulleys are another aspect of tension we should say a word about. If a pulley experiences friction as it rotates, or has some mass (so that its rotation requires the acceleration of some mass), then the tension force is not transmitted around it unaffected. We will actually look at the case of pulleys with mass later in the course, but for now we idealize them as we do ropes – no friction and no mass.

Normal (Contact) Force

What makes the contact force different from some others is that it is nothing more than a "balancer." This means that its magnitude is determined not by independent criteria but rather by the details of the contact. For example, if a block rests on a level tabletop, then the zero acceleration of the block means that the net force on the block is zero. The Earth exerts a gravity force downward on the block, so there must be a balancing force upward to result in zero net force. The only other force available here is the contact force, so its magnitude adjusts to equal the force of gravity on the block. If we now tie a string to the block and pull up on it with a force less than the weight of the block, it maintains contact with the tabletop, and still remains stationary. This time for the forces to balance, the amount of contact force necessary to create balance is reduced by the amount of tension force the string provides.

The primary place where this feature of contact force comes up in problems is when a scale (like one you stand on in your bathroom) is involved. We are used to thinking of a scale as something that measures weight, but this is not the case – it measures contact force! For proof, just imagine pushing a scale against a vertical wall, rather than standing on it.

Kinetic Friction

Probably the most complicated and difficult to understand of the forces we have mentioned so far is friction. The microscopic mechanism is not even that well understood, partly because it has two different modes. This first of these modes comes from the fact that the two surfaces can be highly irregular. The second mode involves adhesive bonds that are formed between the surface molecules for each of the objects involved. Note that the fundamental force involved in both cases is electrical in nature, but it is repulsive in the first case and attractive in the second.

Both of these modes lead to two different forms of friction. The first we will discuss involves two surfaces rubbing across each other. This is called *kinetic friction*. One might imagine a number of contributing factors that affect the amount of rubbing friction force that might exist between two objects, but for our macroscopic model we boil those factors down to just two. The first is another force – the normal force between the surfaces. The harder the surfaces are pushing against each other, the deeper the surface irregularities mesh, and the more molecules are brought together to bond. Experimentally, we find that the friction force grows roughly linearly with this normal force. All the remaining factors we lump into a single unitless constant that multiplies the magnitude of the normal force, known as the coefficient of kinetic friction:

$$f_k = \mu_k N \tag{2.3.3}$$

It is important to note that having the two objects moving while in contact does not ensure that kinetic friction is occurring – the surfaces need to be *sliding across each other*. For example, a ball rolling across a floor that is not slipping in any way is not experiencing kinetic friction.

Static Friction

The second form of friction is a little harder to grasp, though we are all aware of it. Instead of being a force that exists while two surfaces are sliding across each other, this type of friction is a reaction force that attempts to *prevent* two surfaces from sliding across each other. This force, known as *static friction*, is responsible for the phenomenon we experience when we try to slide something and fail until we make enough effort to get it sliding – with any smaller force the object stays put.

Static friction is very similar to contact force, inasmuch as it is merely a balancer. If you push on something hard enough that it begins to slide, then static friction no longer applies, but if it doesn't start sliding, then it is because the static friction force balances that push, resulting in a zero acceleration. Unlike contact force, however, static friction has a maximum – if other forces are sufficiently great, the sliding will begin, and the friction force shifts from static to kinetic. This maximum is also lumped into a constant, but the relation between the friction force and the contact force is now an inequality:

$$f_s \le \mu_s N \tag{2.3.4}$$

The most common place that we encounter friction (kinetic or static) is for an object on a horizontal plane, such as a table on a floor or a plate on a table. When the surface is horizontal, the zero vertical acceleration means that the contact force equals the weight, which means that the friction force is directly proportional to the weight. Put another way, "Heavy things are hard to slide across the floor." But this jump is not a good one to make, since not all surfaces are horizontal (and acceleration is not always zero, e.g. in an elevator). The magnitude (or maximum magnitude) of the friction force is determined by the contact force – not by the gravitational force, and certainly not by the mass of the object.

A phenomenon we are all familiar with is that for the same two surfaces in contact, the coefficient of static friction is greater than the coefficient of kinetic friction. That is, it is more difficult to get an object sliding across a surface than it is to keep it sliding across it. This can be understood conceptually by thinking about the two microscopic modes we discussed earlier. When the surfaces are sliding across each other, the irregularities don't have time to settle into each other, and the molecular bonds are being broken and re-formed, and in the period before re-forming there is no force.

This phenomenon is the main culprit behind what many people mistakenly refer to as "inertia." The physical concepts involved are insidiously subtle: Someone pushes an object across a floor, accelerating it from rest, and notices that it is harder to get it going than to keep it going. Then they try to push a bigger object and finds it even more difficult to get started. They draw the conclusion that the greater mass of the second object means it "has more inertia to be overcome" to get it going. Indeed, these people may have even heard of the "law of inertia," and know that mass is related to inertia.

We now have the tools to debunk this analysis. The "law of inertia" is Newton's first law. This law says that the object being pushed "wants" to remain stationary, but the law does not provide for a minimum force necessary to overcome the object's "desire" to stay in place. The effect of a push on the object is determined by Newton's second law, and no matter how large the mass is, or how small the net force is, a non-zero acceleration will result. The correct explanation for this apparent inertia phenomenon is this: For an object on a horizontal surface, the normal force equals the gravity force. The gravity force is proportional to the mass of the object (it equals mg), so the normal force is proportional to the mass. The maximum static friction force is proportional to the normal force, so it too is proportional to mass. Therefore the amount of force required to get something sliding on a horizontal surface happens to be proportional to the object's mass. But this is not an intrinsic property of mass - there are many important steps between the amount of mass and this thing people call "inertia." It is dangerous to jump to quick conclusions without a careful analysis.

Application: Anti-Lock Brakes

The fact that the coefficient of kinetic friction is greater than the coefficient of static friction for the same surfaces has its greatest application in an invention called "anti-lock brakes." When a tire is rolling perfectly over the road surface, the surface of the tire is not sliding across the surface of the road, which means that if any friction is involved, it is static friction. If you apply the brakes, the tires' rotation will slow down until the static friction maximum is exceeded, at which point the tires will stop turning and the tires will slide across the road surface. The friction force on the tires goes down when this occurs, because the kinetic friction is smaller than the maximum static friction. So an ABS system automatically releases the breaks briefly so that the tires again turn, restoring perfect rolling and allowing the return of static friction. This would be like trying to push a heavy cardboard box across a floor in extremely short bursts – as soon as the box starts sliding (and gets easier to push), you stop and start over. Before the invention of ABS systems, drivers were told to "lightly pump their brakes" in slippery situations to create this same effect. ABS systems do the pumping for us, with a much greater frequency than we could manage, and to great effect.

Air Resistance (Drag)

Air resistance is sometimes referred to as "air friction," but although it has dissipative qualities similar to those of friction (as we will see when we study energy), the mechanism by which it functions is quite different. Air resistance occurs because of countless collisions between the microscopic particles that comprise the atmosphere and a macroscopic object moving through it. There can be no collisions without motion, so unlike friction there is no "static air resistance." In this class we really won't deal with air resistance in a rigorous mathematical way, primarily because fluid dynamics is an especially complicated subject (this is why we simplify projectile problems by assuming no air resistance). But we can determine a few characteristics that make sense:

- cross-sectional area – Anyone who has ever put their hand out the window of a moving car knows that the air pushes it back with greater force when the palm faces forward than when it faces downward. The difference is the cross-sectional area, and that is important because the greater this value is, the more atmospheric particles per second can hit the hand. Obviously if the area is doubled, the number of particles hitting the object per second is doubled, so the drag force is directly proportional to the cross-sectional area.

Alert

"Cross-sectional area" is not the same as "surface area." The surface area of one's hand doesn't change when it is rotated from a palm-forward position to a palm-downward position when the hand is held out the window. What matters is the area perpendicular to the relative velocity of the incoming particles that comprise the air.

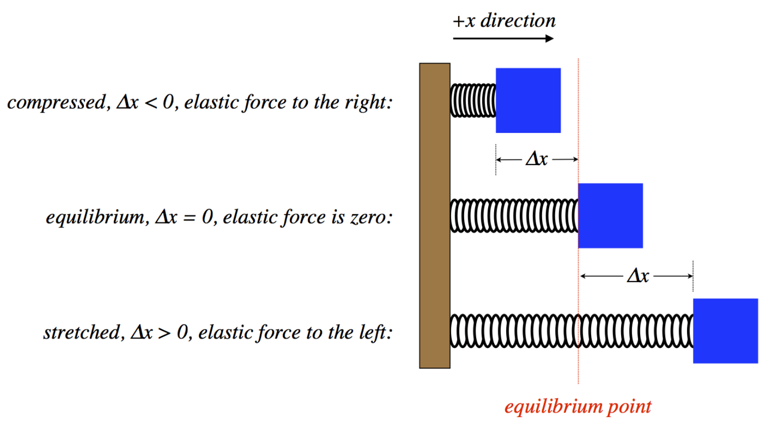
- speed – Returning to the common experience of holding a hand out of a moving car window, we know that the faster the car is going, the greater the force is on our hand. In this case the exact mathematical relationship is not as obvious as it was for cross-sectional area. In fact, it can depend upon the relative air speed differently for different air speeds (or more accurately, whether the air flow is orderly or turbulent).
- air density – Density is a measure of how many particles you find in a given small region of space. If the air is very dense, then there are more particles that can strike the cross-sectional area than if the air is less dense. So the drag force is directly proportional to the density of the air.

We will not write down or have use for a specific equation for air resistance, but we can use the facts above to draw some conclusions. Imagine you have just jumped out of a plane, and the Earth's gravity begins accelerating you downward. This acceleration results in your speed getting greater and greater, and as it does, the air resistance force upward on your body also gets greater. Of course, your downward acceleration is determined not by gravity alone, but rather by the net force on your body, which is reduced as the air resistance (which opposes gravity) grows. Eventually you end up going fast enough that the air resistance force equals your weight, and you stop accelerating. The speed at which this happens is known as *terminal velocity*. Notice that when you pull your ripcord and your parachute deploys, you increase your effective cross-sectional area, and to achieve the same amount of drag force, not so much speed is required. This is good, because it allows you to hit the ground with a speed that is not … um… terminal.

Elastic (Spring) Force

To close out the section on details of forces, we'll look at one not yet mentioned. Actually, it can fall under the categories of both tension and contact force (and is therefore both attractive and repulsive), but it has the nice feature that we can deal with it a bit more precisely mathematically. It is called the *elastic or spring force*. The characteristic feature of this force is that it acts in a restoring fashion (and is therefore a type of *restoring force*), meaning that there exists an equilibrium status, and any deviation from this equilibrium leads to a force acting to return the object to the equilibrium state.

Figure 2.3.1 – Elastic Restoring Force



In particular, the elastic force is proportional to the separation of the object from its equilibrium position:

$$|F_{elastic}| = k|\Delta x| \tag{2.3.5}$$

This force always points from the point to which the object is displaced back to the equilibrium point. The constant k is known as the *spring constant*, and has units of *Newtons per meter*. It is a measure of how "stiff" the spring is (i.e. how difficult it is to stretch or compress). As always, it is important to keep in mind that this is a model for the spring force. Actual springs typically deviate from this behavior, in some cases significantly. But the usefulness of this model in physics cannot be overstated. Indeed this model has found its way into literally every corner of physical theory.

While Equation 2.3.5 describes the elastic force well, there is actually a nice way to express it in a compact manner as a vector equation. Treating displacement from the equilibrium as a vector, we see that the direction of the force is in exactly the opposite direction, regardless of whether the spring is stretched or expanded. Accordingly, we can write:

$$\overrightarrow{F}_{elastic} = -k\Delta\overrightarrow{x} \tag{2.3.6}$$

This equation is commonly known as *Hooke's law*.

Alert

Note that the usage of the "Δ" in Hooke's law is different from how we have used it up to this point – here it refers to a difference in positions, rather than a difference of after and before.

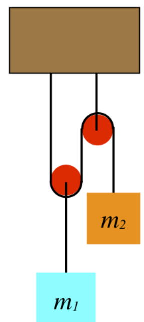
2.4: Problem Solving

Next we are going to focus on elements of problem-solving. We have all the tools we need, so this will not involve any new physics, but the idea is to introduce you to some common themes that come up in physics mechanics problems.

Pulleys

One of the favorite devices for physics problems is the pulley. As was stated in the description of the tension force, to start out we use the simplest model, which means we will assume that pulleys are massless and frictionless. Pulleys get especially interesting in situations like the following example, where at least one of the pulleys is able to move. The two blocks remain at rest in the system of ropes and pulleys shown in the diagram. Given this information, can you conclude how the two masses compare?

Figure 2.4.1 – Blocks Hanging from Multiple Pulleys



By now we know that when it comes to analyzing the forces present in a system, there is no better tool than the FBD. We begin there:

Figure 2.4.2 – FBD of the Block and Pulley

[We have taken the liberty of defining coordinate systems in our FBDs – up is the $+y$-direction for both – which we will need shortly.]

One might ask why there are two tension force vectors drawn for the pulley. The simplest answer is to consider what you would feel if you cut the rope on both sides of the pulley and held one end in each hand. Clearly you would feel both ends of the rope pulling down. Therefore by Newton's third law, both ends of the rope are pulling up on the pulley. With the pulley massless and frictionless, these two tension forces must also be equal, which explains why they are labeled the same. Note that the tension vector on the block is also labeled with the same variable name. This is because it is the *same rope*, and our assumption of massless, frictionless pulleys ensures that everywhere that we measure the tension for a single piece of rope, it will be the same.

Alert

If we were to draw the tension force vector pulling up on the right pulley, we would not be able to label it the same. Not all tension vectors in a single physical system are equal, just the magnitudes of all tension vectors derived from the same rope.

Another curious aspect of this FBD is the weight label of the left pulley. Technically, that force is acting on the block, and the block is pulling on the pulley. The pull on the pulley by the block happens to equal the weight of the block in this case, and the pulley has no weight of its own, so we are justified in taking this little shortcut. Another way to justify it is to treat the block + pulley as a single system, and the gravity force on the system is the force vector shown.

The next step in our analysis is to sum the forces for each object and apply Newton's second law, which in this case involves zero acceleration. In taking the sum of forces, we have to take care to correctly use our coordinate system:

$$\left.\begin{array}{l} 0=a_1=\dfrac{F_{net\ 1}}{m_1}=\dfrac{2T-m_1g}{m_1} \quad\Rightarrow\quad T=\dfrac{m_1g}{2} \\ 0=a_2=\dfrac{F_{net\ 2}}{m_2}=\dfrac{T-m_2g}{m_2} \quad\Rightarrow\quad T=m_2g \end{array}\right\} \quad\Rightarrow\quad m_1=2m_2 \tag{2.4.1}$$

Notice that the light weight m_2 holds up the heavier one because the placement of the pulley allows us to use the tension from the same rope twice on the heavier mass. This trick can actually be repeated as many times as we like (the pulley can have multiple tracks in it), and this enables us to lift very heavy weights with very little force. This invention is called a *block and tackle*. They are used for sailing ships (the heavy sails and boom can be pulled tighter), lifting engine blocks, and many other applications.

Constraints

Next we take on a tricky concept known as *constraint*s. A constraint is a condition that exists for a physical system that restricts how it can behave. What makes the concept "tricky" is that these ultimately play a mathematical role in the solution of the problem, but this role is often difficult to extract from the statement of the problem. Put simply, constraints relate variables in the problem to each other, providing additional equations (beyond Newton's second law) to work with. We have already seen an example of a constraint. It is the relationship between the friction force and the normal force. For kinetic friction this provides an equation that relates these two forces, while for static friction it sets an upper limit on the magnitude of friction for a given normal force.

One of the most common examples of a constraint is related to ropes moving through pulleys. This constraint relates the motion of one object to that of another when they are connected through a system of pulleys. Let us return to the system shown in Figure 2.4.1 and ask the following question: If the block m_2 drops a distance Δy, what happens to the block m_1?

First of all, it should be clear that m_1 rises as m_2 drops, so the only question is, how far? This may not be apparent at first, but think of it this way: When the pulley holding m_1 moves up 1 unit, both segments of string going up from the pulley get shorter by 1 unit. These two units of string don't simply vanish, and in fact they are taken-up by the free end of the string, which is attached to m_2. This means that as m_2 drops a distance of Δy, m_1 must rise only half that far.

What does this say about the comparison of the speeds and accelerations of the two blocks? Well, they are required to move simultaneously, so every unit of length dropped by m_2 is matched by a rise of m_1 by half as much, which means that m_1 always moves at half the speed and accelerates half as much as m_2. If this system is not balanced (as it was above), then applying Newton's second law to both blocks includes two accelerations, but these are *constrained* to be related to each other by a factor of two, providing us with an additional *constraint equation*:

$$2\left|a_1\right|=\left|a_2\right| \tag{2.4.2}$$

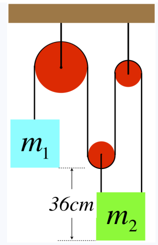
What's with the absolute values, you ask? Well, these variables can have positive or negative values, and we must be careful when it comes to signs. in particular, we have to look at how our constraint relates to our choice of coordinate systems for the two blocks. In Figure 2.4.2, we chose "up" as the positive direction for both blocks. So we need to ask ourselves, "If one block experiences a positive displacement, what is the sign of the displacement of the other block?" In this case it's clear that the displacement of the two blocks have opposite signs. Therefore the constraint equation for the block accelerations is:

$$2a_1=-a_2 \tag{2.4.3}$$

Note that it is perfectly fine to set up different coordinate systems for the two blocks – each FBD is entitled to its own individual coordinate system. How the coordinate systems relate to each other affects the equation of constraint. So for example, if we had instead chosen downward to be the $+y$-direction for block #2 (but left upward as positive for the other block), then there would be no need for the minus sign in the constraint equation – positive displacements of one block correspond to positive displacements of the other block. We see that there is therefore no "correct" choice of coordinate system, but we must take care when the time comes to combine the equations from the two FBDs that the constraint equations relate the variables correctly. We will see an even more striking example of coordinate system choice in the next example.

Example 2.4.1

The heights of the two blocks in the diagram below differ by 36cm. When they are released from rest, the higher block falls while the lower block rises. One of the blocks has a mass that is three times the mass of the other block, the pulleys are massless and frictionless, and the string doesn't stretch.



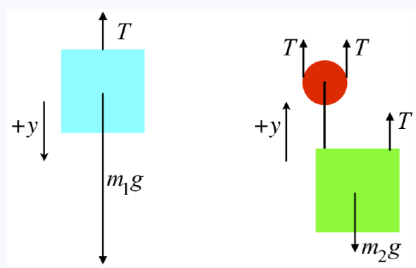
a. *Which is the heavier block? Explain.*
b. *Find the distance that the lower block rises when the two blocks are aligned.*
c. *Find the time it takes for the two blocks to be aligned.*

Solution

a. The tension exerted by the string threaded through the pulleys is the same everywhere, so there is three times as much tension force acting up on the lower block as there is up on the higher one. If the lower block was three times heavier than the higher block, then the system would be balanced, and neither mass could be accelerating. [You should try doing the math of part (c) with the masses the other way, and demonstrate for yourself that it the acceleration would have to be zero.] Given that the system is accelerating, it must be the higher block with the greater mass.

b. As the top block pulls the string down on one side of the large pulley, the same amount of string that is gained on the left side of the pulley is lost from the right side. The string on the right side of the large pulley is divided between the three segments holding up the other block. Therefore the lower block moves up one third as far as the higher block moves down. If the lower block rises a distance y, then the higher block drops a distance $3y$, and since they reach the same height, the sum of those changes is 36cm, which means the lower block rises a distance of $y = 9\,cm$.

c. To find the the time it takes them to align, we need to use their accelerations. We know their relative accelerations already: The lower block accelerates at one third the rate of the higher block. We'll therefore call the higher block's acceleration "$3a$," making the lower block's acceleration equal to a. But we need Newton's laws in order to go any further. FBD of the two systems involved look like this:



The higher mass is three times the lower mass, so we will call m_2 simply "m," which makes m_1 equal to $3m$. Plugging everything into Newton's second law for both FBDs gives these equations:

$$\left.\begin{array}{l} a_1 = \dfrac{F_{net\ 1}}{m_1} \quad\Rightarrow\quad 3a = \dfrac{-T+3mg}{3m} \\ a_2 = \dfrac{F_{net\ 2}}{m_2} \quad\Rightarrow\quad a = \dfrac{3T-mg}{m} \end{array}\right\} \quad\Rightarrow\quad a = \frac{2}{7}g$$

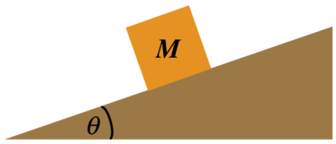
With the acceleration of the lower block, the distance it travels, and the fact that it starts from rest, we can compute the time it takes to make the trip:

$$y = v_o t + \frac{1}{2}at^2 \quad\Rightarrow\quad t = \sqrt{\frac{7y}{g}} = \boxed{0.25s}$$

Inclined Planes

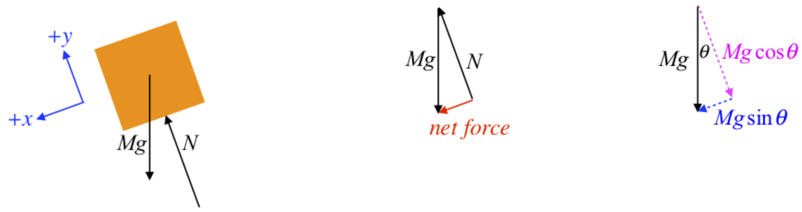
Pulleys (and in particular the block & tackle) are an example of something often referred to as a *simple machine*. This is because we can use such a device to lift a heavy weight with a force less than the weight itself. Another example of a simple machine is the *inclined plane*. These devices also allow one to accomplish the task of raising a heavy object to a higher position using less force than would be necessary in a direct lift. The main feature of problems involving inclined planes is dealing with the coordinate system used for the object on the inclined plane. Let's look at an example. As usual, we start with the very simplest of examples, and complicate it as we go – a block of mass M on a frictionless plane inclined at an angle θ:

Figure 2.4.3 – Block on an Inclined Plane



We begin our analysis, as always, with a free-body diagram. No FBD is complete without a choice of coordinate system, so we must choose one here. If we choose our coordinate system to be horizontal and vertical as we usually do, then when the block slides down the plane its acceleration will have both x and y components. This is fine of course, but it can be a bit cumbersome to work with. There is nothing sacred about the horizontal and vertical directions, so why not choose a coordinate system that is parallel and perpendicular to the plane?

Figure 2.4.4 – Coordinate System for an Inclined Plane



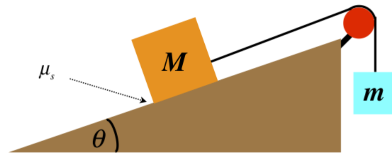
The middle diagram in Figure 2.4.4 shows the tail-to-head sum of the normal and gravity forces resulting in a net force parallel to and down the plane, as it should be, since we know that is the direction the block will accelerate. The right diagram shows the gravity force broken into its x and y components in the chosen coordinate system (you should verify for yourself the geometry that leads to concluding that the angle in this diagram is the same angle that the incline makes with the horizontal). Note that since the block does not accelerate perpendicular to the plane, we can conclude that $N = Mg\cos\theta$. Also it's clear that the net force is the x-component of the gravity force, resulting in an acceleration down the plane of simply $a = g\sin\theta$.

Suppose we wanted to raise this block to some new height. To get it moving up the plane, we would need to apply a force that exceeds the net force shown above, which is less than the force we would have to apply to lift it straight up. Like the case of the block & tackle, this "simple machine" helps us to get a job done while using less force than is required if it is done more directly. In the case of the block & tackle, to raise the mass some distance, the other end of the rope had to be pulled a greater distance. In this case we see a similar thing – the distance we must push the block is the hypotenuse of the triangle in order to raise it the height of the vertical leg of the triangle. This is a common theme in simple machines – less force is required, but it must be applied over a longer distance. In the next chapter we will see why this is the case.

We can complicate this simple example considerably. The most natural adjustment is to incorporate friction. Because of the nature of the two types of friction, adding static friction (which enters as an inequality) can be particularly troublesome. To see why, let's consider the following problem:

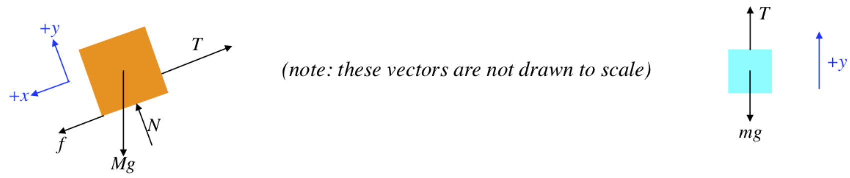
The system depicted in Figure 2.4.5 shows two blocks that remain at rest which are attached by a massless string over a massless, frictionless pulley. The plane is inclined at an angle $theta$ *up from the horizontal, and its surface is rough (i.e. not frictionless). The mass of the hanging block is given, as is the angle of incline and the coefficient of static friction. From these quantities, determine the minimum possible value of the mass of the block on the plane.*

Figure 2.4.5 – A Mechanical System



Starting (as always) with a FBD (including a coordinate system) for each block, we have:

Figure 2.4.6 – Free-Body Diagrams of Blocks



Let's take a moment to comment on the direction of the static friction force. Recall that a static friction force merely reacts to the "attempted" motion of the object along the surface. In this case, if the block was "trying" to slide down the plane, then the static friction force must be up the plane. Here it is drawn pointing down the plane, which means the other forces present must be such that they would accelerate it up the plane... How do we know this is the case? The answer lies in the statement of the question: We are looking for the minimum mass for the block on the plane. Imagine putting in a block whose mass balances the system. if a small mass is added to or subtracted from the block, the system may still remain at rest, as the static friction keeps the balance. If we add too much mass to that block, the static friction will reach its limit and the block will begin sliding down, while if we take away too much mass, it will slide up the plane. The static friction will oppose the intended motion, so for the minimum mass, the static friction force must point down the plane.

Alert

While we are able to determine the direction of the static friction for this problem, in many problems this is not possible. If you know that the static friction force is (or could be) present, just draw it in with either direction. When the problem is complete, if you solve for the value of this force, it will come out positive if you chose the correct direction, and negative if you did not. The FBD is just a tool, and in the end you end up with the answer, so don't waste energy worrying about getting the direction correct on the diagram.

Breaking the vectors into components in our chosen coordinate systems and applying Newton's second law (for zero acceleration) gives:

$$\begin{array}{lll} & x-direction\;forces: & 0=a_x=\dfrac{f-T+Mg\sin\theta}{M} \\ block\;on\;plane: & & \\ & y-direction\;forces: & 0=a_y=\dfrac{N-Mg\cos\theta}{M} \\ hanging\;block: & y-direction\;forces: & 0=a_y=\dfrac{T-mg}{m} \end{array} \tag{2.4.4}$$

Next, apply the constraint that relates the maximum static friction force (which occurs when the minimum mass is on the plane) and the normal force:

$$f\le\mu_S N\quad\Rightarrow\quad(maximum\;f\;for\;minimum\;M)\quad\Rightarrow\quad f=\mu_S N \tag{2.4.5}$$

The rest is algebra with four simultaneous equations, the result of which is:

$$M_{min}=\frac{m}{\mu_S\cos\theta+\sin\theta} \tag{2.4.6}$$

We should now check to see if this answer makes sense. If the angle $theta$ is 90^o, then both masses are hanging, and there is not friction force (because there is no normal force). For the system not to accelerate, the two masses must be equal. Plugging in $\theta=90^o$ indeed results in $M=m$. The $\theta=0^o$ case (a horizontal surface, where the normal force equals the weight of the block on the surface) will require that the friction force equals the weight of the hanging block. That is, we must have $mg=f=\mu_S N=\mu_S Mg\quad\Rightarrow\quad m=\mu_S M$, which is what we get when we plug in zero for θ.

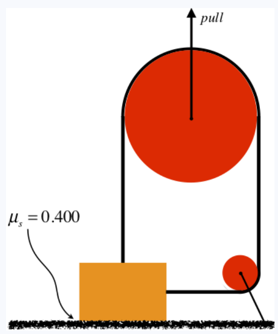
This problem could just as easily have asked for the *maximum* possible value for M. It is left as an exercise for the reader to try this.

There was a lot involved with this problem, but the key is to take it one step at a time and follow the following prescribed steps:

1. **draw a diagram**
2. **isolate the relevant objects and draw free-body diagrams for them**
3. **choose coordinate systems for the diagrams that are convenient**
4. **break forces into components in the coordinate system chosen**
5. **sum the forces in the x and y directions and apply Newton's second law for both directions**
6. **apply the constraints**
7. **solve the algebra**

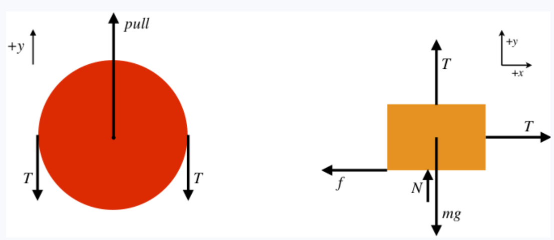
Example 2.4.2

A rope is fastened to a 50.0kg block in two places and passes through a system of two pulleys, as shown in the diagram below. The block rests on a rough (coefficient of static friction is 0.400) horizontal surface. The bigger pulley is then pulled upward with gradually increasing force. Both pulleys are massless and frictionless, and the rope is also massless. The smaller pulley is fastened to the floor and the both pulleys are positioned such that the rope is perpendicular to the floor on one end and parallel to it on the other. When the pull force reaches a certain magnitude, the block just barely begins to slide to the right. Compute the magnitude of this pull force.



Solution

This problem doesn't feature an inclined plane (though it could!), but it is a good example of the importance of following the prescription listed above. Start with the free-body diagrams and coordinate systems. The FBD of the smaller pulley will yield us nothing useful, so there are just two FBDs to draw. Note that the tension on the side of the block comes from the same rope as the tension on the top of the block, so they are equal:



The block is not accelerating at all (nor is the pulley), so the sum of the forces in each of the x and y directions comes out to zero.

$$\begin{array}{ll} block: & \begin{array}{l} x-direction\ forces: \quad 0 = T - f \\ y-direction\ forces: \quad 0 = T + N - mg \end{array} \\ pulley: & y-direction\ forces: \quad 0 = pull - 2T \end{array}$$

If we have to pull "just hard enough" to get the block moving, then this occurs when the horizontal pull equals the maximum static friction force, which gives us a constraint equation:

$$f = \mu_S N$$

Note that the block will have to start sliding before it starts rising, because rising requires than the normal force equals zero, and it will slide when the static friction force is small-but-non-zero. Now solve the equations simultaneously to get:

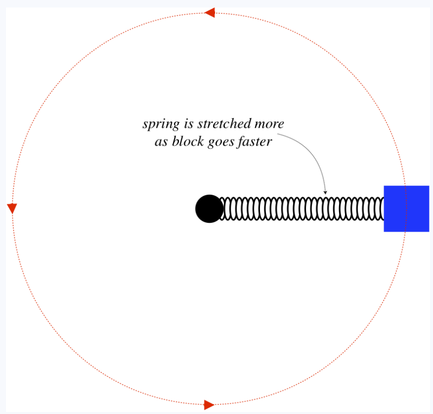
$$pull = \boxed{280N}$$

Incorporating Motion

We've done two examples involving systems for which the acceleration is zero. But of course it's possible that a problem could actually involve accelerating objects. Sometimes we are asked to find this acceleration, and other times the acceleration is a piece of information that is given. The acceleration can be given directly, or possibly it can be calculated in another way, perhaps from kinematics or if the motion is circular. Knowing something about the motion of the object falls under the "constraints" category, because the motion is specified (constrained), bringing in equations that don't result from Newton's second law. We'll look at an example that employs the steps to solving mechanics problems that also includes the added constraint of circular motion. Before we do, give the following example a try:

Example 2.4.3

A block is attached to one end of a massless spring, the other end of which is attached to a vertical fixed peg in a frictionless horizontal surface. The block is spun around a circle, and the spring stretches as a result of this motion. In fact, the faster the motion, the more the spring stretches. To stretch any spring, both ends need to be pulled simultaneously. Clearly the peg is pulling on one end of the spring as the block goes in the circle, but what force is pulled the block outward to stretch the spring?



Solution

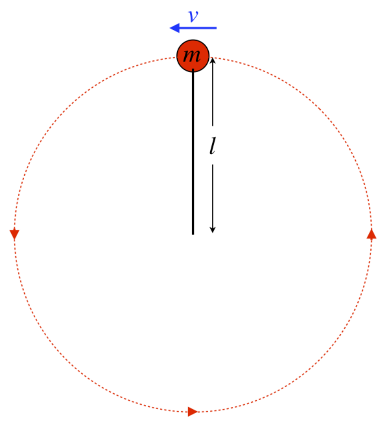
*The block is **not** pulled outward! It is only pulled inward (by the spring). It is not the block that needs to be pulled outward to stretch the spring, but rather the spring that needs to be pulled that way. The spring pulls the block inward (keeping it accelerating centripetally), and the third-law-pair force of the block on the spring is what pulls the spring outward.*

This points out possibly better than any other example the importance of isolating objects with force diagrams. The block here is not a conduit for some mysterious force pulling out on the spring – it is the object pulling out on the spring. You thoroughly need to trust the third law here to get the force between the spring and the block, and you need to thoroughly trust the second law to realize that the block does not require another force on it outward to balance the spring force, because it is accelerating.

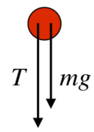
Now for the promised example that incorporates motion following the step-by-step prescription given earlier. What makes this problem interesting is the information that is hidden within the wording...

A rock on a string flies around in a circle in a vertical plane (in the presence of the earth's gravity) such that it just barely gets by the top (the string remains stretched to its full length, but there is no tension) as it continues in its circular path. Find the speed of the rock in terms of the length of the string.

- ***draw a diagram***



- ***isolate the relevant objects and draw free-body diagrams for them***



- **choose coordinate systems for the diagrams that are convenient**

We'll choose down as the positive y-direction.

- **break forces into components in the coordinate system chosen**

Unnecessary here.

- **sum the forces in the x and y directions and apply Newton's second law in both directions**

$$a=\frac{\sum F_y}{m}=\frac{T+mg}{m} \tag{2.4.7}$$

- **apply the constraints**

The first constraint is that the rock barely makes it around. What does this mean? To answer this, think about what would happen if the rock was moving any slower... It would fall out of the circle, which means the string would not remain straight. This would mean that the tension is zero. So the condition of "just making it around" is equivalent to requiring that the tension vanishes. The second constraint is that the rock is traveling in a circle, which requires that the acceleration is centripetal, with the radius of the circle equalling the length of the string (note there is no horizontal force at the top, so there is no tangential component to the acceleration). Stating both mathematically:

$$T=0, \qquad a=\frac{v^2}{l} \tag{2.4.8}$$

- **do the algebra**

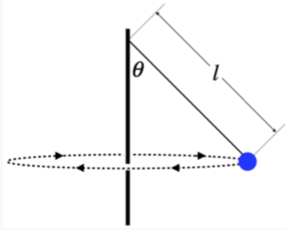
Simple enough:

$$v=\sqrt{gl} \tag{2.4.9}$$

While this is a pretty simple example in terms of the steps taken, it points out one very important aspect of these problems. If you don’t spend some time thinking about what is physically happening, you are likely to overlook the “hidden” information in the wording of the problem. This isn’t intended as a trick to trip you up – this is exactly what you run into in the real world when you need to solve a real problem. You need to be able to convert descriptive aspects of a system into mathematically-analyzable quantities.

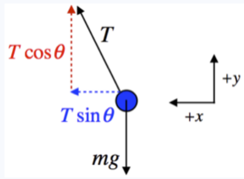
Example 2.4.4

A tetherball swings around a pole, making a full circle once every 1.5s. The total length of the rope is 2.4m. Calculate the angle θ that the rope makes with the pole. The rope has negligible mass.



Solution

Start with a force diagram of the ball, including a coordinate system:



Next sum the forces along the x and y axes and apply Newton's second law:

$$a_x = \frac{T\sin\theta}{m}$$
$$a_y = \frac{T\cos\theta - mg}{m}$$

The acceleration in the x-direction is centripetal, with the radius of its circular motion equal to the horizontal leg of the right triangle whose hypotenuse is the length of the string. The speed is constant, so it equals the circumference of the circle divided by the time for the trip. Putting all this together gives:

$$\left.\begin{array}{l} a_x = \dfrac{v^2}{R} \\ R = l\sin\theta \\ v = \dfrac{2\pi R}{t} \end{array}\right\} \Rightarrow \quad a_x = \frac{4\pi^2 l\sin\theta}{t^2}$$

The acceleration in the y-direction is zero, so plugging the accelerations into the equations from Newton's second law and eliminating $\frac{T}{m}$ from the two equations gives:

$$\left.\begin{array}{l} a_x = \dfrac{4\pi^2 l\sin\theta}{t^2} = \dfrac{T\sin\theta}{m} \quad \Rightarrow \quad \dfrac{T}{m} = \dfrac{4\pi^2 l}{t^2} \\ a_y = 0 = \dfrac{T\cos\theta - mg}{m} \quad \Rightarrow \quad \dfrac{T}{m} = \dfrac{g}{\cos\theta} \end{array}\right\} \Rightarrow \quad \theta = \cos^{-1}\left(\frac{gt^2}{4\pi^2 l}\right) = \boxed{76.5^o}$$

CHAPTER OVERVIEW

3: Work and Energy

Image used with permission (CC BY-SA-NC; anonymous).

3.1: The Work - Energy Theorem

Ignoring Directional Changes

For a large number of applications in mechanics, we are not interested in how a force causes the direction of motion of an object to change. In these cases, we only care about how that force changes the speed of the object. By now we know how much of a pain vectors can be, so having an alternative to Newton's second law to solve problems where only changes in speed are of interest is a welcome improvement. To see how we get to such a place, we need to go back to what we previously said about acceleration, and how it breaks into perpendicular parts – one that is parallel to the velocity (the "speeding-up/slowing-down" part), and the part that is perpendicular to the velocity (the "changing direction" part). We expressed this mathematically in Equation 1.6.12. We will now restrict our attention to the first term. Note that restricting ourselves to the part of the acceleration parallel to the direction of motion means we also restrict ourselves only to the component of the net force parallel to the motion.

Kinetic Energy and the Work-Energy Theorem

We have a neat trick that allows us to relate the change of the speed to the net force. The net force is proportional to the time derivative of the velocity vector, and we can use the product rule for derivatives of dot products of vectors, so let's take a derivative of the square of the velocity:

$$\frac{d}{dt}v^2 = \frac{d}{dt}\left(\overrightarrow{v}\cdot\overrightarrow{v}\right) = \frac{d\overrightarrow{v}}{dt}\cdot\overrightarrow{v}+\overrightarrow{v}\cdot\frac{d\overrightarrow{v}}{dt} = 2\overrightarrow{a}\cdot\overrightarrow{v} \tag{3.1.1}$$

To get to the net force, we multiply both sides by the mass of the object and divide both sides by 2:

$$\frac{d}{dt}\left(\frac{1}{2}mv^2\right) = \left(m\overrightarrow{a}\right)\cdot\overrightarrow{v} = \overrightarrow{F}_{net}\cdot\overrightarrow{v} \tag{3.1.2}$$

This makes some sense. The rate of change on the left side of this equation only depends upon the rate at which the speed changes (it is insensitive to changes in direction), and the dot product on the right side ensures that only the projection of the net force along the direction of motion (i.e. the direction of the velocity) plays a role. The part of the net force that causes the object to change direction is thrown away. We can take this a little bit further by expressing the velocity vector on the right side as a tiny displacement (which we will call $\overrightarrow{dl}$) divided by the tiny time interval. Multiplying both sides by dt then gives an equation that expresses a small change in the quantity $\frac{1}{2}mv^2$ (called the *kinetic energy)* due to a net force acting on the object as it displaces a small amount $\overrightarrow{dl}$.

$$\frac{d}{dt}\left(\frac{1}{2}mv^2\right) = \overrightarrow{F}_{net}\cdot\frac{\overrightarrow{dl}}{dt} \quad\Rightarrow\quad d\left(\frac{1}{2}mv^2\right) = \overrightarrow{F}_{net}\cdot\overrightarrow{dl} \tag{3.1.3}$$

Suppose the object now undergoes several displacements, so that the change in the kinetic energy is no longer infinitesimal. This is tricky business, as each displacement may be the same (if it moves in a straight line), or it may change direction (if it follows a curvy path). Also, the net force on the object might change as the object moves from one place to another. We express the sum of many infinitesimals as an integral, and since the sum of the right side of this equation depends upon the directions of many displacements, this particular type of integral is called a *line integral*. This does not mean that the displacements are along a straight line, however – here the word "line" is rather misleading – the word "trajectory" might be better.

Of course, the left side of this equation is simply a small number, and adding those up does not depend upon anything as complicated as a trajectory, so it ends up being just a change from the beginning of the path to the end. If we call the start of the journey A and the end B, then we can express the totals for the whole journey as:

$$\Delta\left(\frac{1}{2}mv^2\right) = \frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2 = \int_A^B \overrightarrow{F}_{net}\cdot\overrightarrow{dl} \tag{3.1.4}$$

The line integral on the right side of this equation is called the *work done* (by the net force) going from the initial to final positions. We can (and later, will) discuss the work done by individual forces, and the work done by the net force is the total of all of those works. We will often write the above equation with the following abbreviated notation:

$$\Delta KE = W_{tot}\left(A \rightarrow B\right) \tag{3.1.5}$$

In words, this reads: "The change of an object's kinetic energy when it changes its position from A to B equals the work done on it by all forces on it, computed over a well-defined path connecting those endpoints." This is known as the *work-energy theorem*. It does exactly what we set out to do – it expresses the effect forces have on the change in an object's speed, with no regard to its directional changes. It doesn't solve any problem that can't be solved by Newton's second law, and in fact for some cases it isn't even any easier to work with. But for other cases is it *much easier* to work with, as we will see, and these are the cases for which this approach was invented.

These new quantities of kinetic energy and work have units of what we will more generically refer to as energy, and we give energy units their own name:

$$[KE] = \frac{kg \cdot m^2}{s^2} = \text{"}\; Joules \;\text{"} \quad (J)$$

Example 3.1.1

A single force which varies in magnitude and direction in space acts upon an object, and is given by the equation below. Find the change in the object's kinetic energy as it moves from the origin along the $+x$-axis a distance of 2m.

$$\overrightarrow{F}(x,y) = \left(\alpha x^2 + \beta y^3\right)\hat{i} + \left(\beta x^3 + \alpha y^2\right)\hat{j}, \quad where: \;\; \alpha = 2.4\frac{J}{m^2} \;\; and \;\; \beta = 4.5\frac{J}{m^3}$$

Solution

This is a direct application of the work-energy theorem, which means it consists entirely of computing a line integral. To do this, we first need to define the path mathematically, and all of the tiny displacements $\overrightarrow{dl}$ along that path. The path in this case is pretty simple – it is a straight line along the x-axis from the origin ($0m, 0m$) to the point ($2m, 0m$). Along this path, the value of y remains a constant zero. The direction of every infinitesimal displacement is the $+\hat{i}$ direction, and the magnitude of each displacement is simply dx. The work integral therefore becomes:

$$W\left(A \rightarrow B\right) = \int_A^B \overrightarrow{F} \cdot \overrightarrow{dl} = \int_{x=0m}^{x=2m} \overrightarrow{F} \cdot \left(dx\;\hat{i}\right)$$

Now we just need to plug in for the force. The force must be evaluated at each point on the path, and since the value of y is zero on the entire path, we can set $y = 0$ in the force vector, simplifying things greatly:

$$\overrightarrow{F}\left(on\ the\ path\right) = \left(\alpha x^2\right)\hat{i} + \left(\beta x^3\right)\hat{j}$$

The dot product of this vector with the tiny displacement vector simplifies things even more:

$$\overrightarrow{F}\left(on\ the\ path\right) \cdot \overrightarrow{dl} = \left[\left(\alpha x^2\right)\hat{i} + \left(\beta x^3\right)\hat{j}\right] \cdot \left[dx\;\hat{i}\right] = \alpha x^2 dx$$

Finally, we just perform the integral and apply the work-energy theorem:

$$\Delta KE = \int_{0m}^{2m} \alpha x^2 dx = \left[\frac{1}{3}\alpha x^3\right]_{0m}^{2m} = \boxed{6.4J}$$

Line Integrals

As you can tell from the example above, the hardest part of using the work-energy theorem is setting up the line integral. There are several elements that need to be kept in mind:

1. **define a direction for the tiny displacement vectors for every point on the path**

The direction of the tiny displacement vectors (which we will assume to be in the (x, y) plane will have components equal to the displacements in the x and y directions:

$$\overrightarrow{dl} = dx\ \hat{i} + dy\ \hat{j} \tag{3.1.6}$$

2. **write the magnitude of the tiny displacements in terms of the integration variable**

The displacement vector as written above doesn't tell us much. We also need to include the *path* for this to be useful. Since we are assuming that everything is in the (x, y) plane, the path can be expressed as a relationship between the variables x and y. For example, if the path is a straight line, then we can write $y = mx + b$. In this case, we can replace the dy in the displacement vector:

$$\frac{dy}{dx} = m \quad \Rightarrow \quad dy = m\ dx \quad \Rightarrow \quad \overrightarrow{dl} = \left[\hat{i} + m\hat{j}\right] dx \tag{3.1.7}$$

This puts the displacement vector in terms of a single variable (x) for integration (we could of course have instead chosen our integration variable to be y). More generally, the path could be a function: $y = f(x)$, in which case the m above would be replaced by the derivative of the function. Note also that the path may not even be a function, since it could have multiple y values for each x value. [Suffice to say that path integrals have a lot more going on than we will cover in this course, and we'll leave coverage of the more nuanced details to a course in vector calculus.]

3. **evaluate the force vector at each point in the path**

The force vector will be in terms of x and y (i.e. it is defined at all points in space), but in the integral only its value along the path matters, so we can substitute the equation that defines the path (such as $y = mx + b$ in the case of a straight-line) into the force vector so that it is a function of only one variable, allowing us to do the integral.

4. **take the dot product**

We of course know how to do this by now, but it is important to remember that it must be done. This step goes back to the start of our discussion of this method. This dot product assures that we are only using the part of the force vector that lies along the tiny displacement, which means we are only using the part of the force vector that changes the speed of the object.

Of course, much more complicated paths than straight lines are possible. The following example illustrates how this is handled.

Example 3.1.2

Compute the work done on an object by the force given below, along a parabolic path $y = \lambda x^2$ connecting the origin to the point on the path with an x value of 1m, where $\lambda = 0.4m^{-1}$.

$$\overrightarrow{F}(x, y) = \alpha x^2\ \hat{i} + \beta y\ \hat{j}, \quad where: \quad \alpha = 1.5\frac{J}{m^2} \quad and \quad \beta = 3.0\frac{J}{m}$$

Solution

Start by determining the displacement vector as a function of x along the path:

$$\left.\begin{array}{l} \overrightarrow{dl} = dx\ \hat{i} + dy\ \hat{j} \\ y = \lambda x^2 \quad \Rightarrow \quad \dfrac{dy}{dx} = 2\lambda x \quad \Rightarrow \quad dy = 2\lambda x dx \end{array}\right\} \Rightarrow \overrightarrow{dl} = \left(\hat{i} + 2\lambda x\ \hat{j}\right) dx$$

Next write the force vector along the path only (in terms of x):

$$\left.\begin{array}{l} \overrightarrow{F}(x, y) = \alpha x^2\ \hat{i} + \beta y\ \hat{j} \\ y = \lambda x^2 \end{array}\right\} \quad \Rightarrow \quad \overrightarrow{F}_{on\ path}(x) = \alpha x^2\ \hat{i} + \beta\lambda x^2\ \hat{j}$$

Now for the dot product:

$$\overrightarrow{F}_{on\ path} \cdot \overrightarrow{dl} = \left[\left(\alpha x^2\right)(1) + \left(\beta\lambda x^2\right)(2\lambda x)\right] dx = \left(\alpha x^2 + 2\beta\lambda^2 x^3\right) dx$$

And finally integrate between the two endpoints, defined in terms of the x variable that we have put everything in terms of:

$$W(A \rightarrow B) = \int\limits_{x=0m}^{x=1m} \left(\alpha x^2 + 2\beta\lambda^2 x^3\right) dx = \left[\frac{1}{3}\alpha x^3 + \frac{1}{2}\beta\lambda^2 x^4\right]_{x=0m}^{x=1m} = \boxed{0.74J}$$

Lost Information

It is important to note that while the introduction of the work-energy theorem will simplify things for us with a subset of problems, we do sacrifice some information. By throwing out the part of the force that acts to change the direction of the object, we cannot use this method to determine which way the object is moving after the force acts on it – we only know how fast it is going. Also, we lose information about the time element of the motion between the starting and ending points. This should not be surprising – just because we know how fast something is moving, if we don't know the directions it takes to get from start to finish, we still don't know anything about the elapsed time. For example, a projectile thrown into the air will reach the same speed at two different points of time – once on the way up and once on the way down. If we don't know anything about the direction of motion, we don't know which time we are looking at.

To see this another way, consider a situation we are very familiar with – an object moving in a straight line, accelerating at a constant rate. We know that we can write its acceleration in terms of the starting and final velocities using Equation 1.4.3:

$$2a\Delta x = {v_f}^2 - {v_o}^2 \tag{3.1.8}$$

By Newton's second law, the acceleration here must have been caused by a (net) force in the same direction, so substituting the ratio of force/mass for the acceleration gives:

$$2\frac{F_{net}}{m}\Delta x = {v_f}^2 - {v_o}^2 \quad \Rightarrow \quad F_{net}\Delta x = \frac{1}{2}m{v_f}^2 - \frac{1}{2}m{v_o}^2 \tag{3.1.9}$$

This is once again the work-energy theorem (in one dimension, for a constant net force), and we see that it came directly from the kinematics equation from which the time variable had been eliminated.

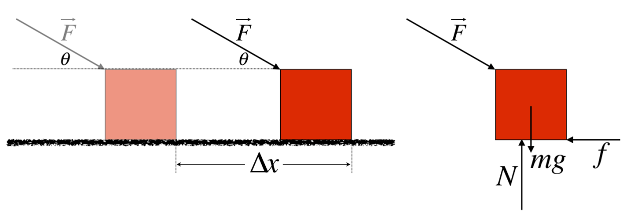
Work Contributions of Individual Forces

It probably isn't immediately clear what is to be gained from this work-energy approach. After all, one still has to determine the net force at each point in the path of the object's motion, so our attempt to escape the tyranny of vectors would appear to be a failure. But there is much more to this story. It begins with the recognition that total work done can be broken into a sum of works done by individual forces:

$$\begin{aligned} W_{tot}\left(A\rightarrow B\right) &= \int_A^B \overrightarrow{F}_{net}\cdot \overrightarrow{dl} \\ &= \int_A^B \left(\overrightarrow{F}_1 + \overrightarrow{F}_2 + \dots\right)\cdot \overrightarrow{dl} \\ &= \int_A^B \overrightarrow{F}_1\cdot \overrightarrow{dl} + \int_A^B \overrightarrow{F}_2\cdot \overrightarrow{dl} + \dots \\ &= W_1\left(A\rightarrow B\right) + W_2\left(A\rightarrow B\right) + \dots \end{aligned} \tag{3.1.10}$$

There are a number of advantages to this, but the one we can see immediately is that if one of the individual forces happens to be everywhere perpendicular to the path of the object from A to B, then the work it contributes is zero, and we can simply ignore it – no need to do the vector addition to add it to the other forces. Consider the following example of pushing a block across a rough horizontal surface. Figure 3.1.1 shows a diagram of what is happening and a FBD of the block.

Figure 3.1.1 – Pushing Block Across a Rough Horizontal Surface



The work done by the net force can be broken down into a sum of the works done by each individual force:

$$W_{applied\ force} = \int_A^B \overrightarrow{F} \cdot \overrightarrow{dl} = F\Delta x\cos\theta \tag{3.1.11}$$

$$W_{friction} = \int_A^B \overrightarrow{f} \cdot \overrightarrow{dl} = f\Delta x\cos 180^o = -f\Delta x \tag{3.1.12}$$

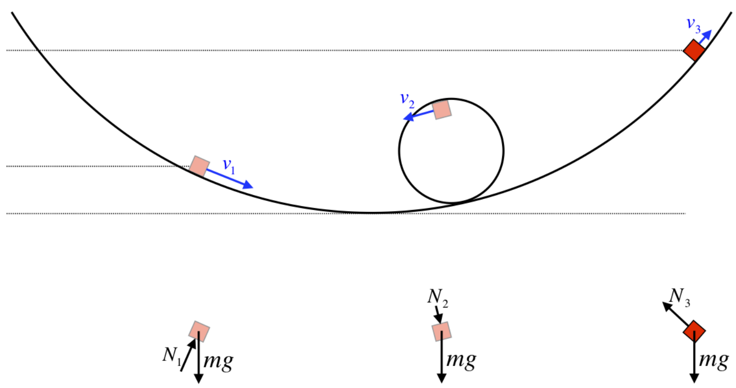
$$W_{gravity} = \int_A^B \left(-mg\,\hat{j}\right) \cdot \overrightarrow{dl} = mg\Delta x\cos 90^o = 0 \tag{3.1.13}$$

$$W_{normal} = \int_A^B \left(N\,\hat{j}\right) \cdot \overrightarrow{dl} = N\Delta x\cos 90^o = 0 \tag{3.1.14}$$

$$W_{tot} = W_{applied\ force} + W_{friction} + W_{gravity} + W_{normal} = (F\cos\theta - f)\,\Delta x = F_{net}\Delta x \tag{3.1.15}$$

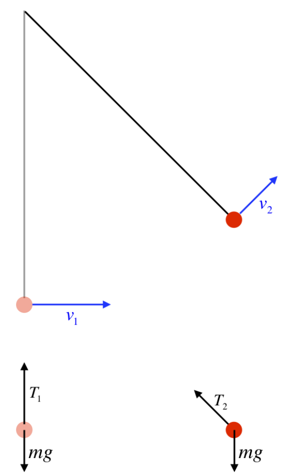
Just by looking at the physical situation it is clear that the gravity and contact forces will play no role in the total work done, as they are always perpendicular to the motion. This greatly reduces the number of forces (and vector addition) we would otherwise have to deal with. Let's look at some even more compelling examples:

Figure 3.1.2 – Loop-de-Loop



For a block sliding around a frictionless loop-de-loop track, the path it follows is quite complicated. The FBD of the block as it travels along the track includes only two forces – gravity and the normal force by the track. The motion of the block is parallel to the track everywhere, which means it is perpendicular to the normal force everywhere. That means that no matter what our starting and ending points are, the normal force does no work on the block! Of the two forces involved, the normal force is by far the hardest to deal with, since its direction and magnitude change everywhere on the track. but if we are only interested in the speed of the block, we only need to worry about the work done by the gravity force, which has a constant direction and magnitude. We'll come back to the simple result that comes from this shortly.

Figure 3.1.3 – Simple Pendulum



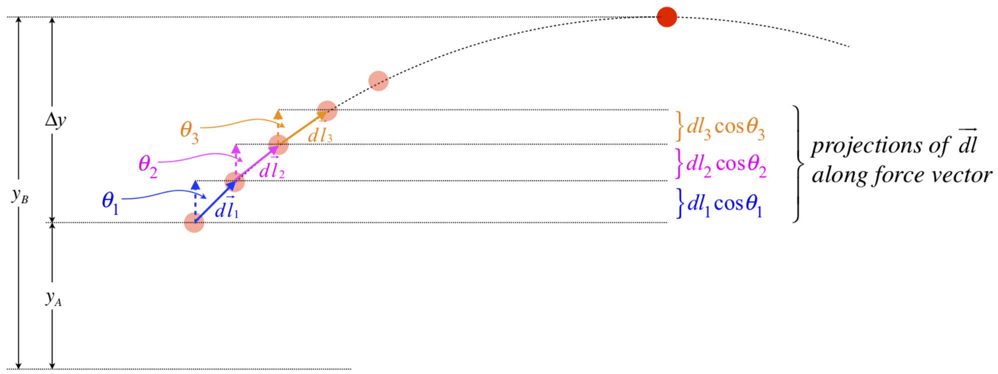
For the simple pendulum, we see the same result for the tension as we found for the normal force in the loop-de-loop example. The tension force remains at right angles to the motion of the bob at the end of the string, so there is no work done by the tension force. If all we care about is the speed of the bob, then we only need to compute the work done by gravity.

3.2: Conservative and Non-Conservative Forces

Gravity is a Conservative Force

Let's take a closer look at the work contributions by some of the various forces we have discussed. We'll start with gravity. In our earlier examples of a loop-de-loop and a pendulum, we saw that the contact and tension forces didn't contribute to work on the object, but what about gravity? Yes, of course it does work, because in both cases, for much of the journey, the gravity force on the object was not perpendicular to its motion. We'll look at these cases again, but first let's consider the work done by gravity on a projectile.

Figure 3.2.1 – Work Done on Projectile by Gravity

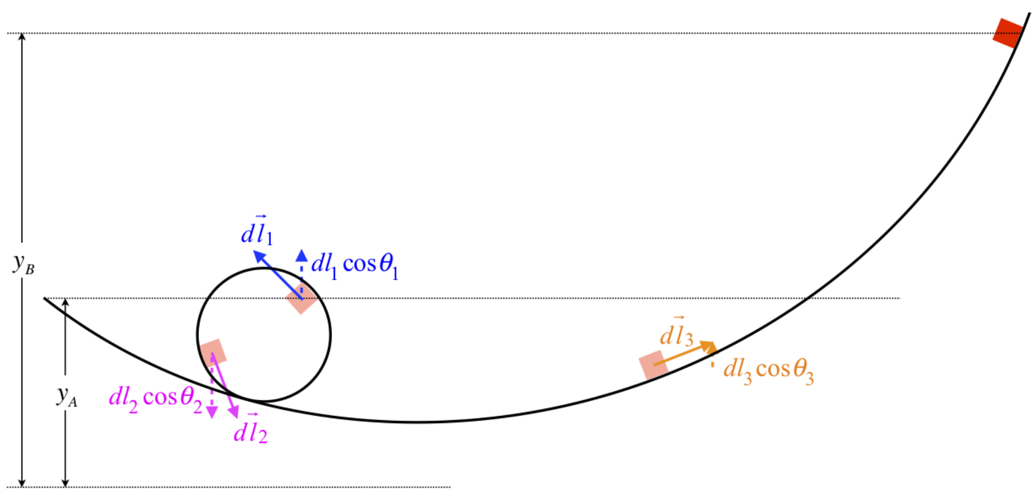


To compute the total work done by gravity as the ball rises to height y_B from y_A, we add up all the dot products of the force with the tiny displacements. Each of these dot products equals the magnitude of the force and the projection of the displacement vector on that force. In this case, the force is down, and the projection is up, so a minus sign also appears. The result is:

$$W_{grav}\left(A\rightarrow B\right)=\int_A^B\overrightarrow{F}\cdot\overrightarrow{dl}=-mg\left(dl_1\cos\theta_1+dl_2\cos\theta_2+dl_3\cos\theta_3+\ldots\right)=-mg\left(y_B-y_A\right)=-mg\Delta y \tag{3.2.1}$$

Now suppose that instead of a projectile, we look at the previous example of the block on the loop-de-loop.

Figure 3.2.2 – More Work Done by Gravity



If we add up all the contributions to the work in this case, while we get some positive as well as negative contributions, we find the same result as before:

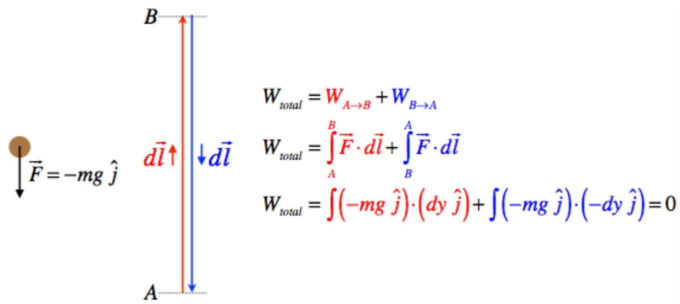
$$W_{grav}\left(A\rightarrow B\right)=\int_A^B\overrightarrow{F}\cdot\overrightarrow{dl}=-mg\left(dl_1\cos\theta_1-dl_2\cos\theta_2+dl_3\cos\theta_3+\ldots\right)=-mg\left(y_B-y_A\right)=-mg\Delta y \tag{3.2.2}$$

We therefore get the remarkable result that the work done by gravity doesn't seem to depend upon the path taken from point A to point B, but only depends upon the positions (specifically, the heights) of those points. We could do the same analysis for the pendulum, and get the same answer. This property of the work done depending only on the endpoints is not true of all forces, but those forces for which this property holds are called *conservative forces* (we'll see where this name comes from shortly). Notice that an object that comes back to the position where it started (i.e. follows a closed path), no matter how crazy the route it takes might be, will have no net work done on it by a conservative force. We can represent this abstractly by:

$$W_{conservative\ force}\left(A\rightarrow A\right)\equiv\oint\overrightarrow{F}\cdot\overrightarrow{dl}=0 \tag{3.2.3}$$

The circle in the integral sign means "around a closed path." Let's show this property explicitly for gravity.

Figure 3.2.3 – Round Trip Work Done by Gravity



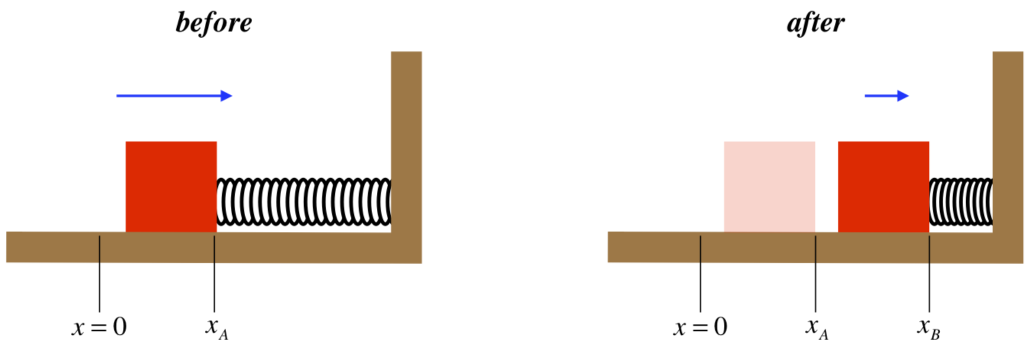
Elastic Force is Conservative

There is a force we have only discussed very briefly that we should also look at – the elastic force. Recall that this force obeys Hooke's law:

$$\overrightarrow{F}_{elastic}=-k\Delta\overrightarrow{x}, \tag{3.2.4}$$

where the magnitude of $\Delta\overrightarrow{x}$ is the amount that the spring is stretched from its equilibrium position. Let's compute the work done by a spring on a block as the spring is compressed. In Figure 3.2.4, $x=0$ represents the position where the spring is at its equilibrium length.

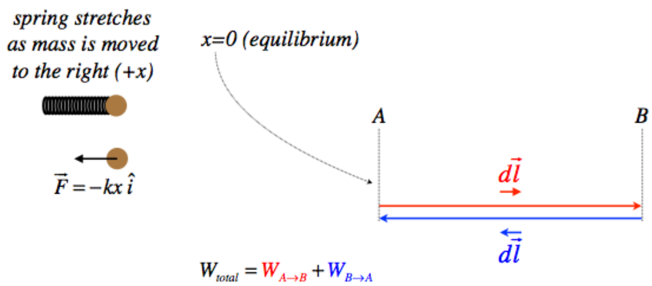
Figure 3.2.4 – Work Done by a Compressing Spring



$$W_{spring}\left(A\rightarrow B\right)=\int_A^B\overrightarrow{F}\cdot\overrightarrow{dl}=\int_A^B\left(-kx\,\hat{i}\right)\cdot\left(dx\,\hat{i}\right)=-k\int_A^B x\,dx=\left[-\frac{1}{2}kx^2\right]_A^B=-\left[\frac{1}{2}kx_B^2-\frac{1}{2}kx_A^2\right]= -\frac{1}{2}k\Delta\left(x^2\right) \tag{3.2.5}$$

While this result isn't exactly the same as that for gravity (the work done depends upon the change of a quadratic rather than a linear function), it is strikingly similar in that the result only depends upon the endpoints. We therefore conclude that this force is also conservative. We can check the round-trip criterion easily enough to be sure. If we move the block from the equilibrium point to the right, compressing the spring, and then back to the equilibrium point, the force always points left, while the displacement vectors reverse direction, giving:

Figure 3.2.5 – Round Trip Work Done by Spring



$$W_{total} = W_{A\rightarrow B} + W_{B\rightarrow A}$$

$$W_{total} = \int_A^B \vec{F}\cdot\vec{dl} + \int_B^A \vec{F}\cdot\vec{dl}$$

$$W_{total} = \int\left(-kx\,\hat{i}\right)\cdot\left(dx\,\hat{i}\right) + \int\left(-kx\,\hat{i}\right)\cdot\left(-dx\,\hat{i}\right) = 0$$

Kinetic Friction is Not a Conservative Force

Let's have a look at kinetic friction next. We'll choose a simple situation – a block sliding on a horizontal rough surface. In this case, the kinetic friction force always opposes the motion of the sliding object, which means the force is *always* opposite to the displacement vector. This prevents the round-trip cancelation of work that we saw for the cases of gravity and elastic force.

Figure 3.2.6 – Round Trip Work Done by Kinetic Friction

$$W_{total} = W_{A\rightarrow B} + W_{B\rightarrow A}$$

$$W_{total} = \int_A^B \vec{F}\cdot\vec{dl} + \int_B^A \vec{F}\cdot\vec{dl}$$

$$W_{total} = \int\left(-\mu_k N\,\hat{i}\right)\cdot\left(dx\,\hat{i}\right) + \int\left(+\mu_k N\,\hat{i}\right)\cdot\left(-dx\,\hat{i}\right) \neq 0$$

We see that in fact the kinetic friction force does not result in zero work done for a closed path. This is an example of a *non-conservative force*. Unlike gravity and the elastic force, work done by the kinetic friction force does not simply depend upon the starting and ending points of the object's motion, but rather depends upon the path followed.

Alert

It only takes one example of the work calculation being dependent upon the path to demonstrate that a force is non-conservative, but to declare a force to be conservative, the work done must be independent of the path for ***all*** *paths. This makes rigorously proving a force is conservative a bit tougher than proving one is non-conservative. Fortunately, in practice it's unusual to find a force that does the same work for two paths with common endpoints, while it does a different amount of work for other paths between those endpoints.*

Example 3.2.1

Show that the force given below is not conservative by computing the work done it would do on an object that moves from the origin to the point (1,1) in the x-y plane for two different paths. The first path goes across-then-up ("atu") – it follows the x-axis from the origin to the point (1,0), then goes parallel to the y-axis up to (1,1). The second path goes up-then-across ("uta") – it follows the y-axis to the point (0,1), then goes parallel to the x-axis over to (1,1).

$$\overrightarrow{F}(y) = \alpha y\hat{i}$$

Solution

For the atu path, in the entire first leg (along the x-axis), the value of y is zero, which makes the force magnitude zero. So there is zero work done by this force in this leg. Then in the second leg of the atu journey, the force magnitude is no longer zero (as y grows from 0 to 1), but the force points in the $+\hat{i}$ direction, which is perpendicular to the direction of motion which is in the

$+\hat{j}$ direction), resulting in a zero dot product. Therefore the force does no work during this leg also, which means no work at all is done for the atu path.

The first leg of the uta path is the same as the second leg of the atu path - the force is perpendicular to the direction of motion, and therefore does no work. In the second leg of the uta path, however, the value of y is 1 (not zero), and the force points in the same direction as the motion, so there is some positive work done during this leg. There is therefore some work done over the uta path, but not over the atu path, which means that the work done is path dependent, and the force is non-conservative.

3.3: Mechanical Advantage and Power

Reconciling Work with Mechanical Advantage

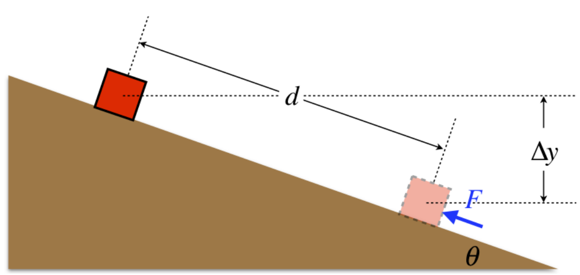
Back when we first talked about pulleys, we said that the block-and-tackle device was useful for lifting heavy objects. Figure 3.3.1 shows two blocks lifted the same distance by pulling on a rope in a pulley system. We know that it requires less force to lift the same mass for the case on the left than the case on the right, but now let's compare the amount of work done by the pull force in the two cases. In both cases, the block is raised the same distance, and in both cases it starts and ends at rest.

Figure 3.3.1 – Work Done with Pulley Systems

The pull force acts downward on the end off the rope, and the direction the end of the rope moves is downward, so there is positive work done in both cases. With identical blocks, the force required to be applied to the rope for the left case is half as great as the force required to lift the block in the right case. However, in order to lift the block the same distance from where it started, the rope must be pulled *twice as far* in the right case than in the left case, thanks to the pulley ratio constraint. With half the force acting over twice the distance, the amount of work done is the same. The same result comes from more pulleys, since the pulley constraint ratio is the same as the mechanical advantage in every case.

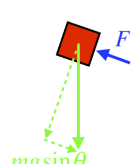
So it appears that the block & tackle (and simple machines more generally) trade effort (force) for displacement, such that the amount of work done remains the same. Let's see if we see a similar result for the case of the inclined plane.

Figure 3.3.2 – Work Done with an Inclined Plane



Let's assume that the inclined plane shown in Figure 3.3.2 is frictionless, that the force applied to the block is parallel to the plane, and that it is just enough so that the block moves at a constant speed up the plane. A FBD (without the irrelevant normal force which acts perpendicular to the motion) looks like:

Figure 3.3.3 – Partial FBD of Block on Plane



For the block to stay at a constant speed, the force up the plane must equal the force down the plane, which means:

$$F = mg\sin\theta \tag{3.3.1}$$

So the force required is reduced by the factor $\sin\theta$ compared to lifting the block straight up. But from trigonometry, the distance the block must be pushed is:

$$d = \frac{\Delta y}{\sin\theta} \tag{3.3.2}$$

When we multiply the force by the distance to get the work done, we get:

$$W = Fd = (mg\sin\theta)\left(\frac{\Delta y}{\sin\theta}\right) = mg\Delta y \tag{3.3.3}$$

This is the amount of work required to lift the block straight up at a constant speed, so once again the simple machine trades extra distance for less force to get the same work. How do we know that $mg\Delta y$ is the work required to lift the block up at a constant speed? We know that to lift it at a constant speed, the net vertical force needs to be zero, which means the upward force needs to balance the gravity force. Multiplying that force by the distance it is raised gives the straight-lift work of (Δy).

Power

We take a moment now to introduce yet another physics word whose common usage in English is very different from its meaning in physics.

Definition: Power

Power is the rate at which work is performed.

Note that just like we can talk about the work done by an individual force or a collection of forces, we can also talk about the power "delivered" to a system by one or more forces. For example, if a car is moving at a constant speed on level ground, its kinetic energy is not changing over time, so no total work is being done on it. If no work is done on it over time, there is no power being delivered to it. But clearly the engine of the car is doing *something*. So it is useful to break up the power delivered by separate sources if we want to isolate the rate at which the engine is doing work, without worrying about the rate at which air resistance and friction are are doing negative work on the car to bring the total to zero.

Mathematically, we therefore have for the power delivered by a given force is:

$$P \equiv \frac{dW}{dt} \tag{3.3.4}$$

Since the units of work is Joules, the units of power is Joules per second, which we rename as: *watts* (W).

One nice shortcut for power involves the force doing the work and the velocity of the object on which the work is being performed:

$$dW = \overrightarrow{F}\cdot\overrightarrow{dl} \quad\Rightarrow\quad P = \frac{dW}{dt} = \frac{\overrightarrow{F}\cdot\overrightarrow{dl}}{dt} = \overrightarrow{F}\cdot\frac{\overrightarrow{dl}}{dt} = \overrightarrow{F}\cdot\overrightarrow{v} \tag{3.3.5}$$

Note that this is the power delivered to the moving object (i.e. the rate at which energy is added to or taken away from the object) at the *instant* that the force and velocity are the vectors given above. If this is integrated from a starting time to a final time, the result is the total work done over that time span by the force.

Example 3.3.1

An object with a mass of $0.400kg$ experiences a net force of $12.0N$ that causes it to speed up as it moves in a circular path of radius $0.800m$. Find the power delivered to the object at the moment that its speed reaches $3.00\frac{m}{s}$.

Solution

The net force consists of two perpendicular components: One that speeds up the object, and one that keeps it moving in a circle. We start by calculating the component of the net force that is causing the centripetal acceleration:

$$F_\perp = m\frac{v^2}{R} = 4.50N$$

With this component of the force and the magnitude of the total force, we can compute the component of the force tangent to the circle:

$$F_{\parallel} = \sqrt{F^2 - F_{\perp}^2} = 11.1N$$

This is the part of the net force that is parallel to the velocity, so the dot product of the net force and the velocity is just this number multiplied by the speed:

$$P = \overrightarrow{F}_{net} \cdot \overrightarrow{v} = F_{\parallel}\ v = \boxed{33.3W}$$

3.4: Energy Conservation Models

Systems

Now that we are talking about work done by individual forces, we should make some mention of Newton's third law. Every individual force is an interaction with two equal-and-opposite forces involved, so how do we know which one of these to use when computing the work done on an object? The work done *on* an object is calculated using the force *on* it and the displacement of that same object. So when one drops an apple to earth, the gravity force on the apple multiplied by the distance it drops is the work done on the apple by the gravity force. There is also work done on the earth, as it experiences an equal-and-opposite force, but it displaces much less, so the work done on the earth is much less. So while we can use work as a sort of proxy for force, it is not the same as force - there is not a "Newton's third law" for work.

One thing that will help us keep things straight going forward is the notion of a *system*. A system is a collection of one or more objects that we have arbitrarily lumped together for accounting purposes. This grouping is isolated from other objects (much like an object in a force diagram), so that forces on it can be taken into account (while other distracting forces can be ignored). With this distinction made, we can also distinguish the work done on a single object as work that is coming from within the system (which may contain other objects) and work that is coming from outside. Breaking the work done into those two types gives a work-energy theorem that looks like this:

$$W_{tot} = \left(W_1 + W_2 + \ldots\right)_{outside} + \left(W_1 + W_2 + \ldots\right)_{inside} = \Delta KE \tag{3.4.1}$$

For reasons we will see soon, we'll rearrange terms in such a way that the work done by forces between objects within the system is on the same side of the equation as the change of kinetic energy:

$$\left(W_1 + W_2 + \ldots\right)_{outside} = \Delta KE - \left(W_1 + W_2 + \ldots\right)_{inside} \tag{3.4.2}$$

This might seem a bit abstract, so let's return to our example of a block pushed along a rough horizontal surface (Figure 3.1.1). Suppose we make the surface + block a single system. Then how is this equation arranged? The applied force comes from outside the system, and actually so does the gravity force, since we did not include the earth as part of the system. The work done by those forces appear on the left-hand side of the equation. The contact and normal forces are between the block and the table, which are both within the system, so the work done by those forces come from inside the system and appear on the right-hand side (with a minus sign).

> **Alert**
>
> *In the future, unless stated otherwise, we will assume that the earth is within the system, so that the work done by gravity always appears on the right-hand side. Note that the work done by gravity on the earth will always be negligible, as the earth's displacement due to a gravitational interaction with a terrestrial object is vanishingly small.*

This may all seem like pointless rearranging, but stay tuned – it will help a great deal with the accounting later.

Total, Mechanical, Potential, and Thermal Energy

So far we have been dancing around the word "energy." We have only briefly mentioned kinetic energy, in the context of the work-energy theorem, but the very existence of such a term implies the existence of another form of energy. Indeed we do define other forms of energy. We do this in terms of the work done by forces between objects within the system. The summary of this model is:

- The work done on a system by forces that come from outside the system changes the energy contained in that system.
- There are three different types of energy changes that can occur within a system:
  - kinetic energy, from the motions of the object(s) within the system
  - changes from work done within the system by conservative forces
  - changes from work done within the system by non-conservative forces

$$\underbrace{W_1 + W_2 + \ldots}_{outside} = \Delta KE \underbrace{- \left(W_1 + W_2 + \ldots\right)_{conservative} - \left(W_1 + W_2 + \ldots\right)_{non-conservative}}_{inside} \tag{3.4.3}$$

So far this is just fancy bookkeeping, but next we'll see how it all comes together. Recall that conservative forces result in work contributions that depend only upon the starting and ending positions of the object moved while under the influence of that force. We therefore define *potential energy* as a quantity of energy that a system possesses due to the position(s) of the object(s), and the change of that value is (the negative of) the work done by the conservative force:

$$\Delta PE = PE_B - PE_A = -W\left(A \rightarrow B\right) \tag{3.4.4}$$

There are many forms of potential energy – indeed there is a form of PE for every conservative force. Notice that we cannot define a potential energy for a non-conservative force, because PE is defined at specific points in space, so the difference in PE for two points in space is always the same. But for a non-conservative force, different paths between the same two points result in different amounts of work. Incorporating potential energy into our model changes Equation 3.4.3 into:

$$W_{ext} = \Delta KE + \Delta PE_1 + \Delta PE_2 + \dots - (W_1 + W_2 + \dots)_{non-conservative} \tag{3.4.5}$$

The work done by non-conservative forces between objects within the system is also accounted-for by a change in a quantity of energy, but this quantity does not depend upon the starting and ending positions of objects. Despite this difference with the case of potential energy, we can still write the energy removed from mechanical energy due to non-conservative forces as an increase of energy of a different form – thermal energy. This converts our energy formula into one that contains only changes of forms of energy on the right-hand-side:

$$W_{ext} = \Delta KE + \Delta PE_1 + \Delta PE_2 + \dots + \Delta E_{thermal} \tag{3.4.6}$$

This framework gives us a totally new picture for thinking about energy, if we just choose our system to be “everything.” Then the work done by forces from outside the system is zero (there are no objects to interact with our system from outside!), and the work-energy theorem morphs into:

$$0 = \Delta KE + \Delta PE_1 + \Delta PE_2 + \dots + \Delta E_{thermal} \tag{3.4.7}$$

When it comes to using this formula, there are few things to note:

- We know ΔKE if we know the mass and the beginning and ending speeds of the object.
- We know ΔPE if we know the type of conservative force and the beginning and ending positions of the object. (We will soon see how we do this.)
- We know $\Delta E_{thermal}$ if we know the specific non-conservative force acting, and the path the object takes. However, it is commonly the case that the details of these forces are not provided, in which case either the thermal energy change is given directly, or it is calculated from the above equation.

You may be getting tired of all the renaming of quantities, but there is one more that is helpful for talking about energy. Kinetic and potential energy share the property that they can be converted back-and-forth freely, but thermal energy is a one-way ticket. Once energy is converted to thermal, it doesn’t return to kinetic or potential – we don’t see objects suddenly cool off and start moving. Accordingly, we give the description *mechanical energy* to the sum of kinetic and potential energy, to distinguish them from thermal energy:

$$0 = \underbrace{\Delta KE + \Delta PE_1 + \Delta PE_2 + \dots}_{\Delta ME} + \Delta E_{thermal} \tag{3.4.8}$$

What we have expressed here is known as a *conservation principle*. In physics, for a quantity to be “conserved” means that its quantity remains fixed over time, and is usually expressed as "$\Delta something = 0$." What is particularly interesting here is that the energy changes form or is reallocated into different groupings, but the total remains unchanged. The interpretation of the above equation is that the total energy of a system with no work done on it by forces acting from outside (sometimes referred to as a *closed system*) is conserved, i.e. the change in its total energy is zero. While all we have done is to rename quantities we have already discussed, this process leads to some useful sub-models for energy conservation that we can use to solve certain characteristic problems. The table below shows the various models we will use.

Figure 3.4.1 – Summary of Energy Conservation Models

outside		*inside*	*conserved*	*typical situation*
W	$=$	$\Delta KE + \Delta PE + \Delta E_{thermal}$	N/A	$\langle choice\ of\ system \rangle$
0	$=$	$\Delta KE + \Delta PE + \Delta E_{thermal}$	$\Delta E\,(total\ energy)$	$\langle choice\ of\ system \rangle$
0	$=$	$\Delta KE + \Delta PE + \cancel{\Delta E_{thermal}}$	$\Delta ME\,(mechanical\ energy)$	$no\ friction\ or\ drag$
0	$=$	$\Delta KE + \cancel{\Delta PE} + \Delta E_{thermal}$	$\Delta E\,(total\ energy)$	$inelastic\ collision$
0	$=$	$\Delta KE + \cancel{\Delta PE} + \cancel{\Delta E_{thermal}}$	$\Delta KE\,(kinetic\ energy)$	$elastic\ collision$

$$\Delta PE = -(work\ done\ by\ conservative\ forces\ between\ objects\ within\ the\ system)$$
$$\Delta E_{thermal} = -(work\ done\ by\ non-conservative\ forces\ between\ objects\ within\ the\ system)$$

1. The first entry allows for outside work to be done, which means that the system is not closed, and energy can enter or exit it via work done by force interactions with objects outside the system. We need to use this model when the statement of the problem does not let us isolate a closed system. In general we try to avoid it whenever possible, so that we don't have to perform any line integrals!
2. The second entry assumes a closed system, which includes both conservative and non-conservative forces. The energy within this system can then change forms, but the total never changes.
3. The third entry is a special case where there are no non-conservative forces within the closed system (or rather, the work done by the non-conservative forces is negligible). This is the model that features "frictionless surfaces" and "no air resistance."
4. The fourth and fifth entries we will use a bit later. Suppose there are two objects in a closed system, which can only interact with each other, but the force has a limited range. When they are far apart, they don't interact. When they get close together, they push off one another, in what can be best described as a "collision," after which they are once again far enough apart that they don't interact. If one of the forces between the two objects is non-conservative, then some kinetic energy is converted to thermal, and even after the objects separate,

this energy does not come back to kinetic form. In the case of conservative force interactions between the objects, however, the potential energy only depends upon the positions of the objects, and they end up in the same state after the collision as they were in before the collision (i.e. too far apart to interact), so for collisions we never have to take into account changes of potential energy. If the only forces present in a collision are conservative, then no thermal energy is produced, and the kinetic energy is conserved.

Now perhaps it is clear where "conservative" and "non-conservative" forces got their names – forces that conserve mechanical energy are conservative, and those that don't are non-conservative.

Potential Energy Function for Gravity

It's all well-and-good to introduce all these models, but unless we have shortcuts for plugging into the conservation equations, we still have to perform work line integrals. The shortcuts we seek come from the presence of the potential energy changes. It is clearly much easier to compute the difference of two numbers as we would in the case of ΔPE than it is to compute the work integral from scratch each time. So we will benefit from taking a few minutes to determine the potential energy functions for various conservative forces ahead of time, which we will then be able to use over and over. We start, as always, with gravity. In Equation 3.2.1, we computed the work done by the gravity force on a projectile, and found:

$$W_{grav}\left(A\rightarrow B\right)=-mg\Delta y \tag{3.4.9}$$

Clearly mg remains constant, so relating this to the negative change of gravitational potential energy immediately gives:

$$\Delta PE_{grav}=mg\Delta y \tag{3.4.10}$$

For reasons we will see later, it is helpful to define the potential energy due to gravity as a function of y. When we get to this point, it's best to drop the two-letter "PE" in favor of a single-letter function. It is traditional to use U. We therefore define:

$$U\left(y\right)=mgy+U_o, \tag{3.4.11}$$

where U_o is an arbitrary constant. This constant comes about because if we take a difference of potential energies at two different altitudes, the constant drops out:

$$\Delta U_{grav}=U\left(y_B\right)-U\left(y_A\right)=\left(mgy_B+U_o\right)-\left(mgy_A+U_o\right)=mg\left(y_B-y_A\right)=mg\Delta y \tag{3.4.12}$$

What this means is that we can reference the zero potential energy to be anywhere, and things will still always work out.

> **Alert**
>
> *This last point bears repeating and expanding. Potential energy only has meaning inasmuch as it* ***changes*** *from one place to another. Remember, it came from work done by a conservative force, and defining work at a single point makes no sense. We define a potential energy function in terms of a number of joules at every point in space, but we can – without changing any physics – add or subtract a fixed number of joules from every position. in other words, we can call any position at all the point of "zero potential energy." Of course, once a position of zero potential energy is selected, the potential energy at every other point in space is fixed (we can't arbitrarily add or subtract a different constant value everywhere, or the $\delta U's$ between points would also change.*

Let's try this out for – you guessed it – the simplest example imaginable. Throw a rock straight up at some initial speed, starting at a height y_o. How fast is it moving after it has risen to a height y_f? We are ignoring air resistance here, so we will use the mechanical energy conservation model, which only involves kinetic and potential energy. First, write down expressions for the kinetic and potential energy at the start and at the finish, then plug them into the conservation equation:

$$\left.\begin{array}{ll} initial: KE_i=\frac{1}{2}mv_i^2 & U_{grav}\left(y_i\right)=mgy_i+U_o \\ final: \;\; KE_f=\frac{1}{2}mv_f^2 & U_{grav}\left(y_f\right)=mgy_f+U_o \end{array}\right\} \;\Rightarrow\; 0=\Delta KE+\Delta U_{grav}=\frac{1}{2}mv_f^2-\frac{1}{2}mv_i^2+mgy_f-mgy_i \tag{3.4.13}$$

$$\Rightarrow\quad v_f=\sqrt{v_i^2-2g\left(y_f-y_i\right)}$$

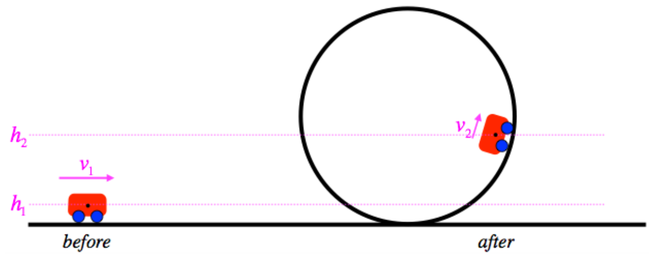
Sure enough, this result matches what we get using kinematics from Chapter 1. But notice that if the object had an x-component of velocity, then the velocity-squared in the kinetic energy includes another piece:

$$\Delta KE=\frac{1}{2}mv_B^2-\frac{1}{2}mv_A^2=\frac{1}{2}m\left(v_{xB}^2+v_{yB}^2\right)-\frac{1}{2}m\left(v_{xA}^2+v_{yA}^2\right) \tag{3.4.14}$$

But of course the x-component of velocity never changes, so we end up with the same change in KE as before. Notice that we don't have to worry about vertical and horizontal components in this example, because we have confined ourselves to only looking at changes in speed, not direction.

Great, so our mechanical energy conservation model works for projectiles, but we already knew how to solve those. Now let's look at something that is significantly harder to solve with Newton's second law and kinematics.

Figure 3.4.2 – Loop-de-Loop



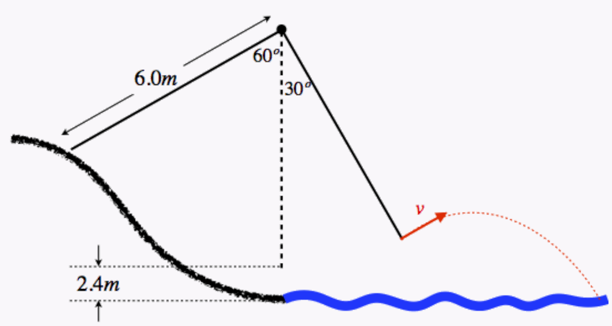
Our system here is the car (and the earth, so that work done by gravity is within the system), but the frictionless track is not included. This means that the normal force on the car by the track is a force coming from outside the system. But this doesn't mean we can't use a mechanical energy conservation model, because the normal force is at all times perpendicular to the motion. So while there is an outside force on the system, it doesn't do any work, which is what matters in our calculation. The car will change height and speed like a projectile, so we have precisely the same result as with the projectile, since it doesn't matter how the car got from point A to point B – rolling on a track or flying through the air, and we get exactly the same result as above:

$$v_2 = \sqrt{v_1^2 - 2g\left(h_2 - h_1\right)} \tag{3.4.15}$$

This required significantly less work than is involved using Newton's second law, and that is the point of constructing these energy conservation models. So long as we only want to know about the speed of the car and not its direction of motion or the time it takes to make its journey, this model gets us quickly to an answer. As an example of this lost information, change the final position of the car to the same height on the other side of the loop (on its way down after going around). The direction of the car's velocity is clearly different from the case above, and the time it took to get there is also different. All we can determine from this model is that the speeds are the same.

ExAmple 3.4.1

There are few things as fun as swinging into a river from a rope swing tied to the limb of a tree on its banks. Suppose you are swinging on such a rope which has a length of 6.0m, and which hangs straight down to the shoreline with its open end 2.4m above the water level. You start your swing at rest from dry land with the rope at a 60º angle with the vertical, and release the rope at an angle of 30º with the vertical (over the water, obviously).



a. *Find your speed at the point when you release the rope.*
b. *Find the distance above the water that you reach at the peak of your flight.*

Solution

a. The tension of the rope does no work here, as it acts perpendicular to your motion throughout. Ignoring air resistance, we therefore can use mechanical energy conservation to find the speed at the point of release. Let us will use as the zero potential energy reference point the tree branch. The distances below this reference point are easy to compute, as the length of the rope L is the hypotenuse for both triangles:

$$\left.\begin{array}{l} starting\ height = y_o = -L\cos 60^o = -\frac{1}{2}L \\ ending\ height = y_f = -L\cos 30^o = -\frac{\sqrt{3}}{2}L \end{array}\right\} \quad \Rightarrow \quad \Delta U_{grav} = mg\left(y_f - y_o\right) = mgL\left(\frac{1-\sqrt{3}}{2}\right)$$

Putting this into mechanical energy conservation with a starting speed of zero gives us the final speed:

$$0 = \Delta KE + \Delta U_{grav} = \frac{1}{2}mv_f^2 - \frac{1}{2}mv_o^2 + mgL\left(\frac{1-\sqrt{3}}{2}\right) \quad \Rightarrow \quad v_f = \sqrt{gL\left(\sqrt{3}-1\right)} = \boxed{6.6\frac{m}{s}}$$

b. We know that at the point of release, the velocity vector makes a 30° angle with the horizontal, so the horizontal component of this velocity (which never changes) is:

$$v_x = v\cos 30^o = 5.6\frac{m}{s}$$

When you hit your peak height, you will have a zero y-component of velocity, so the quantity above will be your speed. You still have the same mechanical energy at this point as you had at the beginning, so we can compute how far below the tree limb you are at this point using mechanical energy conservation from the very beginning to the point where you reach the peak:

$$0 = \Delta KE + \Delta U_{grav} = \frac{1}{2}mv_f^2 - \frac{1}{2}mv_o^2 + mgy_{peak} - mgy_o \quad\Rightarrow\quad y_{peak} = y_o - \frac{v_f^2}{2g} = -4.7m$$

This is measured from our $y=0$ reference point at the tree branch, so since we know how high the tree branch is above the water, have:

$$h = 6.0m + 2.4m - 4.7m = \boxed{3.7m} \tag{3.4.16}$$

Potential Energy Function for Elastic Force

For the mass on the spring we found:

$$W_{spring}\left(A\rightarrow B\right) = -\frac{1}{2}k\Delta\left(x^2\right) \tag{3.4.17}$$

If we go back to the calculation of the work done by a spring, we note the variable x is defined such that when it equals zero, the spring is in its equilibrium position.

As with the case of gravity, this immediately implies a function for elastic potential energy:

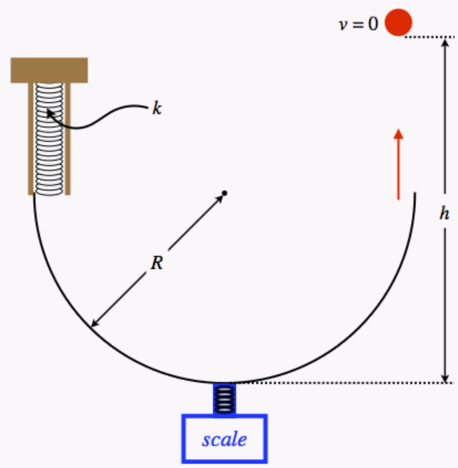
$$\Delta PE_{elastic} = -W_{spring}\left(A\rightarrow B\right) \quad\Rightarrow\quad U_{elastic}\left(x\right) = \frac{1}{2}kx^2 + U_o \tag{3.4.18}$$

If we go back to the calculation of the work done by a spring, we note the variable x is defined such that when it equals zero, the spring is in its equilibrium position. But suppose we happen to define our coordinate system differently, with the spring in equilibrium at a position we'll call x_o? In this case we only need to make the substitution $x\rightarrow x - x_o$, giving:

$$U_{elastic}\left(x\right) = \frac{1}{2}k\left(x - x_o\right)^2 + U_o \tag{3.4.19}$$

Example 3.4.2

A ball is launched straight up into the air with the apparatus shown below. The ball is pushed upward so that it compresses the spring, and is released from rest. It then travels around a frictionless half-circle track, at the bottom of which is a scale that measures the contact force the ball exerts on the track at that point. The mass of the ball is $m=0.400kg$, the stiffness of the spring is $k=22.0\frac{N}{m}$, and the radius of the track is 1.60m. When the ball passes over the scale, it reads 18.5N.



a. *Find the speed of the ball as it passes the scale.*
b. *Find the height h reached by the ball.*
c. *Find the amount that the spring was compressed before the ball was released.*

Solution

a. Start with a force diagram of the ball at the scale (gravity force down, normal force up), then use the facts that the scale measures the normal force and the ball's acceleration is centripetal:

$$N-mg=ma_c=m\frac{v^2}{R} \quad\Rightarrow\quad v=\sqrt{R\left(\frac{N}{m}-g\right)}=7.64\frac{m}{s}$$

b. There is no friction force by the track, and the contact force it exerts is perpendicular to the motion, so it does no work, which means that mechanical energy is conserved. With the speed of the ball at the bottom, we can therefore compute the height it reaches, where it comes to rest:

$$0=\Delta KE+\Delta U_{grav}=\frac{1}{2}mv_f^2-\frac{1}{2}mv_o^2+mgh \quad\Rightarrow\quad h=\frac{v_o^2}{2g}=2.98m$$

c. We know the total energy in the system, either from the KE at the scale, or the PE at the peak height:

$$E_{tot}=mgh=11.7J$$

When the ball compressed the spring, it had no KE, so all of this energy was stored in the PE of gravity and the elastic PE of the spring. Calling the compression Δy, then the height of the ball at the start is $R+\Delta y$. Summing the two PE's and setting the sum equal to the total energy gives a quadratic equation, which we then solve for Δy:

$$E_{tot}=mg(R+\Delta y)+\frac{1}{2}k(\Delta y)^2 \quad\Rightarrow\quad \Delta y=\frac{-mg\pm\sqrt{(mg)^2-2k(mgR-E_{tot})}}{k}=\boxed{0.55m}$$

LibreTexts

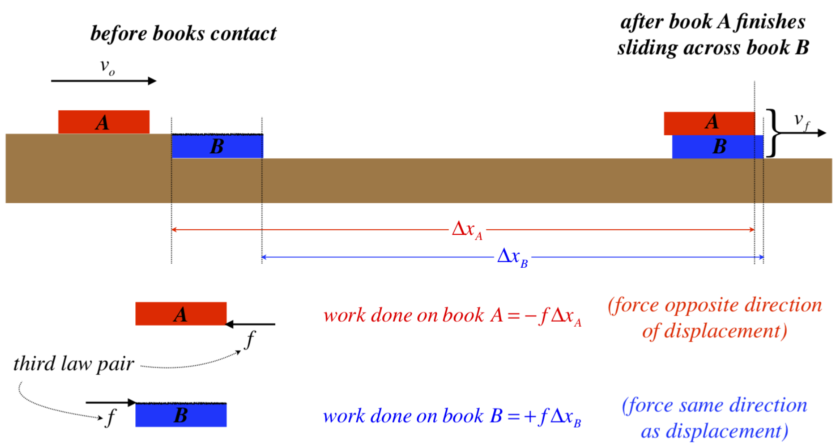
3.5: Thermal Energy

Work Between Colliding Objects

In the previous section we introduced the term *thermal energy*. We used this phrase as a catch-all to describe the form that energy takes when non-conservative forces internal to the system do work. It was not clear at that time why we had to introduce this element to our model, so let's examine it closer here.

When we first introduced the idea of a system, it was mentioned that while internal forces all come in third law pairs, the work done by these pairs are not generally equal-and-opposite. Applying this fact to a non-conservative force like friction is particularly enlightening. Consider the following physical situation: A book slides off a frictionless horizontal ledge directly onto another book which is at rest on a frictionless horizontal surface. There is a kinetic friction force between the two books, and they rub against each other until they are both moving along together at the same speed (see Figure 3.5.1 below).

Figure 3.5.1 – No "Equal-and-Opposite" Law for Work



Newton's third law assures that both books experience equal friction forces in opposite directions, but the work done by the friction force on book A is greater (in magnitude) than the work done on book B, because the displacement is greater. The friction force opposes the displacement of book A, so negative work is done on that book, while positive work is done on book B. Remember that we are interested in the overall effect of internal forces on the *system*, which in this case is the two books. We see that in fact in this case there is a net negative work done on the system:

$$work\ done\ on\ system = -f\Delta x_A + f\Delta x_B = -f\cdot(\Delta x_A - \Delta x_B) = -f \cdot (distance\ top\ book\ slides\ across\ bottom\ book) \tag{3.5.1}$$

This is an isolated system, which means that the negative work calculated above is not coming from outside. Our conservation models require that the total energy of an isolated system doesn't go up or down, though it can change form. The kinetic energy of this system drops as a result of the books coming in contact (this may not be clear at this point, but we will see this is true in the next chapter, so for now let's just accept it as true), and the energy can't go into potential energy, since the kinetic friction force is non-conservative. We put this energy conversion into the third type of energy in our model – thermal energy – and it is precisely equal to the (negative of the) work computed from the friction force acting through the *rubbing distance*.

It should also be noted that this analysis is not exclusive to kinetic friction. We would get the same result if the two books compressed a spring, provided we take as the "after" time the moment when the spring is completely compressed. In that case, the magnitude of the work done by the spring force on one book is not equal to the work done by the spring force on the other, but the difference in work (which is again negative) is stored in the potential energy of the spring.

Microscopic Model

While we will be staying with our macroscopic model of interactions, it helps with our understanding of thermal energy to consider a model of what is happening on smaller scales. We know that a macroscopic amount of matter is comprised of trillions upon trillions of atoms. If this collection of matter is a gas, then the particles don't interact very much with each other, while they do interact (with electromagnetic forces) if the particles make up a liquid or solid. Regardless of the forces acting between them, the particles are able to move, which means that they possess kinetic energy. If they interact electromagnetically, then the presence of this conservative fundamental force means that the particles also possess some potential energy.

So what happens when a system gains thermal energy, like when kinetic friction does net-negative work on the system, as in our example above? This energy is conveyed from the macroscopic to the microscopic modes of kinetic and potential energy. In other words, the energy doesn't mysteriously disappear from the universe, it just disappears from our macroscopic model, and can only truly be taken into account with a microscopic model. But there is something very subtle and fundamental that is going on at the same time.

If the energy that was macroscopically-mechanical (e.g. the kinetic energy of book A before it reaches book B in the example above) simply changes into energy that is microscopically-mechanical (kinetic and potential energy of the atoms in both books), then why refer to thermal energy as being fundamentally different from mechanical energy in our macroscopic model? The reason has to do with the one-way nature of the energy transfer into thermal. The kinetic and potential energies of the atoms are *randomly-distributed and randomly-oriented*. With the atoms vibrating in random directions (they vibrate because molecular bonds create restoring forces similar to tiny springs), all completely out-of-sync, then to give the kinetic energy back to the macroscopic realm, these trillions of particles would have to somehow coordinate their motions. The probability of the random motions of all the atoms being in the same direction at the same time – which is what needs to happen for the macroscopic object to move faster and get back its kinetic energy – is vanishingly small.

So for our macroscopic model, we separate the kinetic and potential energies that apply to large objects (mechanical energy) from those that apply to their randomly-moving atomic constituents (thermal energy). Both are energy, so both measure the same physical property, but conversions of mechanical energy into thermal energy are a one-way street – it's easy to disorder mechanical energy into thermal energy, but the reverse is too improbable to even consider as a possibility.

What we have been calling "non-conservative forces" are simply forces that can't help but bridge the macroscopic model to the microscopic. It's not a coincidence that when we first discussed friction, we took a brief detour into a microscopic model of irregularities of surfaces. Similarly, our first discussion of air drag included a diversion into a microscopic model of tiny particles bouncing off the affected object. The conservative force of gravity, on the other hand, acts on every particle in a macroscopic object at the same time, introducing no random differences between the fates of individual particles.

Energy Transfer

As we head back into our macroscopic model, it is useful to revise our notion of what work is. Up to now, we have thought about it as a way of isolating Newton's second law to changes in speed, in terms of the work-energy theorem. But an even more useful way to think about work is as *a means by which energy is transferred*. This transfer comes in one of three forms:

1. Energy is transferred into or out of a system (through an external force)
2. Energy is transferred from one mechanical form (KE or PE) into another mechanical form (through an internal conservative force).
3. Energy is transferred from a mechanical form into thermal (through an internal non-conservative force).

When it comes to solving problems that involve calculating the energy transferred, in case #1 we generally have no choice but to calculate the work done directly, though a line integral. In case #2, we have the potential energy functions already cataloged, and we don't have to deal with a work calculation at all. But in case #3, there are multiple approaches that might arise. One might be a direct calculation (use the friction force and displacement to find the work done). Another might be simply to find what is "left over" – if the system is closed and the mechanical energy is not conserved, the remaining energy transferred must have gone into thermal energy.

Heat

It turns out that the three cases above do not exhaust the possible ways that energy can be transferred. Suppose we put our system "in contact with" (whatever that means) an external system that exerts no macroscopic force, so that no external work is done between the systems. It is nevertheless possible to transfer energy between the systems *microscopically*. That is, thermal energy can be transferred into or out of a system without any work being done through a direct interaction of the two systems' microscopic energy modes. This form of energy transfer is called *heat*.

Alert

A common misconception is that heat and thermal energy are the same thing. Heat is like work – it is a means for transferring energy, not the energy itself (though it has units of energy). That is, heat is not contained within a system like thermal energy is. If this is confusing, think of it this way: If one system transfers energy to another system mechanically, it doesn't "lose work" while the other system "gains work" – work is not stored within the systems. We know this because work comes from force, which is an interaction between systems. Heat is exactly analogous to work, in that it is not contained in systems, it is an accounting of the energy exchanged.

Since we now know a system can receive energy from outside by a mode other than external work, we need to include this mode in our energy model equation. It is standard to use the variable Q to represent heat, and this value is considered to be positive when energy is added to a system. Thus we have:

$$Q+W_{ext}=\Delta KE+\Delta PE_1+\Delta PE_2+\dots+\Delta E_{thermal} \tag{3.5.2}$$

This equation simply states that the amount of energy transferred into or out of the system microscopically (heat), plus the amount of energy transferred into or out of the system macroscopically (work) equals the amount that the energy within the system changes. This expression of overall energy conservation is known as the *first law of thermodynamics* (although usually when one studies the subject of thermodynamics, the mechanical energy terms are ignored).

Also note that heat can be transferred between objects *within* a system, and in this case the energy transfer is between $\Delta E_{thermal}$'s for different objects within the system. Within the system we therefore have three modes of energy transfer from one type to another: (1) Work done by conservative forces converting between KE and PE, (2) Work done by non-conservative forces converting ME into thermal energy, (3) Heat transferred from one object to another raising the latter's thermal energy while reducing the former's.

We will not delve into the details of exactly *how* heat transfer occurs in this course, except to say that what enables it is a *temperature difference* between the two systems. When two systems at different temperatures exchange heat energy, it flows from the region of higher temperature to that of lower temperature, and this continues until both regions reach the same temperature (a condition that is referred to as *thermal equilibrium*).

Temperature Changes

Up to now, we have carefully avoided an aspect of thermal energy that arises naturally when the word "thermal" comes up. This form of energy must have *something* to do with temperature. Certainly it should come as no surprise that when a system's thermal energy rises, it gets hotter. Indeed, everyone has had experience with work done by kinetic friction causing the temperature of the objects to rise, such as rubbing one's hands together on a cold day.

What makes this interesting is that we can measure temperature changes, and relate these changes back to the amount of energy we know has been transferred. Rather than leap back into the realm of the microscopic to deal with this topic theoretically (we'll leave this to the aforementioned future course), we'll take an experimental approach to this. Suppose we move some energy into an object by the means of heat transfer. This energy can be moved from outside the system or within it. The question is, how is the amount of energy transferred into or out of an object related to that object's temperature change?

- We find that if we move twice as much energy into the object, its temperature rises twice as much. This tells us that the temperature change is proportional to the energy transferred:

$$Q \sim \Delta T \tag{3.5.3}$$

> Alert
>
> *Note that the heat transferred is negative if the object drops in temperature ($\Delta T < 0$), so we define heat as positive when it transfers into an object, and negative when it transfers out.*

Note that the heat transferred is negative if the object drops in temperature ($\Delta T < 0$), so we define heat as *positive when it transfers into an object, and negative when it transfers out*.

- We find that if we double object's mass and repeat the experiment, then the temperature only changes half as much as before. This seems to indicate that how much the the transferred energy is spread out is important. Or put another way, temperature seems to measure the average amount of energy added (or taken away) *per atom*. If the same amount of energy is spread out across more atoms, then the temperature change is less. So we can now include the mass in our proportionality:

$$Q \sim m\Delta T \tag{3.5.4}$$

- We find that if we change the type of material out of which the object is composed, then even with the same mass and same amount of energy transferred, the temperature change can be different. Characterizing the type of material with a constant c (called the *specific heat capacity* – a terrible name, actually), we finally get a full equation relating the energy transferred to the change in temperature:

$$Q = mc\Delta T \tag{3.5.5}$$

The units of this constant involves measurement of temperature. In the SI units we are working with, temperature is measured in degrees centigrade (^oC), so:

$$[c] = \frac{J}{kg^oC}$$

A very common alternative system of units used for energy in the context of specific heat capacity is *calories*. This is a convenient system because it is defined in terms of the temperature change of a specified amount of liquid water. To be precise, one calorie of energy is the amount that the thermal energy of one gram of liquid water must be increased in order to raise its temperature from 14.5^oC to 15.5^oC. That is a far more precise definition than we will require. For our purposes, the added energy that raises the temperature of one gram of water (at any starting temperature) by one degree centigrade is an adequate definition of a calorie. Notice that using calories as our unit for energy gives us a very easy-to-remember value for the specific heat capacity of water – it is simply: $c_{liquid\ water} = 1\frac{cal}{g^oC}$.

We can, of course, convert between calories and joules, and it turns out that:

$$1cal = 4.184J \quad \Rightarrow \quad c_{liquid\ water} = 4184\frac{J}{kg^oC} \tag{3.5.6}$$

> Example 3.5.1
>
> *A 1.2kg block of lead at a temperature of 80^oC is placed within an insulated container containing 0.6kg of water at a temperature of 20^oC. The block warms the water as the water cools the lead, eventually bringing them both to a common temperature. Find this equilibrium temperature. The specific heat capacity of lead is $130\frac{J}{kg^oC}$.*
>
> **Solution**
>
> *When a system is "insulated," it means that it does not exchange energy via heat with the outside surroundings. This means that we can treat the lead + water as a single, closed system, which means that the total energy does not change. With the temperature of the lead falling, it is clearly losing thermal energy, as the increase in the water temperature means it is gaining thermal energy. This is a closed system in which there is clearly no external energy added or changes in mechanical energy, so we have:*

$$Q+W_{ext}=\Delta KE+\Delta PE+\Delta E_{thermal} \quad \Rightarrow \quad 0=\Delta E_{thermal}\left(lead\right)+\Delta E_{thermal}\left(water\right) \quad \Rightarrow \quad \Delta E_{thermal}\left(lead\right)=-\Delta E_{thermal}\left(water\right)$$

The changes in these thermal energies are proportional to their temperature changes, and when they come to thermal equilibrium they end up with the same temperature, so plugging in Equation 3.5.5 for each case:

$$\left.\begin{array}{l}\Delta E_{thermal}\left(lead\right)=Q\left(leaving\; lead\right)=m_Lc_L\Delta T=m_Lc_L\left(T-80^oC\right) \\ \Delta E_{thermal}\left(water\right)=Q\left(entering\; water\right)=m_Wc_W\Delta T=m_Wc_W\left(T-20^oC\right)\end{array}\right\} \quad \Rightarrow \quad T=\frac{m_Lc_L\left(80^o\right)+m_Wc_W\left(20^o\right)}{m_Lc_L+m_Wc_W}=\boxed{23.5^o}$$

Of course we know that work done by non-conservative forces like friction and air drag move energy from mechanical energy into the $\Delta E_{thermal}$ category. This is not quite the same as when heat transfer does this (this requires a temperature difference), but it turns out that the relationship of temperature change with energy transferred in this way is the same, though one has to take care to properly distribute the energy between the objects rubbing against each other (or the object moving through a dragging fluid and the fluid itself).

Example 3.5.2

A $2600kg$ bronze statue is being moved across a marble floor at a speed of $1.5m/s$ on a pad that is effectively frictionless, when suddenly one of the movers accidentally steps on the pad and the statue slides off it, onto the slab of marble below, where it then slows to a stop due to kinetic friction. Assume the surfaces of the statue and floor exchange heat with each other as the slide occurs, so that the surfaces experience the same temperature change during the slide. Also assume that a very short time after the slide, the thermal energy has only distributed itself through a very thin layer of both surfaces, affecting only $2.0kg$ of the statue and $0.80kg$ of the floor. And finally, assume that during this short period, a negligible amount of heat is exchanged with the surroundings (i.e. the statue and marble slab may be treated as a closed system). Compute the temperature change of the affected thin layers of the statue and marble slab. The specific heat capacities of bronze and marble are $370\frac{J}{kg^oC}$ and $880\frac{J}{kg^oC}$, respectively.

Solution

The system starts with kinetic energy, which goes away, replaced with thermal energy. This happens by means of work done by kinetic friction, so this energy transfer can be treated in the same way as if it were transferred as heat. Using the same temperature change for both surfaces, we therefore have:

$$\frac{1}{2}m_{\substack{whole\\statue}}v^2=m_{\substack{statue\\surface}}c_{bronze}\Delta T+m_{\substack{slab\\surface}}c_{marble}\Delta T \quad \Rightarrow \quad \Delta T=\frac{\frac{1}{2}m_{\substack{whole\\statue}}v^2}{m_{\substack{statue\\surface}}c_{bronze}+m_{\substack{slab\\surface}}c_{marble}}=\boxed{2.0^oC}$$

An interesting follow-up question is, "If the coefficient of friction between the statue and the floor is smaller (so that the statue slides farther), or larger (so the statue stops sooner), how does the answer change?" Well, the amount of KE lost is the same in every case, so the answer won't change. The different coefficient of friction changes the friction force, but the work it does is the same, because the change of displacement balances the change of force.

Phase Changes

Thermal energy consists of the microscopic kinetic and potential energies of the microscopic particles comprising an object. These particles are held together by forces that can be overcome if sufficient energy is added, causing (for example) rigidly-held particles in a solid to slide over each other freely, changing the *phase* of the substance to liquid. Even more added energy can disassociate the particles from each other entirely, allowing them to move totally freely as a gas. In both of these transitions, the added energy goes *only* into the phase change – it doesn't change the temperature of the sample.

Of course, these phase transitions can only occur at certain temperatures (pressures also play a role, which we will ignore here), known as the *melting point* [0^oC for water] for the solid/liquid transition, and the *boiling point* [100^oC for water] for the liquid/gas transition. When there is a change in thermal energy (which, as above, typically occurs as a result of heat transfer, though work done by kinetic friction again has the same effect), and this change causes a change in phase, the amount of mass that changes phase is proportional to the amount of energy transferred. We can express it this way:

$$Q=\pm\, L\cdot\delta m, \tag{3.5.7}$$

where δm is the amount of mass that changes phase, and L is a constant that depends upon the specific substance and the type of phase change. The constant L is given another terrible name – it is called the *latent heat*, and obviously has units of J/kg. Because this quantity depends upon the type of phase transition, it is usually split into two types, one for each phase transition:

$$Q\left(melting/freezing\right)=\pm\, L_f\cdot\delta m, \qquad Q\left(boiling/condensing\right)=\pm\, L_v\cdot\delta m \tag{3.5.8}$$

The constant L_f is called the *latent heat of fusion*, and L_v is called the *latent heat of vaporization*. Note that the phase changes can go either way. The convention remains as before, that when heat enters the system, it is a positive value, and this will correspond to phase transitions from solid to liquid (melting) or liquid to gas (boiling). Negative values of heat correspond to energy leaving the system – phase transitions from liquid to solid (freezing), and gas to liquid (condensing).

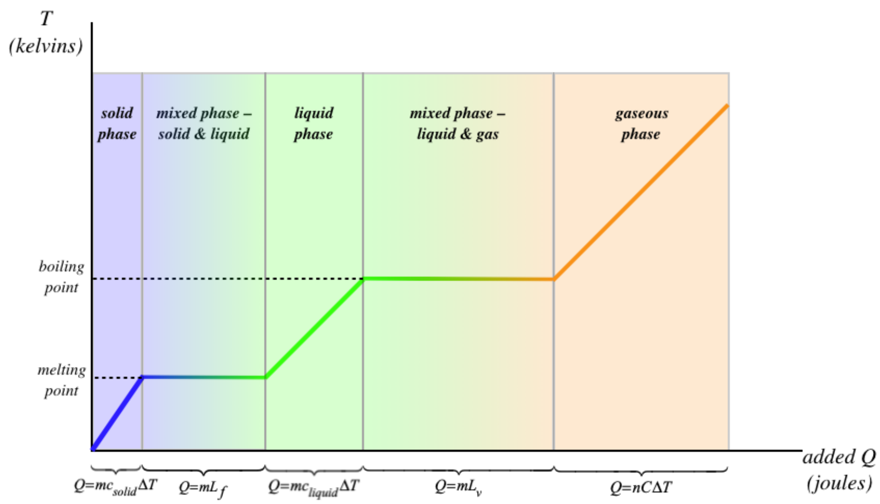
Suppose we wanted to transition a very cold solid all the way to a very hot gas. There would be several steps involved:

- add energy to raise the solid's temperature to the melting point

- add energy to change the phase from solid to liquid (while not changing the temperature)
- add energy to raise the liquid's temperature to the boiling point
- add energy to change the phase from liquid to gas (while not changing the temperature)
- add energy to raise the gas's temperature

Every one of these steps involves a different constant. The specific heat capacities of the solid, liquid, and gaseous phases of the same substance are not the same, and the latent heat of fusion is not the same as the latent heat of vaporization for the same substance. This process is depicted in the graph in Figure 3.5.2.

Figure 3.5.2 – Effect of Heat Transfer on Temperature and Phase



*[Note: In the final section of the graph, you'll note that there is an "n" in the equation relating heat and temperature change. This is because it is more common to measure the amount of a gas in a unit called "moles" rather than in units of mass. The heat capacity in this case is called "**molar** heat capacity," rather than **specific** heat capacity, as we are using for solid and liquid. This is all we will say about moles here, and will leave further discussion of this topic to a course in chemistry or a physics course in thermodynamics.]*

Example 3.5.3

A $0.80kg$ block of ice at a temperature of -6.0^oC is placed into a bucket containing $2.4kg$ of water that is at a temperature of 18^oC. The water and ice are insulated from their surroundings, so they can only exchange heat with each other. Describe the state of the system when it comes to thermal equilibrium. Specifically, is the final state all ice, all water, or a mixture of both, and what is the temperature of the system? The specific heat capacity of ice is $2100\frac{J}{kg^oC}$, and the latent heat of fusion for water is $330,000\frac{J}{kg}$.

Solution

These kinds of problems can be a bit tricky. We start by noting that energy is transferring (as heat) from the warmer water to the colder ice. As this happens, the water gets colder and the ice warmer. When one of them gets to the melting point, it will begin changing phase as the heat transfer continues, so we need to figure out which one reaches the melting point first. We do this by comparing the energy required to get each to the melting point:

$$Q_{ice}=m_{ice}c_{ice}\left(+6.0^o\right)=10,000J,\qquad Q_{water}=m_{water}c_{water}\left(-18^o\right)=-180,000J$$

So clearly the ice gets to the melting point first. When it does, the water is still warmer than the melting point, so the heat transfer continues. Our next step is to determine if all the ice melts before the temperature of the water gets to the melting point. The amount of energy needed to melt all the ice (after it reaches the melting point) is:

$$Q=L_f\Delta m=264,000J$$

But when the water loses $180,000J$ it will be at the melting point, which means the heat transfer will cease (both the ice and water are at the same temperature), so it cannot give the ice enough energy to melt it all. Therefore the final state is a mix of ice and water, but how much of each? To determine this, we simply give all of the energy the water can give to the ice before the water reaches 0^oC. Some of this energy warms the ice, and the rest melts some of it. The first $10,000J$warms the ice to the melting point, so the remaining $170,000J$does the melting:

$$\Delta m=\frac{Q}{L_f}=0.52kg$$

So this leaves a final state containing $0.28kg$of ice and $2.92kg$of water, both at a temperature of 0^oC.

3.6: Force and Potential Energy

Force and the Potential Energy Function

We now have an alternative to the using the work-energy theorem when conservative forces are involved – it consists of computing potential energies and applying mechanical energy conservation. In essence we have developed the idea of potential energy starting from force. Specifically, we have, from Equation 3.4.4 and the definition of work, the following relationship between the potential energy difference between two points and the conservative force that does the work for which the use of potential energy is a shortcut:

$$U_B - U_A = -\int_A^B \overrightarrow{F} \cdot \overrightarrow{dl} \tag{3.6.1}$$

As we saw in Section 3.4, we can express the potential energy of a system as a function of position, so the question arises, "Is there some way to "reverse" Equation 3.6.1 so that we can obtain the functional form of the conservative force from the potential energy function?" We know that derivatives are the "opposite" of integrals, so it should not be too surprising that the reverse of Equation 3.6.1 takes the form of a derivative. To see how this works, let's consider only a very tiny change in potential energy due to a very small displacement. This changes the left hand side of Equation 3.6.1 to an infinitesimal, and the right hand side is no longer a sum of many pieces, but is instead only a single piece:

$$dU = -\overrightarrow{F} \cdot \overrightarrow{dl} \tag{3.6.2}$$

In three dimensions, the tiny displacement can be written as:

$$\overrightarrow{dl} = dx\;\hat{i} + dy\;\hat{j} + dz\;\hat{k} \tag{3.6.3}$$

This means that the dot product with the force vector is:

$$\overrightarrow{F} \cdot \overrightarrow{dl} = F_x dx + F_y dy + F_z dz \tag{3.6.4}$$

Suppose we make our tiny displacement only along the x-axis, so that dy and dz are zero. Then clearly all the work done by the force is given by the first term above, and we get that the small change in potential energy that occurs when the position changes a small amount in the x-direction is:

$$dU\left(x \rightarrow x + dx\right) = -F_x dx \quad \Rightarrow \quad F_x = -\frac{dU}{dx} \tag{3.6.5}$$

This is fine for a potential that changes only in the x-direction, but what happens if the potential energy is also a function of y and z? The answer is that we *treat y and z as though they are constants*, which means that $dy = dz = 0$, and our result above works. When we treat y and z as constants, we have to do something slightly different with our derivative. For example, if we take a derivative of the function $U\left(x,y\right) = xy$ with respect to x, we get, from the product rule:

$$\frac{dU}{dx} = \frac{d}{dx}(xy) = (1)\,(y) + (x)\left(\frac{dy}{dx}\right) \tag{3.6.6}$$

But if we we treat y and z as constants, the derivative of these variables are zero, making the second term above vanish. We call this "hold the other variables constant" derivative a *partial derivative*, and we even use a slightly different symbol to represent it:

$$partial\ derivative\ of\ function\ f\ with\ respect\ to\ x = \frac{\partial f}{\partial x}$$

So following the discussion above, we find that by holding two of the variables constant at a time (so that the displacement for the work is along only one axis), we can obtain all the components of the force from the potential function $U\left(x,y,z\right)$:

$$F_x = -\frac{\partial}{\partial x}U, \quad F_y = -\frac{\partial}{\partial y}U, \quad F_z = -\frac{\partial}{\partial z}U \tag{3.6.7}$$

Example 3.6.1

An object with a mass of 2.00kg moves through a region of space where it experiences only a conservative force whose potential energy function is given by:

$$U(x,y,z)=\beta x\left(y^2+z^2\right), \qquad \beta=-3.80\frac{J}{m^3}$$

Find the magnitude of the acceleration of the object when it reaches the position $(x,y,z)=(1.50m,3.00m,4.00m)$

Solution

We know the mass of the object, so if we can determine the net force on it, we can get its acceleration from Newton's second law. The only force on this object is the conservative force with the given potential energy function, so that is the net force. We compute its components using the partial derivatives:

$$F_x=-\frac{\partial}{\partial x}\left[\beta x\left(y^2+z^2\right)\right]=-\beta x\left(y^2+z^2\right)=95.0N$$
$$F_y=-\frac{\partial}{\partial y}\left[\beta x\left(y^2+z^2\right)\right]=-2\beta xy=34.2N$$
$$F_z=-\frac{\partial}{\partial z}\left[\beta x\left(y^2+z^2\right)\right]=-2\beta xz=45.6N$$

And now for the magnitude of the acceleration:

$$a=\frac{\left|\overrightarrow{F}\right|}{m}=\frac{\sqrt{F_x^2+F_y^2+F_z^2}}{m}=\boxed{55.4\frac{m}{s^2}}$$

We can check to make sure that this method of deriving the force from the potential energy is consistent with the cases we have seen already:

Gravity: $U(x,y,z)=mgy+U_o$

$$\left.\begin{array}{l} F_x=-\frac{\partial}{\partial x}U=-\frac{\partial}{\partial x}(mgy+U_o)=0 \\ F_y=-\frac{\partial}{\partial y}U=-\frac{\partial}{\partial y}(mgy+U_o)=-mg \\ F_z=-\frac{\partial}{\partial z}U=-\frac{\partial}{\partial z}(mgy+U_o)=0 \end{array}\right\} \Rightarrow \quad \overrightarrow{F}_{gravity}=-mg\,\hat{j} \tag{3.6.8}$$

Elastic Force: $U(x,y,z)=\frac{1}{2}kx^2+U_o$

$$\left.\begin{array}{l} F_x=-\frac{\partial}{\partial x}U=-\frac{\partial}{\partial x}\left(\frac{1}{2}kx^2+U_o\right)=-kx \\ F_y=-\frac{\partial}{\partial y}U=-\frac{\partial}{\partial y}\left(\frac{1}{2}kx^2+U_o\right)=0 \\ F_z=-\frac{\partial}{\partial z}U=-\frac{\partial}{\partial z}\left(\frac{1}{2}kx^2+U_o\right)=0 \end{array}\right\} \Rightarrow \quad \overrightarrow{F}_{elastic}=-kx\,\hat{i} \tag{3.6.9}$$

Determining Conservative or Non-Conservative

We know that a potential energy can only be defined for a conservative force, and until now to show that a force is non-conservative we had to do two line integrals between the same two points and show that they yield different results, but this program for finding the force from the potential energy function gives us another less-onerous method for doing this. It goes something like this:

- Start with the force we want to know about, and integrate the x-component with respect to x to "undo" the negative partial derivative of the potential energy function with respect to x. Don't forget to leave an arbitrary constant added to the integration (this is an indefinite integral):

$$U(x,y,z) = -\int F_x dx + constant \tag{3.6.10}$$

- Because we have undone a partial derivative (which assumes the other variables are constant), even the variables y and z are fair game for the arbitrary constant of integration, so write the constant as an unknown function of those variables:

$$U(x,y,z) = -\int F_x dx + h(y,z) \tag{3.6.11}$$

This can readily be shown to be correct by taking the negative partial derivative with respect to x of both sides.

- Use this "candidate" potential energy function to get the other two components of the force vector. If this is possible, then the function $h(y,z)$ can be found (to within a numerical constant). If there is no way to get to the y and z components of the force vector, then it is non-conservative.

Example 3.6.2

Show that the force given in Example 3.2.1 (given again below) is not conservative, using the try-to-integrate-the-force method.

$$\overrightarrow{F}(y) = \alpha y \hat{i}$$

Solution

There is only an x-component of the force, so integrate that with respect to x:

$$U(x,y,z) = -\int F_x dx = -\alpha xy + h(y,z)$$

If we pick the function $h(y,z)$ equal to just zero, aren't we done? Haven't we shown that the force is conservative? After all, its derivative with respect to x gives us the x-component of the force, and that is the only component. Not so fast! The other components are zero, and we must be able to get those components from the partial derivatives as well. Here is where we run into trouble. Taking the partial derivative with respect to y and setting it equal to zero gives:

$$F_y = -\frac{\partial}{\partial y}U = -\frac{\partial}{\partial y}(-\alpha xy + h(y,z)) = \alpha x - \frac{\partial h}{\partial y}$$

This can only equal zero (and give the proper y component of the force) if $\frac{\partial h}{\partial y}$ equals αx. But how can this possibly be true, when the function h depends upon y and z? This is mathematically impossible, which means that this force is non-conservative.

While it's unlikely you have encountered it at this point unless you have taken more math courses than is typical at this point, you should be made aware of a shorthand notation that exists for this process of obtaining the force vector from the potential energy function. Rather than write three equations – one for each component of force – this relationship is often written as a vector equation that looks like this:

$$\overrightarrow{F} = -\overrightarrow{\nabla} U \tag{3.6.12}$$

The funny-looking triangle vector is called the *gradient operator*, or "del," and can be written like this:

$$\overrightarrow{\nabla} \equiv \hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}, \tag{3.6.13}$$

or, in column matrix notation:

$$\overrightarrow{\nabla} \equiv \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \tag{3.6.14}$$

Note that $\overrightarrow{\nabla}$ is not itself a vector – it has to "act upon" a function to create a vector. When it performs this function, the derivatives define vector components which are conveniently multiplied by the unit vectors.

Equipotential Surfaces

Consider the following potential energy function:

$$U(x,y,z) = -\alpha\left(x^2+y^2+z^2\right) \tag{3.6.15}$$

Notice that every point that is the same distance from the origin results in the same potential energy, since the potential energy function is proportional to the square of the radius of a sphere centered at the origin. This means that if an object moves between two points in space, where both points are the same distance from the origin, then (assuming this is the only force present) the object is moving the same speed at both points. This is because mechanical energy is conserved, and the potential energy hasn't changed, so the kinetic energy is also unchanged.

Every value available to the $U(x,y,z)$ above defines the surface of a sphere centered at the origin on which every point corresponds to the same potential energy. But the potential energy function above is not unique. Every such function defines surfaces of equal potential energy. We call these *equipotential surfaces*. A good example of these are represented by the dotted lines you see on topographical maps used by backpackers – each dotted line represents a fixed altitude, and therefore an equal gravitational potential.

Let's compute the force vector for the potential above:

$$\overrightarrow{F} = \hat{i}\left(-\frac{\partial U}{\partial x}\right)+\hat{j}\left(-\frac{\partial U}{\partial y}\right)+\hat{k}\left(-\frac{\partial U}{\partial z}\right) = 2\alpha\left(x\hat{i}+y\hat{j}+z\hat{k}\right) \tag{3.6.16}$$

Hopefully you recognize the part of this vector in parentheses. It is the position vector relative to the origin, Equation 1.6.1. This vector points directly to the point (x,y,z) from the origin, which means that it is *perpendicular to the sphere centered at the origin that contains that point*. It turns out to be a general property that *the conservative force associated with a potential is perpendicular to the equipotential surfaces everywhere in space*.

Notice that for the function $U(x,y,z)$ above, if $\alpha > 0$, the potential energy gets *smaller* as one gets farther from the origin, and the force vector from this potential points away from the origin. This is also a general feature – *the conservative force associated with a potential points in the direction from greater potential to lower potential*. It should be clear on many fronts why this must be the case. If an object moves from a region of higher potential to one of lower potential, this decrease in *PE* must be balanced by an increase in *KE*, which means the object speeds up. Objects speed up when the net force on them points in the same direction that they are moving, so the force must point from where the PE is higher to where it is lower.

3.7: Energy Diagrams

An *energy diagram* provides us a means to assess features of physical systems at a glance. We will examine a couple of simple examples, and then show how it can be used for more advanced cases in physics and chemistry. It's important to understand that there is no new physics in here – what we have learned so far now is simply represented diagrammatically, making it easier in some cases to see the "big picture" of a physical system.

Elements of Energy Diagrams

First of all, it should be noted that we will be confining ourselves to energy diagrams for 1-dimensional motion. This dimension will be represented by the horizontal axis, and the vertical axis will have units of energy. Secondly, the physical systems represented by energy diagrams will involve only one (conservative) force acting on an object.

Construction of an energy diagram entails first graphing the potential energy function for the conservative force on the axes. Note that potential energy function includes an arbitrary additive constant, which means that the entire graph can be moved up or down on the vertical axis as much as one likes without changing the physical system at all. There is one common convention that is followed regarding the height of the graph on the vertical axis, which we will see below, but it should be remembered that this is only convention, and doesn't change any of the physical properties of the system.

The graph of the potential energy function could apply to any object under the influence of this conservative force. To represent a specific system, the diagram also needs to indicate the total mechanical energy of the system, and this is done with a horizontal line with the correct height on the vertical axis.

That's all there is to drawing these diagrams. The real value comes from interpreting them, which we will discuss in the context of a couple of simple examples.

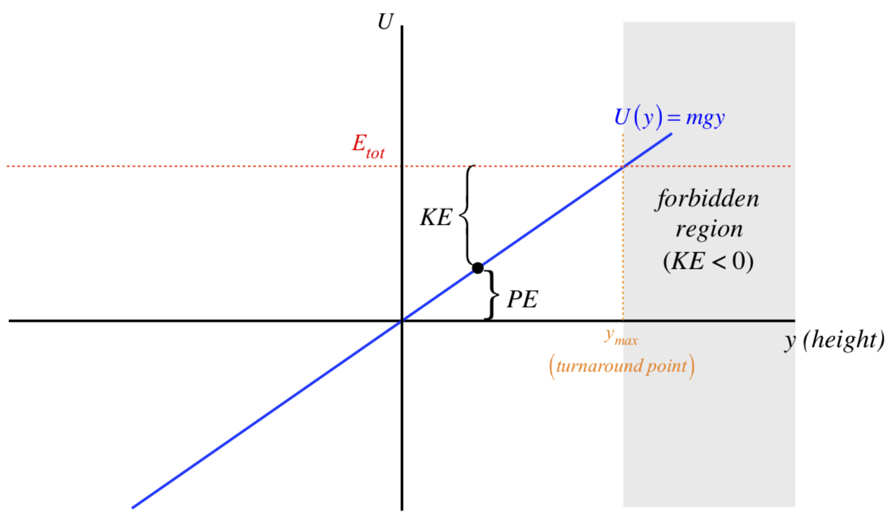
Two Simple Examples

Let's look at the energy diagrams for the two conservative forces we have dealt with so far... gravity and the elastic force.

Gravity

If we choose the arbitrary constant U_o for the gravitational potential energy to be zero, we have as a graph of the potential energy function a straight line that passes through the origin. We then include a horizontal line to represent the total energy of the particular system (which I will label as E_{tot}). Now for interpretation...

Figure 3.7.1 – Energy Diagram for Object Influenced by Gravity Near Earth's Surface



The position of the object (which in this case is the height above some defined zero point), is the value along the horizontal axis. For every position of the object, there is a corresponding value of its potential energy, given by the height of the $U(y)$ graph above (if positive) or below (if negative) the horizontal axis. The total (mechanical) energy of this system is conserved (i.e. it is the same for every position of the object), which explains why the total energy graph is a horizontal line. For a given position, the gap between the total energy line and the potential energy line equals the kinetic energy of the object, since the sum of this gap and the height of the potential energy graph is the total energy.

We can also interpret the intersection point of the total energy and the potential energy graphs. At this point, the total energy equals the potential energy, which means the object has no kinetic energy – i.e. the object is at rest at this position. How can an object under the influence of only gravity be at rest? It can be for just an instant, when it reaches the peak of its flight. Therefore the value on the horizontal axis corresponding to this intersection point is the highest elevation the object can reach. Note that for heights (horizontal axis values) greater than this, the potential energy is *greater* than the mechanical energy, which would require a negative kinetic energy. This is of course impossible, and we call this the *forbidden region* of the diagram, as we will never find the system in one of these states. As time passes, when the object reaches the intersection point, it must have done so from the allowed region, which means that when the object comes to rest here, is reverses its direction of motion. Consequently, this position is often referred to as the *turnaround point*.

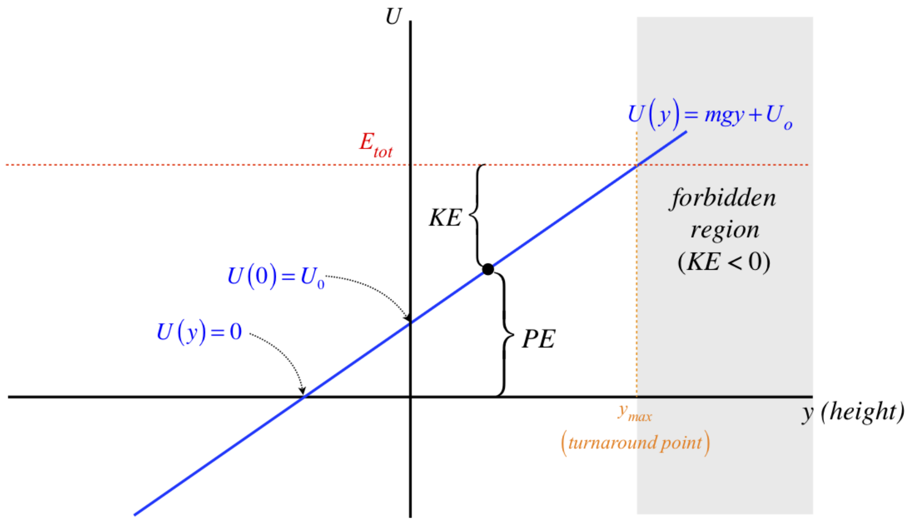
There is one other nugget of information we can extract from this diagram, though in this particular case it is fairly trivial. If we evaluate the negative of the slope of the potential energy graph at the point where the object is at some moment, we know the force acting on the object at that moment. In the case of gravity, the force is the same everywhere:

$$F_y = -\frac{dU}{dy} = -\frac{d}{dy}(mgy) = -mg \tag{3.7.1}$$

[*Note: There is no need for partial derivatives here, as we are only dealing with one-dimensional potential energy functions.*]

If we change the arbitrary constant, the only quantities that change in the entire picture are the potential energy and total energy. Every physically-observable quantity (kinetic energy, turnaround point, and force) remains unchanged. This may not be immediately apparent, but looking at the graph it is easy to see:

Figure 3.7.2 – Redefined Zero Point for Gravitational Potential Energy

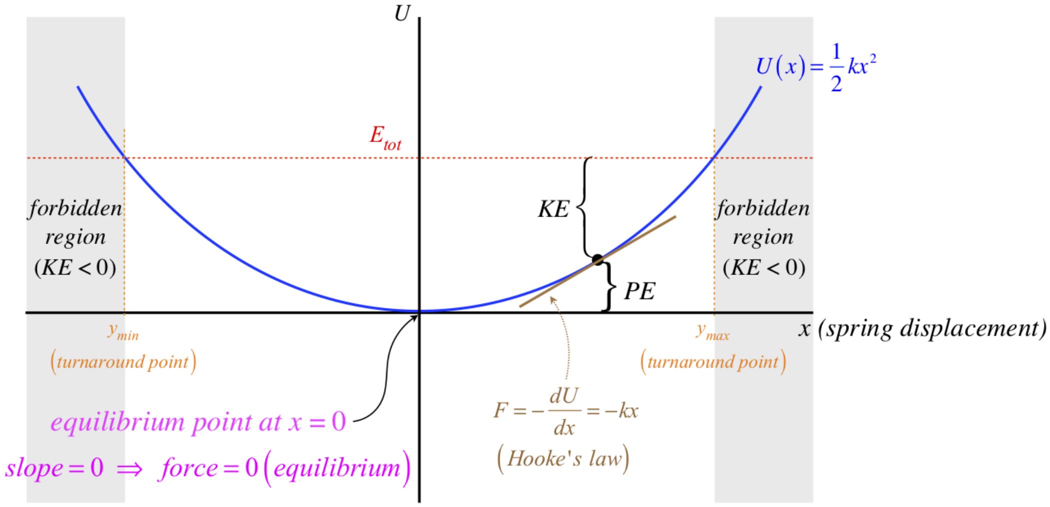


Interestingly, the fact that this potential is a straight line means that a shift of the graph up or down by an additive constant is equivalent to redefining the origin. This is easily seen by noting that this graph can also be viewed as the previous graph shifted to the left by y_o, where $mgy_o = U_o$. So for this simple case, changing the zero point of potential energy is equivalent to changing the position which we call the origin.

Elastic Force

We take precisely the same steps to draw the energy diagram for a mass on a spring, but there are some differences, such as two forbidden regions and a different slope for every position, and there is one additional feature for this potential that doesn't exist for the case of gravity: an *equilibrium point*.

Figure 3.7.3 – Energy Diagram for Object Influenced by Elastic Force



The two forbidden regions arise here because the spring has a maximum stretch and a maximum compression that result in potential energy equaling the total energy. Regions of potential energy confined by two turnaround points like this are often referred to as *potential wells*. Clearly the slope of the potential energy curve is different everywhere, which reflects the fact that the force by the spring is different for every position the mass can have.

An equilibrium point occurs whenever the slope vanishes (at maxima, minima, and inflection points in the potential energy curve) – there are simply places where the force vanishes. For the spring, it is the position where the spring is neither stretched nor compressed from its natural length. This particular equilibrium is referred to as a *stable equilibrium* for the following reason: If the object is at rest at this point, and it is given a small nudge in either direction, the resulting force acts to bring the object back to its original position. We can see this here, because the slope on the (+) side of the equilibrium point is positive, which means the force is in the negative direction. The force on the (–) side of the equilibrium point similarly acts back toward the equilibrium. Forces that create stable equilibrium like this are called *restoring forces*.

It should be clear that any minimum in the potential energy curve will lead to a stable equilibrium, but what about maxima? In this case, the forces that result from small displacements away from the equilibrium point act to push or pull the object farther from its starting point. This type of equilibrium is therefore referred to as an *unstable equilibrium*. Inflection points lie between parts of the curve that are concave (stable) on one side and convex (unstable) on the other, and the resulting equilibrium is referred to as a *meta-stable equilibrium*.

As with the case of gravity, shifting the entire curve up or down by an amount U_o doesn't change any of the physics, including the position of the equilibrium point. Unlike the gravity case, shifting the entire curve left or right (changing the definition of the origin) is not equivalent to the addition of an additive constant to the potential energy curve. But in both cases, the physics is unchanged by the positioning of the curve relative to either of the axes.

Bound States of Two Particles

While our models of terrestrial gravity and the elastic force are useful, we have to keep in mind that force interactions are between two objects, and the energy diagrams we have drawn appear to involve only a single object. The leap to discussing two objects is not a difficult one. Instead of being a position along the x or y axis, the one dimension of freedom becomes the *separation* of the two objects, which we represent with the variable r. If we can express the force as a function of the separation of the two objects interacting, then we can express the potential energy as a function of that variable as well, and voilà – we can draw an energy diagram. For the sake of interpreting such diagrams correctly, we have to keep in mind that the horizontal axis represents a separation, rather than a position, which leads to a big difference from the energy diagrams we created above – the horizontal axis has no negative values. It also should be remembered that these diagrams only relate motion between the particles along the line joining them. If we want to include motion *around* each other (as in the case of an orbit), we require more information than we can get from the energy diagram.

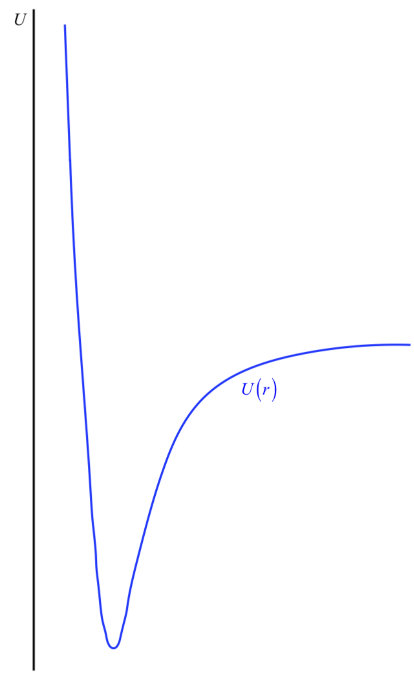
When we look at the universe in both the microscopic and macroscopic realm, we see countless examples of forces holding systems together. Gravity interactions between the sun and the planets keeps them in orbit. Electromagnetic interactions between protons and electrons holds them together in atoms, and electromagnetic interactions between atoms bind them into molecules. Systems of two bodies that possess too little total energy to escape each other's attractive force are in what is called a *bound state*. For the purposes of this section, we will confine ourselves to discussing bound states between microscopic particles, and save the discussion of orbits of celestial bodies for the chapter on gravitation.

Bound states between atoms and molecules in our universe are quite extraordinary. Here is how the famous Nobel laureate Richard Feynman put it:

> *If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generations of creatures, what statement would contain the most information in the fewest words? All things are made of atoms—little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another. In that one sentence ... there is an enormous amount of information about the world.*

From our earlier discussion, we know that this means that forces between atoms (and between molecules, which are clusters of atoms) must be *restoring* in nature, with an equilibrium separation. Indeed this does tell us a lot about the shape of the potential energy curve – it must contain a local minimum. But we know even more than this. Clearly a separation distance of zero is not possible (the particles can't occupy the same space), and this is assured by an ever-increasing force as they continue to get closer. This means that the potential energy curve gets steeper and approaches infinity as the separation distance gets closer to zero (i.e. the graph gets closer to the vertical axis). We know one other thing: The force between the two particles gets weaker as they get farther apart, and drops to zero in the limit as their separation goes to infinity. This means that the potential energy curve "flattens out" to a horizontal line as the distance from the vertical axis goes to infinity. Amazingly, this seemingly insignificant amount of information gives us a general shape for the interaction potential of two particles.

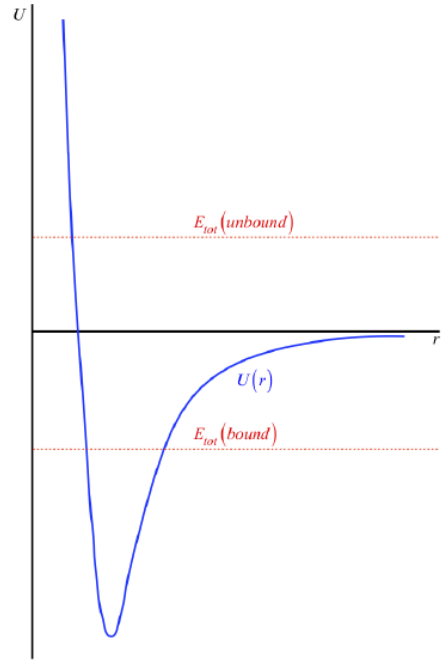
Figure 3.7.4 – Shape of Inter-Particle Interaction Potentials



The horizontal axis has been intentionally left out here, because this is not something we can determine. Recall that this potential energy curve can be raised or lowered an arbitrary amount without changing the physics, so we are actually free to place the horizontal axis (the point of zero energy) wherever we like. It is here that we come to the convention alluded-to at the start of this section: It is generally agreed to place the horizontal axis at the position that the potential energy curve reaches as $r \rightarrow \infty$. That is, it is usually agreed that the potential energy of the interaction vanishes when the particles are separated by an infinite distance (see Figure 3.7.5).

One nice consequence of the $U\left(r \rightarrow \infty\right) \rightarrow 0$ is that it gives us a simple rule-of-thumb to determine whether or not the two particles in this potential are bound to each other. If the total energy of the system is positive (i.e. the horizontal line representing the total energy is above the r-axis), then that means that when the two particles are moving away from each other, the graphs never intersect to give a "forbidden region," and they just keep moving apart – they are not bound. If the total energy is negative, then the total energy horizontal line intersects the potential energy graph in two places, giving two turnaround points, keeping the particles within a range of separations.

Figure 3.7.5 – Bound and Unbound States



All of the physical interpretations we came up with above apply to this function as well, though it might be a bit confusing at first, since for much of the graph the potential energy is negative. But note:

- The kinetic energy for a specific position is still always positive, and equals the gap between the point on the curve and the total energy line.
- The force between the particles is still the negative of the slope, which means it is *attractive (seeks to make r smaller) when the slope is positive, and is repulsive (seeks to make r bigger) when the slope is negative*.
- The equilibrium and turnaround points are defined as the bottom of the dip and the intersection points of the total energy line and potential energy curve, respectively.

Alert

*When two particles are bound to each other, to break the bond, energy must be added to the system. That is, the total energy line must be moved up until it is above the r-axis. Systems of particles must **receive additional energy** to break chemical bonds, and energy only comes out of chemical reactions in which new bonds are formed. For some reason, belief of the exact reverse is a very common misconception.*

A model of the potential energy that works very well for two neutrally-charged atoms or molecules was constructed in 1924 by John Lennard-Jones. This *Lennard-Jones potential* has a simple-but-surprising form (yes, those powers are not typos!):

$$U(r) = \epsilon\left[\left(\frac{r_o}{r}\right)^{12} - 2\left(\frac{r_o}{r}\right)^{6}\right] \tag{3.7.2}$$

Example 3.7.1

Show that in the Lennard-Jones potential, r_o is the separation of the particles when they are at equilibrium, and ϵ is the depth of the potential well (expressed as a positive value).

Solution

The minimum of this curve occurs at the point where the derivative vanishes, so:

$$0 = \frac{dU}{dr} = \epsilon\frac{d}{dr}\left[\left(\frac{r_o}{r}\right)^{12} - 2\left(\frac{r_o}{r}\right)^{6}\right] = 12\left(\frac{r_o}{r}\right)^{11} - 12\left(\frac{r_o}{r}\right)^{5} \quad \Rightarrow \quad r = r_o$$

The value of the potential at the minimum is $U(r = r_o)$, which one can see from a simple plug-in is in fact equal to $(-\epsilon)$.

Example 3.7.2

A system of two particles bound in a Lennard-Jones potential requires that an addition of energy equal to $\frac{3}{4}\epsilon$ in order to become unbound. Find the turnaround points of this bound system in terms of r_o.

Solution

The potential energy equals to the total energy at the turnaround points, and the total energy is given to be $-\frac{3}{4}\epsilon$ (i.e. the energy to unbind the particles is the amount needed to bring the total energy to zero), so:

$$-\frac{3}{4}\epsilon=\epsilon\left[\left(\frac{r_o}{r}\right)^{12}-2\left(\frac{r_o}{r}\right)^{6}\right] \quad\Rightarrow\quad -\frac{3}{4}=\left[x^2-2x\right], \qquad where\;\; x\equiv\left(\frac{r_o}{r}\right)^6$$

Now solve the quadratic for x:

$$4x^2-8x+3=0 \quad\Rightarrow\quad x=\frac{8\pm\sqrt{8^2-4\,(4)\,(3)}}{2\,(4)}=\frac{3}{2}\;\; or \;\;\frac{1}{2}$$

Now solve for the two r values:

$$\left(\frac{r_o}{r}\right)^6=x \quad\Rightarrow\quad r=x^{-\frac{1}{6}}r_o \quad\Rightarrow\quad r_{min}=\boxed{0.93r_o}, \quad r_{max}=\boxed{1.12r_o}$$

Modeling Bonds as Springs

It is quite common to model chemical bonds as springs, but this seems like a strange practice, given that the potential energy function looks like Equation 3.7.2, which is nothing like the potential energy function of a spring. Surprisingly, a spring can nevertheless act as a reasonable replacement for the Lennard-Jones potential when the total energy is deep in the well. Certainly the curve at the bottom of the well *resembles* the parabolic curve of the elastic potential, but one might argue that the bottom of *any* concave curve will resemble the elastic potential energy curve. It turns out that this argument is completely correct – *every smooth concave curve can be approximated by the parabolic curve of the elastic potential energy*!

How well this approximation works depends upon the range of values we confine ourselves to. If we look at the whole curve, then obviously the approximation of the Lennard-Jones potential with a parabola breaks down badly. But as we narrow-down our view to a smaller and smaller range near the bottom of the well, this approximation gets better.

The reason this works is related to some amazing mathematics: Any smooth function of a single variable can be written as a series (often infinite) of powers of that variable, with each term multiplied by a different constant:

$$f\left(x\right)=a_o+a_1x+a_2x^2+a_3x^3+\ldots \tag{3.7.3}$$

From this perspective, it's clear that *every* function has a bit of the x^2 function in it. For the Lennard-Jones potential, we can define x in the expansion above as $\left(1-\frac{r}{r_o}\right)$, which is a measure of how far the particle separation is from the equilibrium. For example, if the separation distance (r) is 90% of the equilibrium distance (r_o), then $\left(1-\frac{r}{r_o}\right)=0.1$, and for r equal to 99% of r_o, $\left(1-\frac{r}{r_o}\right)=0.01$.

In the spring model, it is precisely the distance that the spring is stretched or compressed from the equilibrium that determines the potential energy. So if we write the Lennard-Jones potential as an expansion in powers of $\left(1-\frac{r}{r_o}\right)$, we can more easily compare it to a spring potential energy:

$$\epsilon\left[\left(\frac{r_o}{r}\right)^{12}-2\left(\frac{r_o}{r}\right)^{6}\right]=U\left(r\right)=a_o+a_1\left(1-\frac{r}{r_o}\right)+a_2\left(1-\frac{r}{r_o}\right)^2+a_3\left(1-\frac{r}{r_o}\right)^3+\ldots \tag{3.7.4}$$

To use a spring to model this function, we want the series to look like a quadratic, so we need the contributions of the terms in the series with powers greater than 2 to be small compared to the terms before them. Well, if we restrict ourselves to values of r that are close to r_o (i.e. consider particles that are separated by a distance close to the equilibrium separation, which means the total energy is quite low), then the difference $\left(1-\dfrac{r}{r_o}\right)$ is a small number less than 1. The more times we multiply this small number by itself, the smaller the result, so higher powers provide ever-smaller contributions to the sum of the series.

All that remains if we are going to use the elastic potential energy to approximate the Lennard-Jones (or any other) potential is to figure out how to determine the equivalent "spring constant." Fortunately we have a nice trick for doing this. We can compare the way in which the spring constant enters the potential energy function to the series expansion, and determine what number we need from the series expansion to find the *effective spring constant*:

$$\frac{1}{2}k_{eff}x^2 = a_2x^2 \quad \Rightarrow \quad k_{eff} = 2a_2 \tag{3.7.5}$$

Okay, so we know where to find the effective spring constant – we express the actual potential as an infinite series, then look at the constant that multiplies the quadratic term in the expansion, and multiply it by two. But this still means we need to come up with the series... or does it? Consider the following useful mathematical trick: First, take two derivatives of Equation 3.7.3:

$$f(x) = a_o + a_1x + a_2x^2 + a_3x^3 + \ldots \quad \Rightarrow \quad \frac{df}{dx} = 0 + a_1 + 2a_2x + 3a_3x^2 + \ldots \quad \Rightarrow \quad \frac{d^2f}{dx^2} = 0 + 0 + 2a_2 + 6a_3x + \ldots \tag{3.7.6}$$

Now evaluate this second derivative at the point $x = 0$. All of the terms after the constant term then vanish, leaving what we are looking for:

$$\left.\frac{d^2f}{dx^2}\right|_{x=0} = 2a_2 = k_{eff} \tag{3.7.7}$$

It is left as an exercise for the reader to show that the effective spring constant for the Lennard-Jones potential deep within the well comes out to be: $72\dfrac{\epsilon}{r_o^2}$. [*Note that instead of taking two derivatives with respect to x and evaluating at $x = 0$, where $x \equiv 1-\dfrac{r}{r_o}$, we can perhaps more easily do the equivalent calculation of taking the derivatives with respect to r and evaluate at $r = r_o$.*]

CHAPTER OVERVIEW

4: Linear Momentum

Thumbnail: A pool break-off shot. Image used with permission (CC-SA-BY; No-w-ay).

4.1: Repackaging Newton's Second Law

Definition of Impulse

In chapter 2, we made a point of emphasizing that force is not possessed by objects – it is an interaction between them. One way we know this is that if the same force is exerted on identical objects that start at rest, the two objects are not necessarily moving the same afterward. There is an important ingredient missing here – the *duration* that the force acts. Since the force causes an acceleration, the longer it acts, the more the velocity is affected. So multiplying the force by the amount of time it acts may provide us with a useful quantity. The force may be changing magnitude or direction while it acts, but over a very short time this product is:

$$d\overrightarrow{J} = \overrightarrow{F}\,dt \tag{4.1.1}$$

If we want to know the totality of this quantity over a finite time interval, we need to add up all these little contributions. We give this quantity the name *impulse*.

Definition: Impulse

$$\overrightarrow{J}_{tot}\left(t_A \rightarrow t_B\right) \equiv \int\limits_{t_A}^{t_B} \overrightarrow{F}_{net}\,dt$$

This quantity is the sum of the product of the forces and the times over which those forces act. This certainly sounds very similar to work, which takes a product of forces and displacements. Also, impulse will have an impact on the motion of the object, as work did. But there are also many differences between these two quantities.

The first difference between impulse and work is that they obviously represent different physical quantities, because they have different units. While work has units of energy which we measure in Joules (or Newton-meters), impulse has units of force-times-time, measured in Newton-seconds. A second difference is that the impulse integral (mercifully) is not a line integral – there is no "path" to concern ourselves with when computing impulse. And third, because there is no dot product involved with the impulse integral, the result is a vector, in contrast to work, which is a scalar.

Definition of Momentum

The definition of impulse is not the end of the story, any more than the definition of work was. It needs to be related to the effect it has on the motion of the object. In the case of work, this relationship was expressed as the work-energy theorem:

$$\begin{aligned} \text{actions of pushes and pulls} \;\;&=\;\; W_{tot}\left(A \rightarrow B\right) \\ &=\;\; \int\limits_A^B \overrightarrow{F}_{net} \cdot \overrightarrow{dl} \\ &=\;\; \tfrac{1}{2}mv_B^2 - \tfrac{1}{2}mv_A^2 \\ &=\;\; \Delta KE \\ &=\;\; \text{effect of pushes and pulls} \end{aligned} \tag{4.1.2}$$

For the case of impulse, we find this relationship again by coming back to Newton's second law, and noting that the integral of acceleration is velocity:

$$\begin{aligned} \text{actions of pushes and pulls} &= \overrightarrow{J}_{tot}\left(t_A \rightarrow t_B\right) \\ &= \int\limits_{t_A}^{t_B} \overrightarrow{F}_{net} dt \\ &= \int\limits_{t_A}^{t_B} \left[m\overrightarrow{a}_{cm}\right] dt \\ &= \left[m\overrightarrow{v}_{cm}\right]_A^B \\ &= \Delta\left(m\overrightarrow{v}_{cm}\right) \equiv \Delta\overrightarrow{p}_{cm} \\ &= \text{effect of pushes and pulls} \end{aligned} \tag{4.1.3}$$

We call the quantity $\overrightarrow{p}$ *momentum* (which for a single object is defined as its mass multiplied by its velocity vector).

Definition: Momentum

$$\overrightarrow{p} \equiv m\overrightarrow{v}$$

Equation 4.1.3 is known as the *impulse-momentum theorem*. Like kinetic energy, momentum is related to the motion of the object (and the mass), but besides being a different function of mass and velocity than kinetic energy, it is also different in that it is a vector. This means that the total impulse can lead to a change in the magnitude or direction (or both) of the momentum vector.

The astute reader will undoubtedly realize that all we have really done here is to introduce a new vector, and use it to repackage Newton's second law. Indeed, we can rewrite the second law thus:

$$\overrightarrow{F}_{net} = m\overrightarrow{a}_{cm} = m\frac{d}{dt}\overrightarrow{v}_{cm} = \frac{d}{dt}\left(m\overrightarrow{v}_{cm}\right) = \frac{d}{dt}\overrightarrow{p}_{cm} \tag{4.1.4}$$

Given we are making comparisons between momentum and kinetic energy, it is useful to point out a direct mathematical relationship, which not only points out the difference between the two quantities, but will also be quite useful later on:

$$KE = \frac{1}{2}mv^2 = \frac{(mv)^2}{2m} = \frac{p^2}{2m} = \frac{\overrightarrow{p}\cdot\overrightarrow{p}}{2m} \tag{4.1.5}$$

Systems and Momentum Conservation

Let's continue following the trajectory from our discussion of work-energy, by returning to the idea of defining a system. As before, we define a system as an arbitrarily-grouped collection of objects, which can experience forces between themselves, or from outside the system. Previously we said that forces between objects within the system were responsible for internal work and forces exerted on objects within the system from outside provide external work. We will now similarly define internal impulse as coming from forces between objects within the system, and external impulse as coming from objects outside the system.

When it comes to forces between objects within our defined system, we know that the work done on one object does not cancel the work done on the other object. If the internal force is conservative, the non-zero total work done between the objects equals the energy converted between kinetic and potential energy. If the internal force is non-conservative, then the non-zero total work done between the objects equals the energy converted from mechanical to thermal energy. Is there an analogous process for impulse?

To answer this question, we need to determine whether impulses internal to a system don't cancel out, as in the case of work. We again start with Newton's third law, which ensures that the two forces involved in creating the pair of impulses are equal-and-opposite. Impulse vectors have the same directions as their associated force vectors, so the third law pair of forces results in a pair of impulses that are in opposite directions. But what about the magnitudes? Well, the force magnitudes are equal thanks to the third law, so all that remains is the time interval. There is never a moment when a force is acting that its third law pair isn't also acting, so the time intervals are the same. This leads to the following very important result: *All of the impulses internal to a system cancel each other out*. This means that there is no momentum analog to potential or thermal energy within a system. There is only momentum, and if the system experiences no external impulses, then momentum is conserved for the system. Comparing to what we got for energy, it looks like this:

$$\begin{array}{cc} \textbf{work-energy} & \textbf{impulse-momentum} \\ W_{ext} = \Delta KE \underbrace{- W_{cons} - W_{non-cons}}_{from\;internal\;forces} & \overrightarrow{J}_{ext} = \Delta \overrightarrow{p}_{cm} \underbrace{- \overrightarrow{J}_{cons} - \overrightarrow{J}_{non-cons}}_{from\;internal\;forces} \\ W_{ext} = \Delta KE + \Delta PE + \Delta E_{thermal} & \overrightarrow{J}_{ext} = \Delta \overrightarrow{p}_{cm} + 0 \quad + \; 0 \end{array} \tag{4.1.6}$$

There are two important features of this result:

1. ***It doesn't matter what forces are acting internally***. The result we obtained made no mention of whether the internal force was conservative or non-conservative – all forces satisfy Newton's third law, and the pairs act for equal periods of time, so the impulses cancel regardless of the nature of the force.
2. ***The quantity (momentum) that is conserved within the closed system is a vector***. This means that adding up all of the momentum vectors of a system at one point in time, then doing so again at another point in time, will give the same total vector in both cases, if the system is isolated from external impulses. This means that the total magnitude and direction don't change, or equivalently that the components measured in a given coordinate system don't change.

Example 4.1.1

Two blocks of different masses are attached to identical springs that are horizontal to the frictionless surface on which the block rests. If the springs are stretched the same distance and the blocks are released from rest, how do the following quantities compare for the two blocks when they reach the equilibrium point?

a. *kinetic energy*
b. *momentum*
c. *velocity*

Solution

a. The springs are stretched an equal amount, which means they both store the same potential energy. That means that when they get to the equilibrium point where they both have zero potential energy, they must have the same kinetic energy, since the mechanical energy is conserved.

b. We can determine the difference in momenta for the two blocks in two ways. First, we can consider the impulse given to each block by the spring. In the case of the more massive block, the spring force will accelerate it less, which means it will take longer to get to the equilibrium point. At every point during their journeys, the two blocks experience the same amount of force, but since the time interval for the heavier block is longer, it must experience the greater impulse. Therefore the heavier block gains more momentum, and since both blocks started with zero momentum, the heavier block must have more momentum at the equilibrium point. The second solution is much simpler: We already know that the two blocks end with the same KE, so since $KE = p^2/2m$, the block with the greater mass must have more momentum.

c. With the same kinetic energy, using $KE = \frac{1}{2}mv^2$, we see that the block with the greater mass must have the lower velocity.

The moral of this story: Although we tend to use kinetic energy, momentum, and velocity as proxies for motion, they are all quite different quantities.

Partial Momentum Conservation

We have to give some extra thought to what we mean by a conserved vector. Since a vector has both magnitude and direction, then to be conserved, both of those properties must remain unchanged. An equivalent way of saying this is that for the vector to be conserved, every component of that vector must be individually conserved. If the full momentum vector is not conserved, it is still possible for one or two of its components to be conserved, if the components of the external impulse in those directions is zero. So for example, a projectile (with no air resistance) conserves momentum in the two horizontal directions, but not in the vertical direction. This allows us to use momentum conservation to solve a much broader range of problems than if we can only consider complete momentum conservation.

4.2: Center of Mass

Systems with Multiple Particles

When we studied the work-energy theorem, we found that the main advantage to this approach was the nice "shortcut" available to us when we selected a system large enough that the energy within the system was conserved. So it is with the impulse-momentum theorem. We will therefore largely set aside cases where the system has an external impulse acting on it, and focus only on systems that conserve momentum.

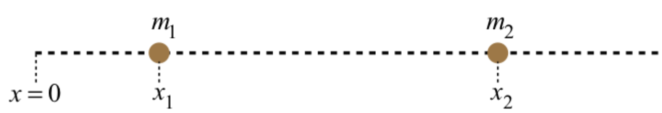
In many cases we have dealt with so far, our "system" has been only a single object, or perhaps two objects, where one of them (i.e. the earth) is so large that we can ignore it because it basically just sits there as it provides a potential energy that affects the motion the other object. But now we will embrace multiple *active* objects, meaning that their masses will not be so different that one object's motion can be ignored. Now we will need to account for the motions of all objects within the defined system. Note that we make no claims at this point about the relative interactions of these objects – they may be securely attached to each other in a rigid embrace (a solid object), or they may be totally free to move independently of one another (a volume of gas).

Consider the impulse-momentum theorem for a system consisting of several objects. The momentum part of this theorem consists of two pieces: The mass, which we presume is the sum of the masses of all the objects within the system; and the velocity of the center of mass of the system. We have a pretty intuitive notion of center of mass for a single object, but how do we talk about the center of mass of a system with many separate moving objects? We take some time away from our discussion of momentum conservation to address this.

Center of Mass (of a Collection of Particles)

In some sense, one can think about the center of mass of a single object as its "average position." Let's consider the simplest case of an "object" consisting of two tiny particles separated along the x-axis, as in Figure 4.2.1.

Figure 4.2.1 – Center of Mass for Two Point-Particles



If the two particles have equal mass, then it's pretty clear that the "average position" of the two-particle system is halfway between them.

> **Alert**
>
> *Center of mass is a mathematical construct, not an actual position that resides on a physical object. The center of mass of a system often lands at a position consisting of empty space, whether that is because the system consists of multiple objects, or because the single object in the system is bent or has a hole in it.*

If the masses of the two particles are different, would the "average position" still be halfway between them? Perhaps in some sense this is true, but we are not looking for a *geometric center*, we are looking for the average placement of mass. If m_1 has twice the mass of m_2, then when it comes to the average placement of mass, m_1 gets "two votes." With more of the mass concentrated at the position x_1 than at x_2, the center of mass should be closer to x_1 than x_2. We achieve the perfect balance by "weighting" (no pun intended) the positions by the fraction of the total mass that is located there. Accordingly, we define as the center of mass:

$$x_{cm} = \left(\frac{m_1}{m_1+m_2}\right)x_1 + \left(\frac{m_2}{m_1+m_2}\right)x_2 = \frac{m_1x_1+m_2x_2}{M_{system}} \tag{4.2.1}$$

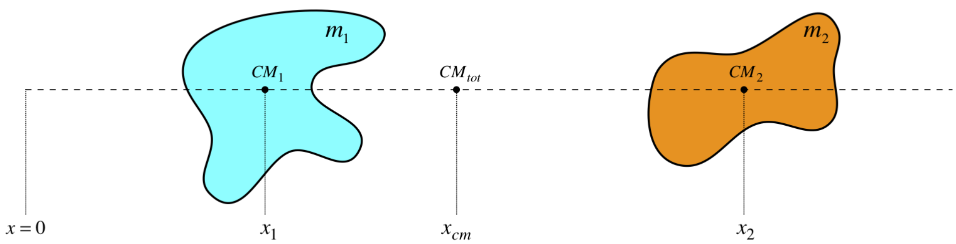
If there are more than two particles, we simply add all of them into the sum in the numerator. To extend this definition of center of mass into three dimensions, we simply need to do the same things in the y and z directions. A position vector for the center of mass of a system of many particles would then be:

$$\begin{aligned}\vec{r}_{cm} &= x_{cm}\,\hat{i}+y_{cm}\,\hat{j}+z_{cm}\,\hat{k}\\ &= \frac{[m_1x_1+m_2x_2+\ldots]\,\hat{i}+[m_1y_1+m_2y_2+\ldots]\,\hat{j}+[m_1z_1+m_2z_2+\ldots]\,\hat{k}}{M}\\ &= \frac{m_1\left[x_1\hat{i}+y_1\hat{j}+z_1\hat{k}\right]+m_2\left[x_2\hat{i}+y_2\hat{j}+z_2\hat{k}\right]+\ldots}{M}\\ &= \frac{m_1\vec{r}_1+m_2\vec{r}_2+\ldots}{M}\end{aligned} \tag{4.2.2}$$

Center of Mass (of a Collection of Objects)

Suppose now we want to know the center of mass of multiple extended objects, where all the heavy-lifting has already been done – the centers of mass of the objects are already known (see below for how to do this heavy-lifting). How do we determine the center of mass of such a system? It turns out to be pretty easy when you know the locations of the centers of mass of the two objects – just treat them as if they are point particles with all of their mass concentrated at their own centers of mass, and then do the calculation above.

Figure 4.2.2 – Center of Mass for Two extended objects



For proof of this, let's treat two extended objects (A and B) as collections of lots of point particles (atoms, if you like), and write down their centers of mass (measured from a common origin) in terms of the masses and positions of their atoms.

$$\left.\begin{aligned}\vec{r}_{cm\ A} &= \frac{m_{1A}\vec{r}_{1A}+m_{2A}\vec{r}_{2A}+\ldots}{M_A}\\ \vec{r}_{cm\ B} &= \frac{m_{1B}\vec{r}_{1B}+m_{2B}\vec{r}_{2B}+\ldots}{M_B}\end{aligned}\right\} \quad\Longleftrightarrow\quad \vec{r}_{cm}=\frac{M_A\vec{r}_{cm\ A}+M_B\vec{r}_{cm\ B}}{M_A+M_B} \tag{4.2.3}$$

The left-hand side equations are those of the center of mass for each object in terms of its atoms' masses and positions. The right-hand side gives the center of mass of the two-object system in terms of the masses of the objects and the positions of their individual centers of mass. When the expressions for $\vec{r}_{cm\ A}$ and $\vec{r}_{cm\ B}$ from the left side are plugged into the right-hand side equation, then all the atoms of both objects are come together into a single center of mass formula, as if they were part of a single system with total mass M_A+M_B, proving the contention above.

Example 4.2.1

Two thin circular disks made from the same material lie flat on a horizontal surface, with their outer edges in contact with each other. One disk has a larger radius (R) than the other (r), and have equal thicknesses. Find how far the center of mass of the two-disk system lies from the center of the larger disk.

Solution

The disks are made from the same uniform material, so they have equal mass densities. That means that the mass of the larger disk is larger than that of the smaller disk by the same factor as the ratio of their areas. That is, if the larger disk has twice the area of the smaller one, then it has twice as much mass. We therefore have the following relationship between the masses and radii of the disks:

$$\frac{M}{m}=\frac{\pi R^2}{\pi r^2} \quad\Rightarrow\quad M=\frac{R^2}{r^2}\,m$$

Let's choose the center of the larger disk as the origin, and have the center of the other disk lie on the $+x$-axis. The disks are uniform, so their individual centers of mass lie at their geometric centers, and we can compute the center of mass of the system by treating the disks as point masses located at these centers. The distance of the center of mass from the origin is what we are looking for, so:

$$x_{cm}=\frac{Mx_1+mx_2}{M+m}=\frac{M\left(0\right)+m\left(R+r\right)}{M+m}=\frac{m\left(R+r\right)}{\frac{R^2}{r^2}m+m}=\boxed{\frac{\left(R+r\right)r^2}{R^2+r^2}}$$

We can double-check this answer by looking at an obvious special case: $R=r$. If the disks are identical, then the center of mass must be halfway between their centers, which is the point where they are in contact, a distance R from the center of the larger disk. Plugging in R for r indeed gives this answer.

Center of Mass (of Continuous Objects)

We now turn to the problem of computing the position of the center of mass of an object whose distribution of mass is known. What follows is pure math, but it is important math that returns over and over in physics.

Alert

The important thing to gain from this discussion is to understand how the set-up process works. It culminates in an integral, but performing the integral is mere busywork compared to the task of setting it up. It's easy to be overwhelmed by the thought of the integral that is being constructed, but if you understand each step that leads up to it (and don't try to just jump to an answer that looks like something you have seen before), it will go fine.

We will keep this simple by restricting ourselves to objects for which the position of the center of mass in two of the three dimensions is obvious, which means we don't need to concern ourselves with the whole vector described in Equation 4.2.2 – just the x-component will do. A good model for this is a simple thin, cylindrical rod. This rod's mass distribution is completely cylindrically symmetric, which means that the center of mass lies on the axis passing through its center. But the mass distribution as a function of position on this axis may not be uniform. For example, it may be more dense on one end than on the other. Put another way, the particles located within the rod may be packed together more tightly in one region of the rod than in another, which means that the center of mass will not necessarily lie at the point halfway between the ends.

We need to say a few words about *mass density* before we proceed. Density is a measure of how closely-packed in space a quantity of something is. This quantity can be many different things. Here we will be considering mass, but in later physics classes you will deal with density of electric charge (and even, bizarrely, probability!). A *uniform density* for a region in space means that the quantity (whatever it happens to be) is evenly-distributed everywhere within that region. The way we define an average density for a region in space is to add up how much "stuff" is there, and then divide it by the total space it occupies. This gives an average density, but of course densities can vary from one point in space to another, in which case a *density function* is defined. We will deal with only the simplest variable densities here. As we will mainly be looking at thin rods for our examples, we will only consider densities that might vary along the length of the rod – this simplifies the process to a single dimension.

The mass density function in this case is a function of a single variable, has units of kg/m, and is called a *linear mass density*. This mass density function is typically denoted as $\lambda\left(x\right)$. If it is uniform, then the function is a constant λ, and the amount of mass m within a given length l is simply given by:

$$m=\lambda l \tag{4.2.4}$$

If the density is not uniform, then it is only a constant over an infinitesimal length dx, so Equation 4.2.4 can only apply to a tiny piece of mass dm, and the relationship is different at every position x because the density is different at every position:

$$dm=\lambda\left(x\right)dx \tag{4.2.5}$$

Now that we can write down how much mass is at every position, we are ready to do our calculation. We begin by drawing a diagram with the rod in a coordinate system along the x-axis such that one end is at the origin and the other is at $x=L$. Figure 4.2.3 provides a fully-labeled diagram that is very helpful for solving such problems.

Figure 4.2.3 – Setup Diagram for Computations Involving Mass Density of a Thin Rod

The center of mass is found by multiplying the amount of mass at each point by the x-coordinate of that mass, then adding up all of those products and dividing by the total mass. Of course, in this case we have an infinite number of point masses, so the sum is infinitely long, but the masses are infinitesimally small, so we solve this by converting the sum into an integral, in which we add up all the pieces from $x=0$ to $x=L$:

$$x_{cm} = \frac{dm_1 x_1 + dm_2 x_2 + \dots}{dm_1 + dm_2 + \dots} = \frac{\int\limits_{x=0}^{x=L} dm\ x}{\int\limits_{x=0}^{x=L} dm} \tag{4.2.6}$$

Now we plug in Equation 4.2.5 to give the following formula for center of mass (in one dimension) for a thin rod with a linear mass density that varies with x:

$$x_{cm} = \frac{\int\limits_{x=0}^{x=L} \lambda(x)\, x dx}{\int\limits_{x=0}^{x=L} \lambda(x)\, dx} \tag{4.2.7}$$

Okay, so let's do a couple of examples...

A Uniform Rod

As was stated above, if the rod is uniform, then the density is a constant (which we will call simply λ). Plugging this into Equation 4.2.7 leads to a simple calculation and an unsurprising result:

$$x_{cm} = \frac{\int\limits_{x=0}^{x=L} \lambda x\, dx}{\int\limits_{x=0}^{x=L} \lambda\, dx} = \frac{\cancel{\lambda}\int\limits_{x=0}^{x=L} x\, dx}{\cancel{\lambda}\int\limits_{x=0}^{x=L} dx} = \frac{\left[\frac{1}{2}x^2\right]_0^L}{\left[x\right]_0^L} = \frac{1}{2}L \tag{4.2.8}$$

So we have calculated what we already knew – that for a thin rod with a uniform mass density, the center of mass is at its center (which on our coordinate system lies at $L/2$).

A Non-Uniform Rod

Next we'll look at an example of a rod which has a mass density that varies from one end to the other. This variable density is expressed in its density function:

$$\lambda(x) = \lambda_o\left(\frac{x}{L} + 1\right) \tag{4.2.9}$$

Before we do the math, let's try to make sense of this function. The easiest way to do this is to consider the endpoints. At $x=0$, the density equals the constant λ_o, while at $x=L$ that density has grown to twice that much. This increase of density happens linearly with the variable x. What should we *expect* to see when we compute the center of mass? Well, the rod is more dense near the $x=L$ end than the $x=0$ end, so the center of mass should be at an x value greater than $L/2$. Okay, so let's plug the density function into Equation 4.2.7 and see what we get:

$$x_{cm} = \frac{\int\limits_{x=0}^{x=L} \left[\lambda_o \left(\frac{x}{L}+1\right)\right] x\, dx}{\int\limits_{x=0}^{x=L} \left[\lambda_o \left(\frac{x}{L}+1\right)\right] dx} = \frac{\cancel{\lambda_o} \int\limits_{x=0}^{x=L} \left(\frac{x^2}{L}+x\right) dx}{\cancel{\lambda_o} \int\limits_{x=0}^{x=L} \left(\frac{x}{L}+1\right) dx} = \frac{\left[\frac{x^3}{3L}+\frac{x^2}{2}\right]_0^L}{\left[\frac{x^2}{2L}+x\right]_0^L} = \frac{5}{9}L \tag{4.2.10}$$

Interestingly, the center of mass doesn't depend upon the density constant λ_o.

4.3: Momenta of Systems

Momentum Conservation

Our reason for studying center of mass was to make sense out of its role in momentum conservation. We are now able to unpack the "$\Delta \overrightarrow{p}_{cm}$" from Equation 4.1.3. We know that the delta means "after minus before," so let's focus on the rest of it:

$$\overrightarrow{p}_{cm} = M\overrightarrow{v}_{cm} = M\frac{d}{dt}\overrightarrow{r}_{cm} = \cancel{M}\frac{d}{dt}\left[\frac{m_1\overrightarrow{r}_1 + m_2\overrightarrow{r}_2 + \dots}{\cancel{M}}\right] = m_1\frac{d}{dt}\overrightarrow{r}_1 + m_2\frac{d}{dt}\overrightarrow{r}_2 + \dots = m_1\overrightarrow{v}_1 + m_2\overrightarrow{v}_2 + \dots = \overrightarrow{p}_1 + \overrightarrow{p}_2 + \dots \tag{4.3.1}$$

So we see a very important result:

The momentum of a system, defined as the total mass of the system multiplied by the velocity of the center of mass of that system, equals the (vector) sum of the momenta of the individual objects contained in the system.

Alert

Reiterating the previous alert... The velocity of the center of mass of a system is the velocity of a mathematical construct, not of a particular piece of matter. The objects within the system move as they do, and the center of mass's motion is derived from all these motions. Indeed, it is possible for the center of mass to remain stationary even though every object in the system is moving, if their momenta happen to cancel out to zero.

If we now put the delta back in, and assume that the external impulses sum to zero, we get the following expression of the momentum conservation of the system:

$$0 = \Delta\overrightarrow{p}_{cm} = \Delta\overrightarrow{p}_1 + \Delta\overrightarrow{p}_2 + \dots \tag{4.3.2}$$

When we are talking about multiple objects, it is probably more intuitive to take an equivalent "before = after" viewpoint. This consists of breaking up all of the $\Delta\overrightarrow{p}$'s into $\overrightarrow{p}_{after} - \overrightarrow{p}_{before}$, and collecting all of the "befores" on one side of the equation, and the "afters" on the other. This looks like:

$$\underbrace{\overrightarrow{p}_1 + \overrightarrow{p}_2 + \dots}_{before} = \underbrace{\overrightarrow{p}_1 + \overrightarrow{p}_2 + \dots}_{after} \tag{4.3.3}$$

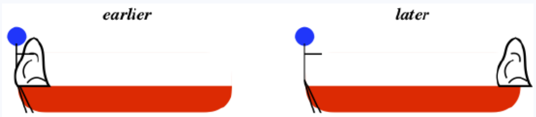
The sum on each side is over the several objects in the system. So adding up the momentum vectors of all the objects before some event, and then doing it again after the event gives the same vector. This of course assumes that the "event" does not involve an external impulse, though it can include as many internal impulses as you like.

Alert

It is important to remember that this equation does not mean that each of the terms remains unchanged. Rather, they change in such a way that the changes all compensate for each other, and the sum of the all-new momentum vectors comes out to the same sum that came out before.

Example 4.3.2

A 30.0kg child sits on the rear end of a 12.0kg, 3.50m long sled (whose mass is uniformly-distributed along its length) with a 18.0kg block of frozen snow at rest in her lap. The sled is sliding forward on the horizontal, frictionless snow at constant a speed of 1.60m/s, when the girl suddenly shoves the block forward in the sled (she remains firmly planted on the sled). After a time of 2.50s, the block comes to rest in the front of the sled.



a. *Find the speed of the sled after the block stops sliding forward.*

b. *Find the position of the center of mass of the girl + sled + block before the block is pushed. Reference this position from the rear of the sled.*
c. *Find the distance (relative to the ground) that the center of mass of the girl + sled + block moves during the period of time that the block slides forward.*
d. *Find the distance that the sled moves during the period of time that the block slides forward.*

Solution

a. One can do all the "in-between" math to get this answer, but the simplest method is to note that all the forces between the child, snow, and sled are internal to the system, and the system experiences no net external force, which means that the motion of its center of mass remains unaffected. The center of mass of the system was moving at 1.60m/s before the child shoved, and it will be moving at the same speed as the snow moves forward, and the same speed after the snow stops. The sled will slow be moving slower than its initial speed as the snow is sliding forward, because the sled is not moving with the center of mass of the system, but when the snow reaches the front, the whole system is once again moving together.

b. To find the center of mass of the system, we treat each object in the system as if it is a point mass existing at its own center of mass, and plug into the center of mass equation. Choosing the rear of the sled as the origin (our reference point), the girl and snow are at the origin, while the center of mass of the sled is half its length from the origin, giving:

$$x_{cm}=\frac{m_{girl}x_{girl}+m_{snow}x_{snow}+m_{sled}x_{sled}}{m_{girl}+m_{snow}+m_{sled}}=\boxed{0.350m}$$

c. The speed of the system remains unchanged during this time (as indicated in part a), so the distance that the center of mass moves during this period is just the speed times the time:

$$\Delta x_{cm}=v_{cm}t=\boxed{4.00m}$$

d. However far the sled travels, the girl travels the same distance, and the snow travels that distance plus the length of the sled. Plugging in all these changes into the equation for the change of center of mass (of the whole system) gives:

$$\Delta x_{cm}=\frac{m_{girl}\Delta x_{girl}+m_{snow}\Delta x_{snow}+m_{sled}\Delta x_{sled}}{m_{girl}+m_{snow}+m_{sled}}=\frac{m_{girl}\Delta x_{sled}+m_{snow}\left(\Delta x_{sled}+L\right)+m_{sled}\Delta x_{sled}}{M}$$

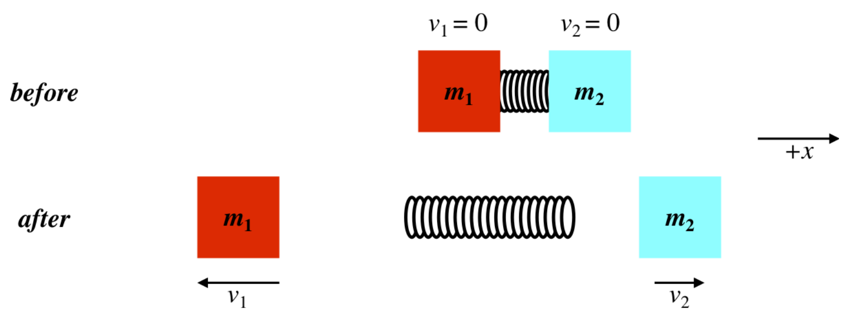
Solving for the displacement of the sled gives:

$$\Delta x_{sled}=\Delta x_{cm}-\frac{m_{snow}L}{M}=\boxed{2.95m}$$

Using Center of Mass

Let's look at an example of how we can use what we know about center of mass the analyze a case of two blocks of different masses that squeeze a (massless) spring between them until they are released from rest.

Figure 4.3.1 – Repelling Masses



Intuitively one can probably tell that for this situation $m_2 > m_1$. When a light object pushes off a heavy one (a flea jumping off a dog, a bullet leaving a gun, etc.), the lighter object's motion is always affected more. With our physics training, we can explain it with Newton's second and third laws: The blocks push on each other with equal forces (third law), and with equal forces, the block with less mass will accelerate more. They both start from rest and are pushed for equal periods of time, so the one with the greater acceleration will be going faster when they separate, sending it a greater distance in the same time period.

Okay, now let's look at it from the perspective of momentum conservation. Treating the two blocks as a single system, the spring force produces only internal impulses, which means that the momentum of the system is conserved. The momentum before the spring unloads is zero, so it must be zero afterward. If v_1 and v_2 are the speeds of the two blocks (i.e. these are positive numbers), then we have for our conservation equation:

$$momentum\ before = momentum\ after \quad \Rightarrow \quad 0 = m_1 v_1 \left(-\hat{i}\right) + m_2 v_2 \left(+\hat{i}\right) \quad \Rightarrow \quad v_1 = \frac{m_2}{m_1} v_2 \tag{4.3.4}$$

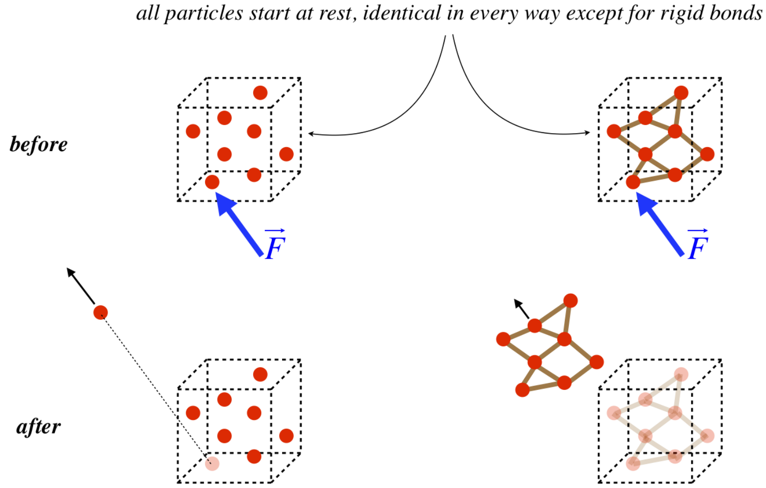
Since it's clear from the diagram that $v_1 > v_2$, it must be that $m_2 > m_1$. We can also use what we know about center of mass here. The system experiences no external net impulse and its center of mass is stationary, so it must remain stationary! We don't know exactly where the center of mass is before the repulsion, but since it stays put, we can draw a vertical line down into the second diagram to find where it is after the repulsion. This clearly results in the center of mass being closer to m_2, which means that is the larger mass.

Center of Mass Acceleration

Let's see if we can incorporate what we have learned about center of mass to make sense of Newton's second law. Consider the two systems shown in Figure 4.3.2. Each consists of a collection of 8 identical particles in close proximity to each other (the boxes shown are just used as a reference for later motion – they are not physical objects). In the left system, the particles are floating freely (there is no gravity or other forces), while in the right diagram, the particles are bound together with rigid, massless rods. The two systems are identical in every way except for the presence of these rods – the particle all have the same positions and masses as their counterparts, and are all at rest.

Now for the experiment: Suppose we exert the same force on the same particle in both systems. Clearly the reaction is different in the two cases – in the left case, only the particle given the push accelerates away, while in the right cases the entire group of particles accelerates. The question is, in which case does the center of mass of the system of particles accelerate *more*?

Figure 4.3.2 – Forces on Free and Rigid Systems



Here is the short answer: The forces that are (or are not) between the particles defining the system are internal, and therefore have no effect on the velocity of the system's center of mass. The only external force on each system is $\overrightarrow{F}$, and each system has the same mass, so Newton's second law says that both systems should react with the same acceleration of their center of mass.

But that is unconvincing when we see only one particle move in one case, and the whole conglomerate move in the other! Let's suppose the forces act for some small period of time. The acceleration of the single particle will be eight times greater than that of the conglomerate, so in the same time interval it will move eight times as far as the conglomerate. Let's call the initial position of the center of mass the origin. The seven particles left behind experience no change in their position relative to this origin, and the one particle's position relative to the origin travels eight units of distance, while all eight of the particle in the other system travel just one unit from their original positions relative to the origin. Treating the direction of motion as the $+x$ direction, and plugging the masses and distances into Equation 4.2.1, it should be immediately clear that both centers of mass move by the same amount. As strange as it sounds, Newton's second law works for any system of particles, whether they bond together to form a solid object, or are completely independent of each other, like particles in a gas.

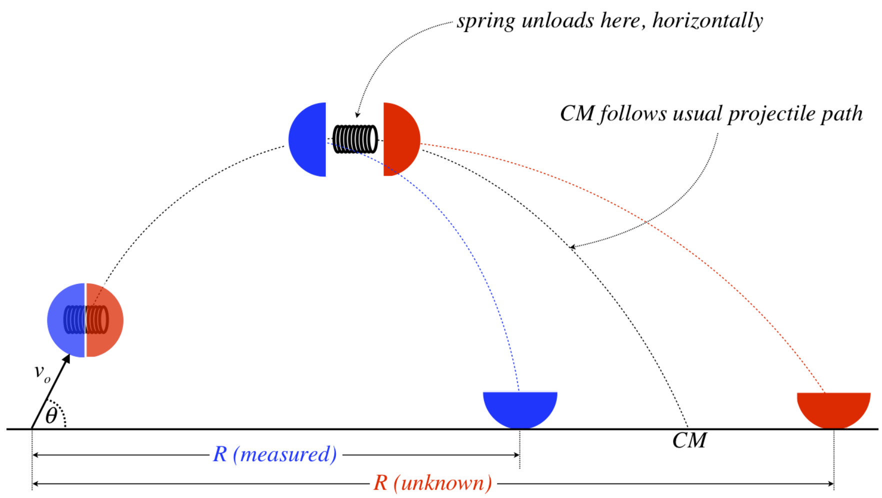
Center of Mass Frame

Sometimes analysis of problems that involve multiple objects interacting with each other is simplified by using what is called the *center of mass frame* of reference. Here's an example.

A child's toy called a "hot potato" consists of two hemispherical shells that close on a spring and are held together by a latch on a timer. When the time expires, the latch is released and the spring is allowed to expand, shooting the two shells in opposite directions, exposing the toy company to a product liability lawsuit from the family of the child that holds the hot potato when it goes off.

Let's suppose a child throws this hot potato through the air, and the peak of its projectile motion, it explodes so that the two shells are propelled horizontally, as shown in Figure 4.3.3. The landing point of the shell that lands closest to where the toy was thrown is noted, but the other shell flies off into some tall grass and is lost. Naturally the child knows the starting speed they gave the toy as well as the exit angle, and she can easily measure the distance that the closer shell travels form the launch point. From this information and her vast knowledge of physics, she conceives of a plan to find the other shell that is far more elegant than searching for it in the tall grass.

Figure 4.3.3 – Exploding Projectile

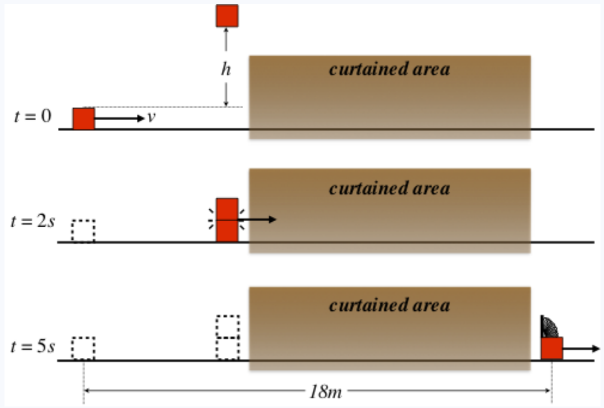


The forces on the shells by the spring are internal to the two-shell system, so assuming air resistance is negligible, the center of mass of the system will behave exactly as it would if the internal forces didn't exist. With the starting angle and speed known, the child can use the range equation (see Example 1.7.4) to calculate the landing point of the center of mass of the system. Then with the actual landing point of one piece of the toy, she can use the center of mass formula to compute the landing point of the other piece.

You might think we can do the same even if the spring unloads in an orientation that is other than horizontal, but this is not the case. The center of mass motion still follows the same parabolic trajectory, but naturally the center of mass is always between the two shells. In the case above, the shells land simultaneously (they both start with zero vertical component of velocity when the explosion occurs), so the center of mass lands at the same time, between the shells. When the explosion is not horizontal, one shell lands before the other, then friction stops is horizontal motion while the other shell keeps moving horizontally. This makes calculating the landing point of the center of mass using the usual range equation impossible.

Example 4.3.3

A 1.4kg block slides along a frictionless horizontal surface at a speed of v = 4.8m/s, starting at position x = 0 and time t = 0. At time t = 2.0s, an identical block lands directly on top of it. The surfaces of the blocks that are sticky, so the top block adheres to the bottom block when it lands on it, and they continue along together. The blocks slide together into a curtained-off area, after which you hear a spring noise and a "thud." At time t = 5.0s, the bottom block emerges from the curtain at position x = 18.0m *without the top block on it, after apparently having its top lid sprung open from within. Find the x position of the other block at this time (which is presumed to be on the frictionless surface, still within the curtained area).*



Solution

For the collision of the two blocks, the vertical momentum is not conserved (gravity and the normal force from the floor don't cancel during the collision, so there is a net vertical impulse), but the horizontal momentum is conserved. This because the horizontal frictional (or "adhesive force," if you prefer) is internal to the two-block system. The blocks have equal mass, so setting in incoming horizontal momentum of a single block equal to the outgoing horizontal momentum of both blocks gives:

$$mv_o + 0 = 2mv_f \quad \Rightarrow \quad v_f = \frac{1}{2}v_o = 2.4\frac{m}{s}$$

The horizontal momentum of the two-block system remains constant for the entire time, which means that this system's center of mass continues at a constant horizontal velocity. This center of mass velocity is clearly the "final" velocity computed above, since at $t = 2.0s$ the two blocks are moving together. All we need therefore is the starting x-position of the system's center of mass and the time elapsed, and we can determine the new position of the center of mass. The starting position of the two-block system's center of mass is the x-position halfway between the blocks (since they have equal masses), and we can set the starting position of the sliding block to be the origin, but what is the x-position of the falling block?

We are given that the time that the sliding block takes to reach the falling block, and we have its speed, so we know how far it the two blocks are separated in the x-direction, and with it, the starting position of the center of mass:

$$x_{cm}\left(t = 0s\right) = \frac{1}{2}v_o t = 4.80m$$

With the speed of the center of mass of the two-block system, the starting position of the center of mass, and the time elapsed, we know where the center of mass is later:

$$x_{cm}\left(t\right) = x_{cm}\left(t = 0s\right) + v_{cm}t = 16.8m$$

The blocks have equal mass, so this center of mass positon lies halfway between the two blocks, and since we know the position of the sliding block, we can get the position of the falling block:

$$x_{cm} = \frac{1}{2}\left(x_{sliding} + x_{falling}\right) \quad \Rightarrow \quad x_{falling} = \boxed{15.6m}$$

Rocketry

While we are on the topic of two parts of a system going their separate ways by pushed off each other, this brings us to the topic of rocketry. A rocket that is stationary in space somehow is able to accelerate itself by firing its engines. How can the center of mass of the rocket system accelerate without any external forces acting on it? Well, it can't of course, but the rocket (or rather, its fuselage) is not an isolated system. It expels fuel (in the form of very hot gas) backward. If we include the fuel as part of the system, then the center of mass of the system doesn't accelerate at all! All that matters in the end is that the fuselage of the rocket is propelled forward. Note also that the rocket has more mass than the fuel, but the ignited fuel sends particles away at very high speeds, and this momentum balances the momentum of the fuselage in the opposite direction (which has more mass and lower velocity).

4.4: Momentum and Energy

Revisiting the Work-Energy Theorem

The work-energy theorem discussed in the previous section was derived from Newton's second law, which carries with it a reference to the center of mass of the object on which the force is acting. So technically, the velocity and displacement that appear in the work-energy theorem are the velocity and displacement of the center of mass, which would suggest altering Equation 4.1.4 to:

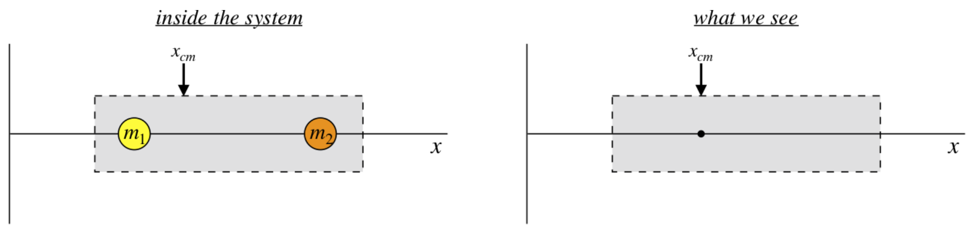
$$\Delta\left(\frac{1}{2}mv_{cm}^2\right)=\int_A^B \overrightarrow{F}_{net}\cdot\overrightarrow{dl}_{cm} \tag{4.4.1}$$

While accurate, this introduces a lot of cumbersome subscripts, which are entirely unnecessary for the kinds of physical situations we had in mind – a force acting on a solid object to accelerate every part of the object the same as its center of mass. But this does suggest an interesting extension of the idea: What if the force acts on one part of a system that includes multiple objects? Indeed, even a solid object is technically comprised of lots of particles, and many forces that act on such an object are only exerted on a fraction of the particles. We now explore this idea.

An Instructive Model – A System of Two Particles

Let's explore the work-energy theorem in the context of a system of two particles of differing masses. For simplicity, we'll keep everything in one dimension – the particles can only move along the x-axis, and the force that does the work can only act parallel to the x-axis. The thing to remember about our two-particle system is that we will be watching how the *system's* motion evolves, regardless of what happens to the individual particles. One can imagine cloaking the details of the particle motions and just watching the motion of the center of mass of the conglomerate.

Figure 4.4.1 – A System of Two Particles

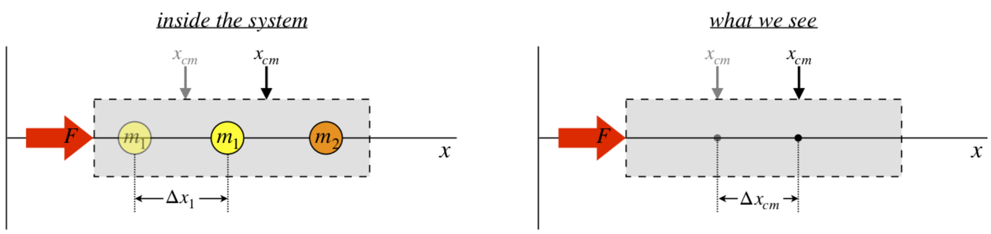


For the two-particle system shown in Figure 4.4.1, the center of mass is closer to m_1 than m_2, which means that $m_1>m_2$. In terms of the positions of the two particles, the center of mass location is found using Equation 4.2.1:

$$x_{cm}=\frac{m_1x_1+m_2x_2}{m_1+m_2} \tag{4.4.2}$$

Now let's apply a force on this system (which we will assume starts from rest) from the left, which means it is exerted only on m_1, but as far as we are concerned, it is exerted on the system (we can't see m_1).

Figure 4.4.2 – Work Performed on a System of Two Particles



After m_1 starts moving, the velocity of the center of mass is:

$$v_{cm}=\frac{d}{dt}x_{cm}=\frac{m_1\frac{d}{dt}x_1+m_2\frac{d}{dt}x_2}{m_1+m_2}=\left(\frac{m_1}{m_1+m_2}\right)v_1 \tag{4.4.3}$$

We can combine this force with the displacement of the center of mass to find the work done on the system (m_2 doesn't displace at all, so $\Delta x_2=0$):

$$W_{on\ system}=F\Delta x_{cm}=F\frac{m_1\Delta x_1+m_2\Delta x_2}{m_1+m_2}=\left(\frac{m_1}{m_1+m_2}\right)F\Delta x_1 \tag{4.4.4}$$

We want to check to see if this work results equals what we measure for the change of kinetic energy of the system (which starts at zero), so we calculate the final kinetic energy using the total mass of the system and the speed of the center of mass from Equation 4.4.3:

$$\Delta KE_{system}=\frac{1}{2}(m_1+m_2)v_{cm}^2=\frac{1}{2}(m_1+m_2)\left[\left(\frac{m_1}{m_1+m_2}\right)v_1\right]^2=\left(\frac{m_1}{m_1+m_2}\right)\left[\frac{1}{2}m_1v_1^2\right] \tag{4.4.5}$$

Comparing Equation 4.4.4 and Equation 4.4.5, we see that the work-energy theorem applies to the system as a whole if:

$$W_{on\ system}=\Delta KE_{system}\quad\Longleftrightarrow\quad F\Delta x_1=\frac{1}{2}m_1v_1^2 \tag{4.4.6}$$

But we recognize the equation as the work-energy theorem applied to m_1, so we have demonstrated that the work-energy theorem is equally applicable to systems of particles as individual ones.

Mechanical and Internal Energy

The reader may be puzzled about why the same force acting on the same system appears to transfer two different amounts of energy, depending upon one's perspective. From the particle point of view, the energy transferred to the two particle system is $\frac{1}{2}m_1v_1^2$ to one particle and zero to the other for a total energy transfer of $\frac{1}{2}m_1v_1^2$. From the view of someone looking at the system as a whole from outside, the system gains the same amount of energy, reduced by a fraction of $\frac{m_1}{m_1+m_2}$. It clearly can't be both, so which amount of energy does the system really get?

If we imagine a system as a closed box with a bunch of particles in it, the box has an energy equal to the sum of the masses of the particles. If all of the particles happen to be moving in the same direction at the same speed, then the box must also be moving, and the kinetic energy of the system equals the sum of the kinetic energies of the particles. But suppose while all the particles move at the same speed, half are going in the opposite direction as the other half. In this case, the center of mass remains stationary, and the kinetic energy of the box is zero (one-half the total mass times the square of the center of mass velocity). Clearly the energy in the system is not zero, but from our outside-the-box perspective, we are unable to witness it directly.

The kinetic energy of the system as a whole is what we have been referring to as "mechanical" in nature. The remaining energy that is hidden to us due to individual motions of the particles being concealed within the box we refer to as *internal energy*. To see how this internal energy is defined, let us return again to the two-particle model above. We already have the (mechanical) kinetic energy of the system, given by Equation 4.4.5. Now let's place ourselves within the system by changing reference frames to the rest frame of the system. That is, suppose we are moving along with the center of mass of the system, and measure the total kinetic energy of the two particles. To make this change of frames, we use the method described back in Section 1.8. The velocity of each particle in the new frame is the velocity vector in the "laboratory" frame, minus the velocity vector of the center of mass. The kinetic energy of the two particles in this frame is:

$$KE_{internal}=\frac{1}{2}m_1(v_1-v_{cm})^2+\frac{1}{2}m_2(v_2-v_{cm})^2 \tag{4.4.7}$$

In the example above, $v_2=0$. Plugging this and Equation 4.4.3 into Equation 4.4.7 and doing a little bit of tedious algebra gives the result:

$$KE_{internal}=\left(\frac{m_2}{m_1+m_2}\right)\left[\frac{1}{2}m_1v_1^2\right] \tag{4.4.8}$$

Comparing this result with that for the system's kinetic energy (Equation 4.4.5), we see that the sum of the mechanical energy given to the system and the internal energy given to the system is indeed the total energy given to the system. Put mathematically:

$$W_{external} = \Delta KE_{mechanical} + \Delta KE_{internal} \tag{4.4.9}$$

It is important to note that this is not a modification of the work-energy theorem. The right hand side of this equation still equals the total kinetic energy of the system. It just repackages it to accommodate what we can readily see (the mechanical energy) and what we cannot (the internal energy). We can once again take the steps we outlined previously to construct our energy conservation models.

Alert

Use of the word "system" in Section 3.4 is subtly different from how the word is used here. Here it refers to a collection of particles within a single object, allowing us to distinguish mechanical energy from internal energy for that object. Our previous use of the word "system" referred to a collection of objects. What we refer to as "external" work here could be work done between objects that are within the same collection-of-objects system, and should not be confused with work that could be done from outside the system of objects. If this external work between objects doesn't result in any internal energy created within the objects, then the force between the objects is conservative, and can be expressed as a potential energy. It can be confusing to keep these different system definitions straight, and it might help to remember our current discussion as "systems of particles," and the previous discussion as "systems of objects," with objects being collections of particles.

Notice how important the model we use is to the definition of what energy is mechanical and what energy is internal. If our model is constructed to take into account the movements of all the particles, then all of the energy is mechanical. But as soon as we have to lump together particles into a system (whether by choice or out of necessity), we define this division of energy types. No matter how we define our system and do our accounting, however, the total energy is still conserved.

Notice also that this doesn't just apply to internal *kinetic* energy. If the particles within the system interact with each other through some internal force, then the potential energy that results goes into the accounting of the internal energy. In our two-particle model, we might imagine a spring connecting the two particles that is compressed when the force is applied. Even in this case, the particles get closer together, which means that the work done on m_1 is greater than the work done on the system as a whole (the center of mass doesn't move as far as m_1), resulting in some of the energy going internal. In this case, the internal energy is manifested by the two particles vibrating back-and-forth as the center of mass of the system moves along at a steady speed.

Demystifying Non-Conservative Forces and Thermal Energy

When we discussed non-conservative forces and how they lead to thermal energy, the mechanism by which this occurred was rather mysterious. Perhaps now with the insight we have into work done by a force resulting in both mechanical energy change and internal energy change, we can get a better sense of how all this plays out.

First of all, it should be clear that thermal energy is a form of internal energy. The only additional property thermal energy requires is that it involves a random distribution among the particles in the system. With this criterion, one can hardly consider the internal energy of the two-particle example above to be "thermal," while it's clear that we have no choice but to treat the shared internal energy of trillions of particles in that manner. Still, there is nothing fundamentally different between these two cases. What makes thermal energy so interesting is that while we can't "see inside the box" to follow the intricate details of the internal energy, we do have a way to measure it – through the temperature.

We see from the example above that in order for work done on a system to contribute to its internal energy, the force acting on the system must accelerate various particles in the system differently. In our two-particle example, internal energy arose because the force acted on only one of the two particles. It is nevertheless possible for work done on a system can go purely into mechanical energy (i.e. no thermal or other internal energy created) in one of two ways:

- The system of particles is a solid, rigid, object, so that any force on one part of the system accelerates every particle in the system in precisely the same way. (We will see an important exception to this in Chapter 5.)
- The force applied to the system acts on every particle in proportion to its mass, so that even though the particles are not rigidly bound to each other, they all accelerate the same.

For the vast majority of situations in classical mechanics, we resort to the first of these criteria in cases of conservative forces: A solid mass is pushed by a spring; a solid magnet is repulsed by another magnet, and so on. It is possible, however, to avoid a change in internal energy for a system of particles that are not rigidly held together, such as gases and liquids, if the second criterion is met. The latter criterion is clearly satisfied by gravity. Sure enough, we know that work done by gravity changes only the mechanical energy of a system, and not its thermal energy (it is "conservative").

Digression: Gravitation Contributes to Thermal Energy

It is well-known that the moon of Jupiter named "Io" exhibits extensive volcanic activity. It is believed that the source of the thermal energy is an imbalance in the gravitational forces on different parts of the moon. This imbalance comes from Io's gravitational interaction with its sister moons Europa, Ganymede, and Callisto, along with its primary gravitational interaction with Jupiter.

Io, with two plumes erupting from its surface. NASA

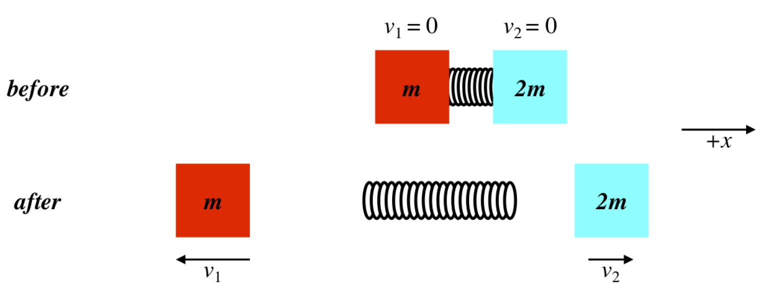
Forces other than gravity that act on a gas, however, will have an effect on the internal energy of the system. For example, if we compress a gas in a confined space, the piston doing the compressing only does work on the particles with which it comes in contact. After the piston is done moving, the center of mass of the gas comes back to rest, which means the piston added nothing to the gas system's mechanical energy. But clearly work was done, and this goes into the internal energy of the gas, the change of which we can measure with an increase in temperature.

Consider next the case of kinetic friction. We know that the mechanism for this force involves microscopic interactions of irregularities in the two surfaces involved. This means that the particles only at the surfaces are exerting forces on each other – all the particles in the objects that are sliding are not involved equally. We cannot resort to saying that the objects involved are systems of particles held rigidly together, or they would stop moving immediately when surface irregularities interacted. These irregularities must be capable of some deformation for any sliding to occur. For static friction, the particles can be considered to be held rigidly in place, and therefore no thermal energy is generated from static friction as there is with kinetic friction.

Kinetic Energy Distribution Within a System

Let's return once again to an example we looked at in the previous section (Figure 4.3.1), and ask a new question about it (the example has been simplified slightly by giving one block exactly twice the mass of the second block).

Figure 4.4.3 – Kinetic Energy Distribution for Repelling Blocks



The spring stored some potential energy when it was compressed, and it gave this energy to the kinetic energy of the two blocks. What fraction of this energy is given to each of the blocks? One might be inclined to believe that since the spring exerts equal forces on both blocks, they both get equal amounts of kinetic energy. But by now we know better! They only get the same amount of energy if the spring does the same amount of work on both, and it's clear here that the lighter mass is pushed a longer distance before losing contact with the spring than the heavier mass, so with equal forces acting on each, more work is done on the lighter mass. Specifically, the lighter mass is accelerated twice as much by the equal force, so it displaces twice as far, and therefore gets twice as much energy as the heavier block.

Another way to see it is to note that both blocks must have the same magnitude of momentum after the spring expands (since the momenta must cancel to equal zero and remain conserved), so using Equation 4.1.5 we can compare their kinetic energies:

$$KE_1 = \frac{p_1^2}{2m_1} = \frac{p^2}{2m} \quad \Rightarrow \quad KE_2 = \frac{p_2^2}{2m_2} = \frac{p^2}{2\left(2m\right)} = \frac{1}{2}KE_1 \tag{4.4.10}$$

This confirms what we reasoned above.

Now we can see more clearly why we are able to refer to the gravitational potential energy of the system of a small object and the earth as simply the gravitational potential energy "of the object," ignoring the fact that the earth is also involved. This is because when the potential energy is converted to kinetic energy, virtually all of the kinetic energy goes to the object, and none of it to the earth (imagine the heavier block above being *much* heavier).

4.5: Collisions

One-Dimensional Collisions

We know that in the case of a collision, the force acting between the two objects is irrelevant to momentum conservation, but is very important to determining the amount of energy converted to thermal energy. For example, if two blocks collide with a spring between them, then all the kinetic energy they come in with, they will also go out with, since the internal spring force involved in the collision is conservative, and there is no spring potential energy before or after. On the other hand, if an internal non-conservative force is present between the colliding objects, then some of the incoming kinetic energy is converted into thermal energy. The former sort of collision (where kinetic energy is conserved) we call *elastic*, and the second type of collision we call *inelastic*.

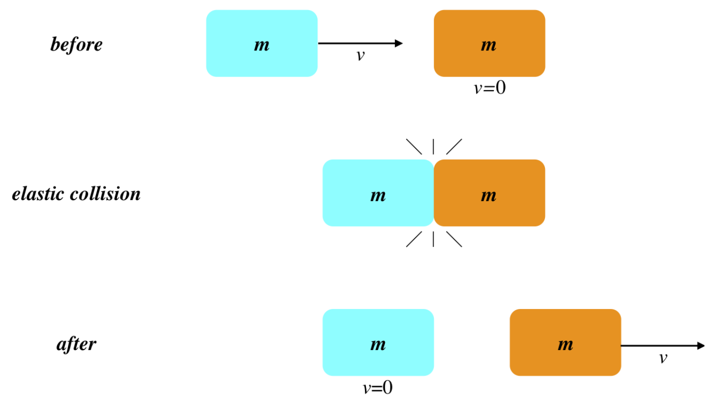
From our discussion in Section 4.4, it's clear that what determines the inelasticity of a collision is the deformation of the colliding objects. When a colliding object deforms, it's because the particles directly involved in the contact are accelerated more than other particles in the same object, thus introducing internal energy, and reducing the amount of mechanical energy available to go back into the motions of the objects.

A collision where the objects continue together with the same velocity after the collision (i.e. they remain stuck together), is often referred to as *totally* or *perfectly inelastic*. This of course does not mean that all of the kinetic energy is lost (the objects do continue moving at the end in most such collisions), only that they don't bounce off each other. From the perspective of the center of mass frame, we can see that such a collision maximizes the amount of internal energy that the collision can create: In this frame, the objects stop entirely after the collision, so all of the mechanical energy becomes internal. Changing frames doesn't change the amount of internal energy created (it only changes the mechanical energy we see), so having the objects stick together results in the largest possible creation of internal energy.

Elastic Collisions

If we are told that a given collision is elastic (or at least can be approximated as such), then that gives us an additional condition that we can use to solve the problem. Let’s see a couple of examples. in each case, the diagram will show the experimental result, which we will then show mathematically using the combination of momentum and kinetic energy conservation.

Figure 4.5.1 – Elastic Collision of Equal Masses, Target Stationary

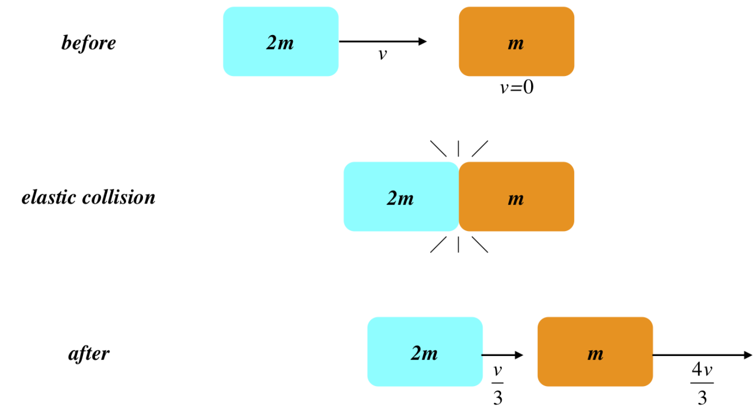


We see that the incoming cart stops completely and the target cart moves off with the same velocity as the original cart (note that the center of mass continues moving at a constant speed, as it should). We now show this mathematically... Dropping the vector arrows, since the motion is in one dimension, and choosing to the right as the (+) direction, we have:

$$\left.\begin{array}{lll} momentum\;conservation: & mv+0=mv_1+mv_2 & \Rightarrow \quad v=v_1+v_2 \\ elastic\;collision & \frac{1}{2}mv^2+0=\frac{1}{2}mv_1^2+\frac{1}{2}mv_2^2 & \Rightarrow \quad v^2=v_1^2+v_2^2 \end{array}\right\} \Rightarrow \quad v_1=0,\;\; v_2=v \quad or \quad v_1=v,\;\; v_2=0 \tag{4.5.1}$$

Wait, why do we get two solutions? That is, why can *either* velocity equal zero? Well, if the incoming cart were to *miss the target cart*, then that too is an elastic “collision,” inasmuch as the momentum and kinetic are both conserved, so the math takes into account that as a possibility.

Figure 4.5.2 – Elastic Collision of Unequal Masses, Target Lighter and Stationary



The algebra is only a little tougher this time:

$$\left.\begin{array}{lll} momentum\ conservation: & 2mv+0=2mv_1+mv_2 & \Rightarrow\ 2v=2v_1+v_2 \\ elastic\ collision & \frac{1}{2}2mv^2+0=\frac{1}{2}2mv_1^2+\frac{1}{2}mv_2^2 & \Rightarrow\ 2v^2=2v_1^2+v_2^2 \end{array}\right\} \Rightarrow\ 4v_1=v_2 \ \Rightarrow\ v_1=\frac{v}{3},\ \ v_2=\frac{4v}{3} \tag{4.5.2}$$

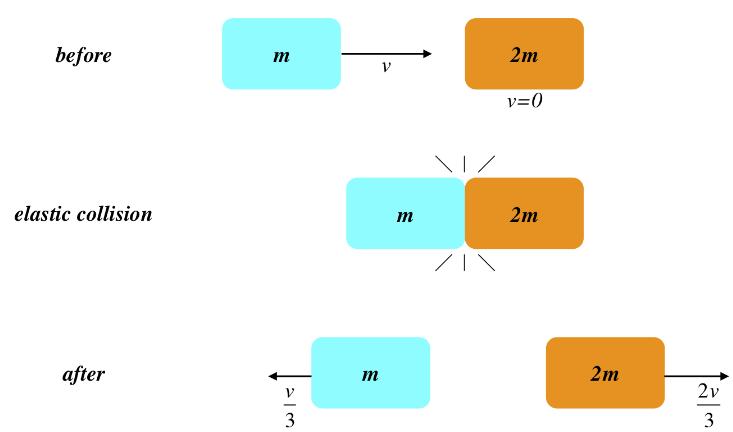
Both carts continue forward, the lighter one at 4 times the speed of the heavier one. Note that once again $v_1=v,\ \ v_2=0$ is a solution (the incoming cart misses the target).

A clear application of this principle comes in bowling. Clearly we want the bowling ball to have more mass than a pin, so that it can carry through to the pins behind the front pin(s). If we consider collisions in two dimensions (which we will do later), we will find that the angular deflection of the ball when it doesn't strike the pin head-on will be less when the ball is heavier, which is one reason heavier bowling balls are more effective than lighter ones.

As a second example of this, suppose we are passengers in one of two vehicles involved in a head-on collision. Which vehicle would we rather be in, the lighter one or the heavier one? Intuitively we know we would rather be in the heavier vehicle, but why? Well, we would want to experience as little force as possible (force is what breaks bones). The force that our dashboard or steering column exerts on us is going to equal our mass times our acceleration (as it constitutes our net horizontal force), and we are constrained to experience the same acceleration as our car. So compare the accelerations of the two carts here. The heavier cart goes from a speed v down to a speed of v/3, for a change of 2v/3. The lighter cart’s velocity changes from 0 to 4v/3 in the same period of time, which means it experiences twice the acceleration. More acceleration for our car means more acceleration for us, which means more force on us, which is bad.

Lastly, we look at the lighter object bouncing off the heavier one:

Figure 4.5.3 – Elastic Collision of Unequal Masses, Target Heavier and Stationary



The math:

$$\left.\begin{array}{lll} momentum\ conservation: & mv+0=mv_1+2mv_2 & \Rightarrow\ v=v_1+2v_2 \\ elastic\ collision & \frac{1}{2}mv^2+0=\frac{1}{2}mv_1^2+\frac{1}{2}2mv_2^2 & \Rightarrow\ v^2=v_1^2+2v_2^2 \end{array}\right\} \Rightarrow\ 2v_1=-v_2 \ \Rightarrow\ v_1=-\frac{v}{3},\ \ v_2=\frac{2v}{3} \tag{4.5.3}$$

The lighter cart bounces off the heavier one at half the speed that the heavier one continues forward (or the incoming cart misses the target). There is actually a clever way we could have solved this case more quickly by using the solution of the previous case and what we know about relative motion. If we move along with the incoming block and declare ourselves to be "stationary," then we see the heavier mass coming toward us at a speed v, which is exactly the same physical situation as we had above. After the collision, we will see the heavier mass continuing in the same direction at a speed of $v/3$, while the target block moves in the same direction at a speed of $4v/3$. That is what *we see*. Going back to the original frame, these two speeds change by v, which means the heavy object is not going left at $v/3$ – it is going *right* at $v-v/3=2v/3$, while the smaller block is moving left at a speed of $4v/3-v=v/3$.

Inelastic Collisions

Okay, so let's look at some inelastic collisions. As we said above, these can include some component of an elastic collision, where the objects bounce off each other or they can be totally inelastic, where they stick together and continue as one mass. In both cases, some of the kinetic energy contained in the system is converted to thermal energy. [*Note: From this point on, "kinetic energy" will refer to mechanical kinetic energy, and all internal energy will be called thermal. We will no longer talk about "internal kinetic energy."*] This can be expressed in a couple of different ways:

$$KE_{after}=KE_{before}-\Delta E_{thermal}\,,\quad or\quad KE_{after}=(x\%)\,KE_{before} \tag{4.5.4}$$

Of the two types of inelastic collisions, the totally inelastic case is easier to handle, because we we don't require the information of what fraction of the internal force was conservative and what fraction was non-conservative. We can see the simplicity of the totally inelastic case comes from the fact that all of the information about the resulting motion is contained in just the momentum conservation equation, because both objects have the same final velocity:

$$m_1v_1+m_2v_2=(m_1+m_2)\,v_f\quad\Rightarrow\quad v_f=\frac{m_1v_1+m_2v_2}{m_1+m_2} \tag{4.5.5}$$

Note that (naturally) the final speed of the combined masses is the speed of the center of mass, since everything is moving together. The kinetic energy converted to thermal energy can be computed from this result in terms of the incoming velocities:

$$-\Delta E_{thermal}=KE_f-KE_o=\frac{1}{2}(m_1+m_2)\,v_f^2-\left(\frac{1}{2}m_1v_1^2+\frac{1}{2}m_2v_2^2\right), \tag{4.5.6}$$

where v_f is given in terms of v_1 and v_2 in Equation 4.5.5.

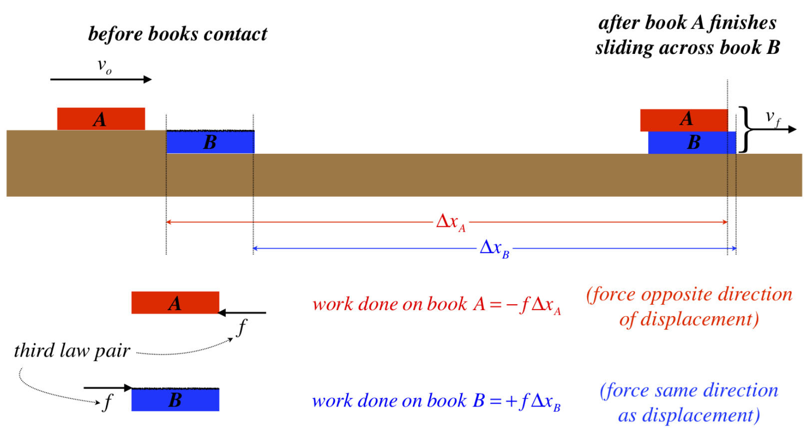
If the perfectly inelastic head-on collision involves an incoming object with mass m_1 and a stationary target with mass m_2, it is easy to use momentum conservation and Equation 4.1.5 to derive a simple relationship between the starting and ending kinetic energy of the system:

$$\frac{KE_{after}}{KE_{before}}=\frac{\dfrac{\cancel{p^2_{after}}}{2\,(m_1+m_2)}}{\dfrac{\cancel{p^2_{before}}}{2m_1}}=\frac{m_1}{m_1+m_2} \tag{4.5.7}$$

This certainly makes sense from the perspective of dropping a pile of mashed potatoes on the floor. The earth is the stationary target with a very large mass, and after the perfectly inelastic collision, pretty much all of the kinetic energy in the mashed potatoes becomes thermal. Of course, you may prefer to warm your mashed potatoes in a microwave oven instead.

Notice that nowhere here do we mention the details of the force acting between the two masses. This is strange, because we have seen that the specifics of the force determine the work done, which in turn determines the amount of energy converted to thermal. Let's return to our two book example one more time:

Figure 4.5.4 – Work Done in Perfectly Inelastic Collision



Suppose we change the friction force by changing the coefficient of kinetic friction. In fact, let's assume we double that constant – what will happen? Well, the friction force will double, since the normal force is the same as before. From the reference frame of book B, this means the acceleration of book A has doubled. With double the acceleration and the same starting speed, in book B's reference frame it will take half as long for book A to come to rest. Therefore the time that both books experience double the acceleration is half as long as before. Putting this into the equations of motion for the two books gives:

$$\left.\begin{array}{l} book\ A: \ \Delta x_A = v_o t - \frac{1}{2}at^2 \\ book\ B: \ \Delta x_B = \frac{1}{2}at^2 \end{array}\right\} t \rightarrow \frac{t}{2},\ a \rightarrow 2a \Rightarrow \Delta x_A \rightarrow \frac{\Delta x_A}{2},\ \Delta x_B \rightarrow \frac{\Delta x_B}{2} \tag{4.5.8}$$

So both books travel half as far when the coefficient of friction is doubled. The friction force is twice as great, so the work done on each is the same as before! This shows that the same amount of energy is lost in both cases, and in general it will turn out to be the same no matter what sort of non-conservative force is acting, so long as the collision is perfectly inelastic.

Another way to see this is to use our favorite trick of going into the center of mass reference frame. In this frame, the two books are moving toward each other, and after they slide across each other, they come to a stop. This means that all of the initial kinetic energy becomes thermal. But notice that this is true no matter what force was acting between them! Different forces (like changing the coefficient of friction) will lead to different *rates* at which the kinetic energy becomes thermal, but all of it converts in the end.

Example 4.5.1

A large sled of mass 12.0kg is at rest on a horizontal, frictionless sheet of ice, when a heavy rock with a mass of 7.50kg is thrown onto it from behind. The rock is moving purely horizontally at a speed of 2.40m/s when it comes into contact with the sled, and it skids across the rough top surface of the sled until it and the sled are moving forward together at the same speed. The scrape marks on the sled indicate that the rock skidded across it a distance of 1.60m.

a. *Find the final speed of the sled.*
b. *Find the magnitude of the friction force between the sled and the rock.*

Solution

a. The only horizontal force present is the friction force between the rock and the sled, which is internal to the rock+sled system, so the system's horizontal momentum is conserved. We therefore have:

$$\underbrace{m_{rock}v_{rock} + m_{sled}v_{sled}}_{p_{before}} = \underbrace{(m_{rock} + m_{sled})V}_{p_{after}} \quad \Rightarrow \quad V = \frac{m_{rock}v_{rock}}{m_{rock} + m_{sled}} = \boxed{0.92\frac{m}{s}}$$

b. The work done by friction equals the energy converted to thermal, which is the energy lost from kinetic. We therefore compute the kinetic energy lost first. We can do this the long way, but because this is the special case of a stationary target and a perfectly inelastic collision, we can use a shortcut to get the energy lost, using Equation 4.5.7:

$$\frac{KE_{after}}{KE_{before}} = \frac{m_1}{m_1 + m_2} = 0.385 \quad \Rightarrow \quad \Delta E_{thermal} = -\Delta KE = KE_{before} - KE_{after} = 0.615 KE_{before} = 0.615\left(\frac{1}{2}m_{rock}v_{rock}^2\right) = 13.3J$$

The distance the rock scrapes across the sled is given, so since the work done during this scraping is the energy converted to thermal, we can compute the friction force:

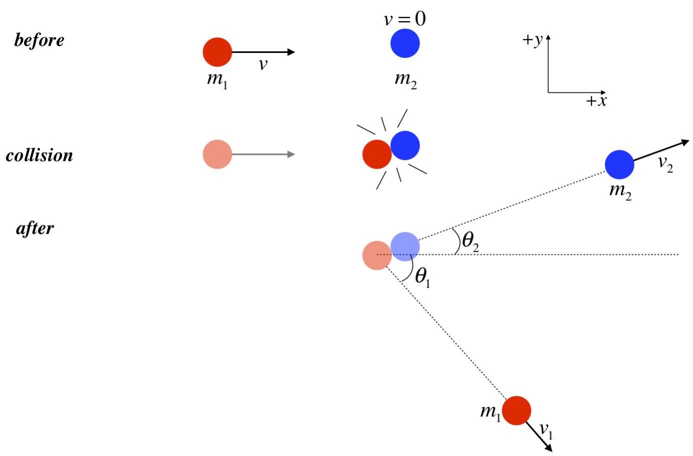
$$\Delta E_{thermal} = f \cdot \Delta x \quad \Rightarrow \quad f = \boxed{8.1N}$$

General Two-Dimensional Collisions

We have been saying for awhile now that one of the big differences between momentum conservation and energy conservation is the fact that momentum is a vector while energy is not. This means that there are actually three momentum quantities that are equal before and after (if the full momentum vector is conserved). Here we will look at what this entails.

Let's look at a standard two-dimensional collision. In this example, we will have a stationary ball struck by another. The two balls have different masses, and they collide off-center, so that they emerge from the collision in directions angled off the original direction of motion. We'll set up the geometry and label all the known and unknown variables with a diagram, and then do the physics:

Figure 4.5.5 – General Two-Dimensional Collision in the Target Frame



Now we need to apply momentum conservation. Since momentum is a conserved vector, each of its components are individually conserved, which means that momentum conservation provides us two separate equations to work with. In the "before" case, we have an x-component of momentum that is simply the incoming mass times the incoming velocity ($m_1 v$), while the y-component of momentum is zero. in the "after" case, we need to resolve the momenta into components. Setting before equal to after gives:

$$\begin{array}{ll} x-direction: & m_1 v = m_1 v_1 \cos\theta_1 + m_2 v_2 \cos\theta_2 \\ y-direction: & 0 = -m_1 v_1 \sin\theta_1 + m_2 v_2 \sin\theta_2 \end{array} \tag{4.5.9}$$

You'll note the minus sign for the component in the $-y$-direction. This is not strictly necessary, as this negative sign could be absorbed into θ_1, but it is generally less confusing to put the signs in explicitly, and let all the angle values be positive.

Let's consider what would be required to solve a problem that looks like this. We have two equations, and seven distinct variables. If this is all we know about the collision, then to completely unravel this physical situation, we need to know five of these quantities. So for example, we could be given the two masses, the incoming speed, and the outgoing speed and direction of one of the balls, and we can solve for the outgoing speed and direction of the other ball. If we also provided the target ball a starting velocity, or a y-component to the incoming ball's velocity, then there would be even more unknowns. But we can quickly reduce this problem back to the one above, by first rotating our coordinate system so that the incoming velocity is once again in the x direction, and then changing the reference frame to the rest frame of the target ball. It is also sometimes useful to change to the center of mass reference frame.

Notice that once such a problem is solved, one can then check to see if the collision is elastic, by comparing the kinetic energy before and after the collision:

$$KE_{before} = \frac{1}{2} m_1 v^2 \qquad KE_{after} = \frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2 \tag{4.5.10}$$

This comparison could be a difference (determining how much kinetic energy is lost), or a fraction (determining the percentage of kinetic energy remaining or the percentage lost). Note that a collision *can* result in an increase of kinetic energy, but this can only happen if there is some potential energy stored within the colliding objects that is unleashed by the collision. This is such an uncommon occurrence (the circumstances need to be quite contrived), that it is safe to assume that a collision is either elastic (conserves kinetic energy) or is inelastic such that kinetic energy is lost.

Not all problems are posed with five of the seven variables given. The energy condition can be given instead, which provides a third equation, requiring only four of the seven variables in the statement of the problem. Needless to say, these problems can require a lot of tedious algebra, but getting the equations set up using momentum conservation and the fate of the system's kinetic energy is where the physics is.

Elastic Two-Dimensional Collisions

As daunting as the full-blown problem shown above can be, there are cases where shortcuts or simplifications exist. We look first at the case of elastic collisions. If we want to know all the information shown above, we have no choice but to go through the algebra involved. But we can achieve an interesting result without recourse to the coordinate system at all. Namely, it turns out that the ratios of the masses of the colliding objects and their outgoing speeds completely determine the angle *between* the outgoing velocity vectors, $\theta_1 + \theta_2$. To get this result, we will use Equation 4.1.5 extensively...

Let's call the incoming momentum $\overrightarrow{p}$ and the mass of the incoming object m_1. Then the kinetic energy of the system (in the frame where the target is stationary) is:

$$KE_{before} = \frac{1}{2m_1} \overrightarrow{p} \cdot \overrightarrow{p} \tag{4.5.11}$$

Now let's define the outgoing momenta of the two objects as $\overrightarrow{p}_1$ and $\overrightarrow{p}_2$, with the latter being for the target object after collision. The kinetic energy after the collision is therefore:

$$KE_{after} = \frac{1}{2m_1}\overrightarrow{p}_1 \cdot \overrightarrow{p}_1 + \frac{1}{2m_2}\overrightarrow{p}_2 \cdot \overrightarrow{p}_2 \tag{4.5.12}$$

Now we apply momentum conservation:

$$\overrightarrow{p} = \overrightarrow{p}_1 + \overrightarrow{p}_2 \quad \Rightarrow \quad KE_{before} = \frac{1}{2m_1}\left(\overrightarrow{p}_1 + \overrightarrow{p}_2\right) \cdot \left(\overrightarrow{p}_1 + \overrightarrow{p}_2\right) = \frac{1}{2m_1}\overrightarrow{p}_1 \cdot \overrightarrow{p}_1 + \frac{1}{2m_1}\overrightarrow{p}_2 \cdot \overrightarrow{p}_2 + \frac{1}{m_1}\overrightarrow{p}_1 \cdot \overrightarrow{p}_2 \tag{4.5.13}$$

Applying kinetic energy conservation (remember, we are assuming an elastic collision):

$$KE_{before} = KE_{after} \quad \Rightarrow \quad \cancel{\frac{1}{2m_1}\overrightarrow{p}_1 \cdot \overrightarrow{p}_1} + \frac{1}{2m_1}\overrightarrow{p}_2 \cdot \overrightarrow{p}_2 + \frac{1}{m_1}\overrightarrow{p}_1 \cdot \overrightarrow{p}_2 = \cancel{\frac{1}{2m_1}\overrightarrow{p}_1 \cdot \overrightarrow{p}_1} + \frac{1}{2m_2}\overrightarrow{p}_2 \cdot \overrightarrow{p}_2 \tag{4.5.14}$$

Now multiply through by m_1 and rearrange things a bit to get:

$$\overrightarrow{p}_1 \cdot \overrightarrow{p}_2 = \frac{1}{2}\left(\frac{m_1}{m_2} - 1\right)\overrightarrow{p}_2 \cdot \overrightarrow{p}_2 \tag{4.5.15}$$

Now write the dot products in terms of the magnitudes of the vectors and the angles between them:

$$p_1 p_2 \cos\theta = \frac{1}{2}\left(\frac{m_1}{m_2} - 1\right)p_2^2 \tag{4.5.16}$$

The angle θ is of course the angle between the two outgoing velocity vectors (which point the same direction as the momentum vectors). The $p_2 = 0$ solution to this corresponds to the case of the incoming object missing the target entirely (because the target remains stationary), so assuming the target is not missed, we can divide both sides by p_2 and if we also plug in $p_1 = m_1 v_1$ and $p_2 = m_2 v_2$, we get the promised relationship of the *scattering angle* in terms of the masses and outgoing speeds:

$$\theta = \cos^{-1}\left[\frac{1}{2}\left(1 - \frac{m_2}{m_1}\right)\frac{v_2}{v_1}\right] \tag{4.5.17}$$

We can extract some interesting information from this result:

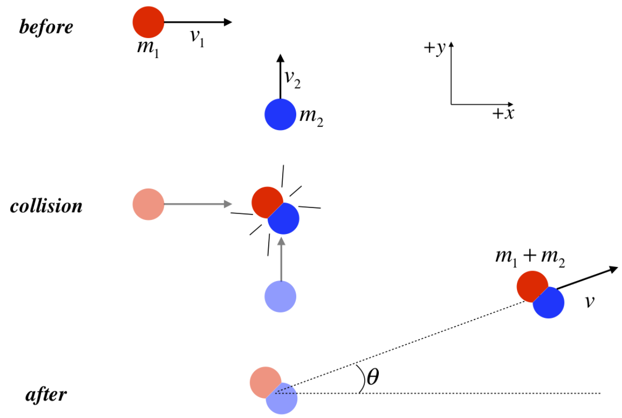
- We see that if the masses are equal, then the scattering angle is precisely 90^o, since the cosine of this angle vanishes. In this case, the scattering angle doesn't depend at all on how off-center the collision is (except that a direct head-on hit naturally leads to an angle of 0^o or 180^o). The degree of how off-center the collision is (which is measured by a quantity known as the *impact parameter*) *does* effect the angles θ_1 and θ_2 in Figure 4.5.5, but not the sum of those angles. If the masses are not equal, then the impact parameter does play a role in the scattering angle, because it has a say in the ratio of the outgoing speeds.
- If $m_2 > m_1$, the argument of the inverse cosine is negative, so the angle must be greater than 90^o. This makes sense, because if the target mass is greater than the incoming mass, the incoming mass "bounces back," rather than "plowing through" (a result we found for the one-dimensional elastic collisions we examined above), and since the target mass has a forward component to its final velocity, the angle is greater than 90^o.
- The argument of the inverse cosine can never be larger than +1 or smaller than –1, which places limits on the outgoing speeds given the masses. For example, if the incoming mass m_1 is twice the target mass m_2, then the largest possible ratio of the two outgoing velocities is 4. This ratio occurs when $\theta = 0$, and indeed we have seen this result already above (Equation 4.5.2).

It should be noted that this result could also be achieved using the formulas resulting from Figure 4.5.5, but it would require an unnatural desire to slog through trigonometric identities.

Perfectly Inelastic Two-Dimensional Collisions

As much as we were able to do with elastic collisions, perfectly inelastic collisions are even easier to handle. This is because the outgoing motions of the two objects are constrained to be the same (i.e. they stick together and have the same final speed and direction). This constraint means that if we are given all of the incoming conditions (the masses of the two objects, and their incoming velocity vectors), we can determine the result completely. That is, the amount of energy lost in the collision does not need to be given – it is unique and can in fact be computed. Figure 4.5.6 is a diagram for an example of a perfectly inelastic collision. [*This is somewhat simplified by having the incoming objects approach each other at right angles, but not as simple as the case of looking at it from the target frame, which makes the collision one-dimensional!*]

Figure 4.5.6 – A Perfectly Inelastic Two-Dimensional Collision



We follow the same procedure as we did for Figure 4.5.5, this time with the simplification that we have a single outgoing momentum:

$$\begin{aligned} x-direction: \quad & m_1 v_1 = (m_1+m_2)\,v\cos\theta \\ y-direction: \quad & m_2 v_2 = (m_1+m_2)\,v\sin\theta \end{aligned} \tag{4.5.18}$$

The amount of energy converted to thermal from this collision equals the loss of kinetic energy from the system, and as we saw in the one-dimensional case, this amount doesn't depend upon the details of the internal non-conservative force. It only matters that eventually (after the two objects end their tumultuous collision) settle into moving off together with a common velocity. The amount of energy converted is:

$$\Delta E_{thermal} = KE_{before} - KE_{after} = \left[\frac{1}{2}m_1 v_1^2 + \frac{1}{2}m_2 v_2^2\right] - \left[\frac{1}{2}(m_1+m_2)\,v^2\right] \tag{4.5.19}$$

For the case above where the two incoming objects have velocities are right angles to each other, we can turn this into an equation that includes only the masses and incoming speeds. Sparing the reader the algebra, the result is:

$$E_{thermal} = \frac{1}{2}\left(\frac{m_1 m_2}{m_1+m_2}\right)\left(v_1^2+v_2^2\right) \tag{4.5.20}$$

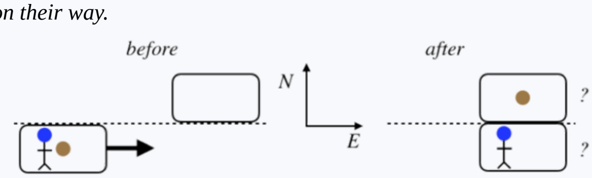
Notice that since the two velocities are perpendicular, the sum of their squares is actually the square of their *relative* velocity. This is not a surprising result, and in fact will translate into collisions at any angle (though the equation will look different), because we would not expect the post-collision blob to be any hotter when the collision is viewed in one frame as opposed to another. As mentioned above, we can always view this collision from the target frame, making the collision one-dimensional, and the total kinetic energy of the system before the collision is a function of the relative velocity. In that case, we can use Equation 4.5.7 to compute the energy converted to thermal:

$$\Delta E_{thermal} = KE_{before} - KE_{after} = KE_{before} - \left(\frac{m_1}{m_1+m_2}\right)KE_{before} = \left(\frac{m_2}{m_1+m_2}\right)KE_{before} \tag{4.5.21}$$

So suppose we drop a ball of clay to the ground. Viewing this from the earth's rest frame, the earth becomes the stationary target with mass m_2, and essentially all of the clay's incoming kinetic energy is converted to thermal (because $m_2 \approx m_1 + m_2$), and the clay's (and earth's) temperature goes up a bit. If we view it from the clay's rest frame, then the kinetic energy of the earth is enormous (same relative speed, much larger mass), and after the collision we might therefore expect the temperatures to go up a lot, but making the clay the stationary target now makes the target mass m_2 very small compared to $m_1 + m_2$, which makes the fraction multiplying the earth's kinetic energy very small – exactly small enough to give the same energy change as before.

Example 4.5.1

A cart slides along a frictionless surface in an easterly direction at a speed of 2.40m/s. The cart has a mass of 100kg, and it contains a person who has a mass of 60.0kg, as well as a medicine ball that has a mass of 15.0kg. The cart slides past an identical (but empty) stationary cart, also on the frictionless surface. When the carts are side-by-side, the person throws the medicine ball into the other cart by pushing the ball in the north direction. At the moment of release, the person sees the ball is moving away from their hands at a speed of 3.20m/s. The ball comes to rest inside the other cart, and both carts continue on their way.



Find the speed and direction of both carts after the medicine ball has been exchanged. Express the directions as angles that are north or south (indicate which) of east.

Solution

Treating the two carts, the person, and the medicine ball as a single system, we know that all the forces on them are internal. We have some additional information to use. First, we are told that the medicine ball is pushed north. This will affect the north-south component of the momentum of the ball and the cart it lands in. The equal-and-opposite push of the ball on the person will affect the north-south component of momentum of the person and the cart they are in. But the east-west components of momentum are unaffected. We therefore know that the person + cart continues with the same eastward component of velocity as before, and the eastward momentum of the ball gets transferred to the ball + cart system, giving the final eastward component of the ball + cart:

$$\begin{array}{lcccc} & before & & after & \\ east-west: & m_bv_o+m_c\left(0\right) & = & \left(m_b+m_c\right)v_{bc(east)} & \Rightarrow \quad v_{bc(east)}=\dfrac{m_bv_o}{m_b+m_c}=0.313\,\tfrac{m}{s}\end{array}$$

Now for the north-south components. Initially nothing has a north-south component of momentum, so the final total north-south component of momentum must also be zero for the whole system. We are given the relative velocity of the ball and the person as it leaves their hands, but we are working in the earth frame, so we have to be careful. Calling the northward component of the ball's velocity relative to the earth v_b, and the southward component of the person + cart after the ball is released v_{pc} (which has a negative value), we can relate these through their relative motion by:

$$v_{b(north)}=3.20\frac{m}{s}+v_{pc(south)}$$

Now apply momentum conservation (in the earth frame) along the north-south direction to get the north-south component of the velocity of the person + cart:

$$\begin{array}{lccccc} & before & & after & & \\ north-south: & 0 & = & m_bv_{b(north)}+\left(m_p+m_c\right)v_{pc(south)} & = & m_b\left(3.20\tfrac{m}{s}+v_{pc(south)}\right)+\left(m_p+m_c\right)v_{pc(south)} \\ & & & \Rightarrow \quad v_{pc(south)}=-\dfrac{m_b\left(3.20\tfrac{m}{s}\right)}{m_b+m_p+m_c}=-0.274\dfrac{m}{s} & & \end{array}$$

This southward velocity of the person + cart gives us its final magnitude and direction of motion:

$$v_{pc}=\sqrt{v_o^2+v_{pc(south)}^2}=\boxed{2.42\frac{m}{s}}$$

$$\theta_{pc}=\tan^{-1}\left(\frac{v_{pc(south)}}{v_o}\right)=\boxed{6.51^o\;\;south\;of\;east}$$

The northward component of the momentum of the ball + cart is the same as the southward momentum of the person + cart, since the whole system's north-south momentum remains zero. We therefore have for the ball + cart:

$$0=\left(m_b+m_c\right)v_{bc(north)}+\left(m_p+m_c\right)v_{pc(south)} \quad \Rightarrow \quad v_{bc(north)}=-\left(\frac{m_p+m_c}{m_b+m_c}\right)v_{pc(south)}=0.381\frac{m}{s}$$

Now with both components of the velocity of the ball + cart, we complete the solution:

$$v_{bc}=\sqrt{v_{bc(east)}^2+v_{bc(north)}^2}=\boxed{0.493\frac{m}{s}}$$

$$\theta_{bc}=\tan^{-1}\left(\frac{v_{bc(north)}}{v_{bc(east)}}\right)=\boxed{50.1^o\;\;north\;of\;east}$$

4.6: Problem Solving

In this chapter we consider three classic problems in momentum and energy conservation.

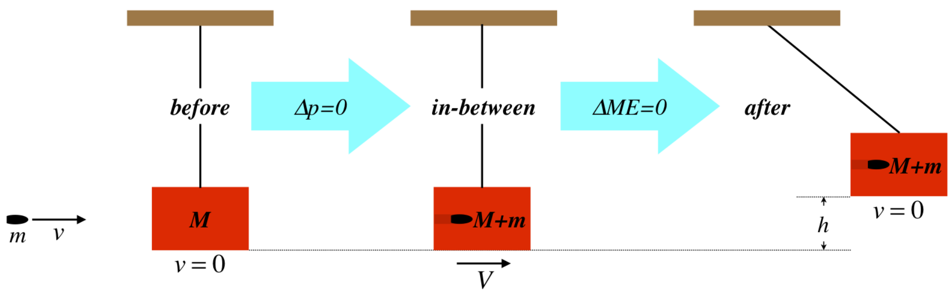
The Ballistic Pendulum

A bullet is fired into a block of wood that is hanging by a string from the ceiling. The mass of the bullet and the block are given, as is the incoming speed of the bullet. Find the height to which the block + bullet swings before stopping.

This problem involves both momentum and energy, but it requires great care that we don't invoke energy conservation when non-conservative forces are acting, and that we don't invoke momentum conservation when external forces are acting. Basically the trick is to dance around the problem only using the right tool for the job. The bullet colliding with the wood block involves a non-conservative force (the friction that slows the bullet within the block), so the mechanical energy of the bullet + block system is not conserved. But the force between the bullet and the block is internal to the bullet + block system, so momentum is conserved (we assume that the bullet digs-in and lodges so quickly that the block doesn't yet swing very far, so the external force from the string has no horizontal component during the collision). So we use momentum conservation to determine the speed of the block + bullet as they begin to swing. Then once it is swinging, the tension force from the string kills momentum conservation, but it does no work (it acts perpendicular to the direction of motion), so mechanical energy is conserved, and we use that to find the height.

Figure 4.6.1 – Ballistic Pendulum

Figure 4.5.1 – The Ballistic Pendulum



Applying momentum conservation for the first stage gives the following:

$$mv+0=(M+m)V \quad \Rightarrow \quad V=\frac{m}{M+m}v \tag{4.6.1}$$

Mechanical energy conservation for the second stage yields:

$$\frac{1}{2}\cancel{(M+m)}V^2 = \cancel{(M+m)}gh \quad \Rightarrow \quad h=\frac{V^2}{2g} \tag{4.6.2}$$

Plugging V from Equation 4.6.1 into Equation 4.6.2 gives an answer:

$$h=\frac{1}{2g}\left(\frac{m}{M+m}\right)^2 v^2 \tag{4.6.3}$$

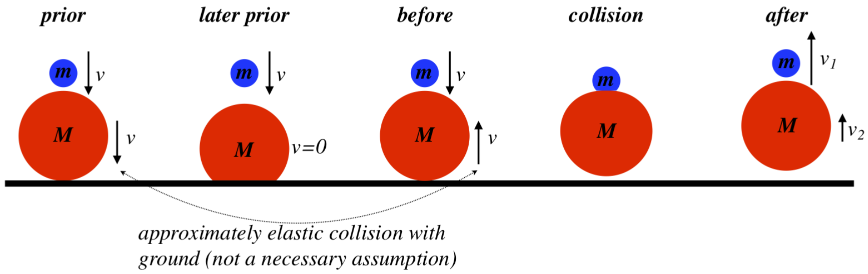
Notice that using conservation of energy from the beginning will not give the right answer – it *must* be broken into two parts, because we can only use the appropriate physical principles when they are applicable.

Stacked Balls

Two balls are dropped to the floor, with the lighter ball atop the heavier one. The balls collide approximately elastically with each other and with the floor. We see a reaction we are not used to – the small ball flies up to a height higher than it was dropped. Does this violate mechanical energy conservation?

Figure 4.6.2 – Stacked Balls

Figure 4.5.2 – The Stacked Ball Launch



As with the previous problem, we need to break this up into workable parts. If we assume an elastic collision with the floor (this is not necessary, but it makes it easier to compare the initial and final heights of the smaller ball), then the large ball leaves the floor at the same speed at which it struck the floor (which we are calling v). The balls were dropped together, so the small ball is moving at the same speed down as the large ball is moving up, and a collision occurs (depicted in the "before" portion of Figure 4.6.2). The momentum and kinetic energy conservation of this collision allows us to solve for the final velocities of the two balls:

$$\begin{array}{ll} \textit{momentum conservation}: & Mv - mv = mv_1 + Mv_2 \\ \textit{elastic collision}: & \frac{1}{2}mv^2 + \frac{1}{2}Mv^2 = \frac{1}{2}mv_1^2 + \frac{1}{2}Mv_2^2 \end{array} \tag{4.6.4}$$

We can now solve for v_1 and v_2 in terms of v, which we can determine from the height that the balls are dropped from. If v_1 comes out to be greater than v, then the small ball will rise to a height greater than that from which it is dropped. The algebra required to solve for v_1 and v_2 is rather daunting and too long to provide here, but here is the result:

$$v_1 = \left(\frac{3M - m}{M + m}\right) v\,, \qquad v_2 = \left(\frac{M - 3m}{M + m}\right) v \tag{4.6.5}$$

We see that the small ball must rise to a height greater than that from which it was dropped, because the fraction in front of v is always greater than 1. Interestingly, if the larger ball is exactly three times the mass of the smaller one, the larger ball comes rest immediately after the collision with the smaller ball.

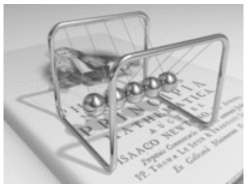
Perhaps you are worried about how we can assume a head-on collision of two balls going in opposite directions at the same speed, when they are in fact in contact with each other throughout? You are right to worry! This is one of those cases where we have *greatly* simplified the conditions in order to get a solution. Clearly if the balls were glued together for the collision with the floor, and then just as they leave contact with the floor the glue dissolves, then the balls would rise together at the same speed, staying in contact.

How do we explain this difference? When the balls remain in contact during the collision with the floor, the large ball is pushed from above by the smaller ball, and compressing the larger ball against the floor more than if the balls were not in contact. This greater compression leads to a greater impulse delivered by the floor. Indeed, we can compute this difference: With the balls in contact, they behave like a single ball with mass $M + m$, so the momentum change due to an elastic collision with the ground is $2\left(M+m\right)v$. If the balls lose contact, and the bottom ball bounces off the floor by itself, the momentum change is only $2Mv$. So the momentum of the system of two balls as they move up is greater when the large ball is in contact with the small ball during the bounce than when it bounces off the floor while not in contact with the small ball. All of the collisions are elastic, so regardless of whether or not the balls are in contact, the kinetic energy of the system as it leaves the floor is the same.

A system of two or more particles can of course have many different momenta for the same total kinetic energy. [*For example, a system of two identical particles moving at equal speeds toward each other has zero momentum, but if they are moving in the same direction at the same speed, the momentum is not zero, while the kinetic energy is the same in both cases.*] This is the case here – the two-ball system comes off the floor with the same kinetic energy in both cases, but the circumstances provide different momenta for those two cases, and this leads to different behaviors for the two objects in the system.

These are not the only two possibilities. If the balls are separated by a very small distance, so that the bottom ball does not fully reflect off the floor before colliding with the higher ball, then contact with the smaller ball will slow the larger ball's departure from the floor, increasing the impulse delivered by the floor. The increased impulse is not as great as if the balls were in contact the whole time, but it is more than the case where the larger ball bounces off the floor unimpeded. In this case, the momentum of the system is less than the case of the connected balls, and more than the case of the disconnected balls. The kinetic energy is still the same, and the result is that the small ball rises faster than the large ball, but not as fast as it does in the disconnected case. The case we calculated above is the *fastest* the smaller ball can be going at the end, and therefore the *highest* it can rise.

Newton’s Cradle



Everyone is familiar with this desktop toy. As fun as it is to watch, it also comes with a puzzle: How does the other side of the line of balls know how many balls to send up? Obviously, if we send two balls down it will provide more momentum to come through the other side, but why can’t that momentum come from a single ball that comes out twice as fast? Or four balls half as fast? The key is in the fact that the collisions are (nearly) elastic. If two balls go in and one ball comes out with that momentum, then it will have a different kinetic energy:

$$\begin{array}{ll} momentum\;in:\;2p & KE\;in:\;2\left(\dfrac{p^2}{2m}\right) \\ \underbrace{momentum\;out:2p}_{\substack{momentum\;conserved\\ all\;\;forces\;internal}} & \underbrace{KE\;out:\left(\dfrac{(2p)^2}{2m}\right)}_{\substack{if\;one\;ball\;comes\;out,\;it\;does\\ so\;with\;twice\;the\;momentum\\ of\;a\;single\;ball\;coming\;in}} = 4\cancel{\dfrac{p^2}{2m}}\;(too\;much) \end{array} \tag{4.6.6}$$

We see a similar thing if too many balls come out (the total outgoing KE is too small). So the cradle “knows” how many balls to send out because it is the only way it can satisfy momentum conservation for the elastic collisions.

CHAPTER OVERVIEW

5: Rotations and Rigid Bodies

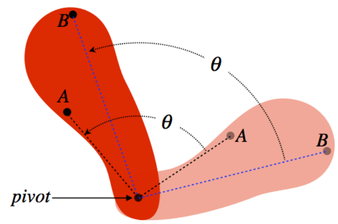
5.1: Rotational Kinematics

Our first foray into linear motion was with kinematics, and we start our discussion of rotation with the same topic.

Rigid Body Rotation

Whenever we talk about “rotation,” there is something that is generally implied – we are not talking about a point mass or a collection of independently-moving point masses. Instead, we are generally referring to the rotation of a rigid object. A rigid object is nothing more than a collection of point objects that are confined to stay at specific positions relative to each other. When we talk about rotation, all these point objects follow different paths and travel different distances, but they all have something in common.

Figure 5.1.1 – Motion of Two Points on a Rotating Rigid Body



Drawing a straight line from the fixed point (called the *pivot*) to two different points on the object, we see that the angles through which these straight lines sweep are the same, and indeed this is true for *every* point on the object. So as we talk about rigid body rotation, our old language of linear motion (displacement, velocity, acceleration) that is based on units of distance and time, will have to give way to a new language for rotational motion, based on the units of radians (the most common unit of angular measure) and time. This language will be very similar to what we used for the linear case, usually with the word "angular" or "rotational" appended in front of the usual words.

Just because we are going to a new language, it doesn't mean we throw out the physical principles we have learned so far. But to apply them in our new area of study, we need to develop some way to translate between the two. Back in Section 1.7, in our discussion of circular motion, we came up with a translation between the arclength traveled by an object in circular motion and the angle is motion sweeps out. Certainly the points *A* and *B* in Figure 5.1.1 are following a circular path (they remain a fixed distance from the pivot), so this relation applies to them. If a given point on a rigid body is a distance r from the pivot, then the relationship between the distance it travels along the arclength and the angle measured in radians is given by Equation 1.7.2, and the relationship between its linear speed and the rate at which the angle is changing (in radians per second) is given by Equation 1.7.3, both of which we'll reiterate here:

$$s=R\theta, \qquad v=\frac{ds}{dt}=R\frac{d\theta}{dt}=R\omega \tag{5.1.1}$$

While s and v are different for every point on the rigid object, we see that θ and ω are common to all of them. We therefore embrace these as our *angular displacement* and *angular velocity* measurements, respectively, for the rigid body as a whole. We can similarly define an *angular acceleration* (α) in terms of the change of the linear speed of a spot on the rotating object:

$$a=\frac{dv}{dt}=r\frac{d\omega}{dt}=r\alpha \tag{5.1.2}$$

While each point mass comprising the rigid object may have its own linear velocity/acceleration, they all share a common angular velocity/acceleration. We therefore can simplify our discussion of rigid body rotation from tracking the many different motions of all of the individual parts of the object to one simple parameter common to all of them. We therefore (for the moment) step away from the translation between linear and angular motion – which we have already discussed in earlier sections – and instead focus on purely rotational motion, following exactly the same path as we did for linear motion. You'll note that as a rule the convention for rotational motion, we stick with Greek variables, in contrast to the Latin variables we used for linear motion.

Alert

Whenever the word "acceleration" is combined with circular motion, one naturally thinks of centripetal acceleration. Be careful not to make that association here! The link between linear acceleration and angular acceleration is through the component of acceleration responsible for speeding up the spot on the rigid object, not the acceleration responsible for changing its direction of motion (which is centripetal acceleration). So for example, an object rotating at a constant rate has no point on it that is *speeding up (and has zero angular acceleration), but every point on it (except at the pivot) experiencing a centripetal acceleration. Conversely, a rotating object that slows down, stops, and reverses its direction of motion is experiencing angular acceleration at all times, including the moment it stops, but the centripetal acceleration of points on the object is zero at the moment that it stops. And finally, the difference should be clear mathematically. For a point on the object, its acceleration has two components:*

$$\overrightarrow{a}=\overrightarrow{a}_{\perp}+\overrightarrow{a}_{\parallel},\qquad where:\begin{cases}a_{\perp}=a_c=r\omega^2\\ a_{\parallel}=r\alpha=r\dfrac{d\omega}{dt}\end{cases}$$

Rotational Equations of Motion

We define the following angular (rotational) versions of what we studied previously in kinematics:

$$\begin{aligned} position &: \theta(t)\\ displacement &: \Delta\theta=\theta_2-\theta_1\\ average\ velocity &: \omega_{ave}=\frac{\Delta\theta}{\Delta t}\\ instantaneous\ velocity &: \omega(t)=\frac{d\theta}{dt}\\ average\ acceleration &: \alpha_{ave}=\frac{\Delta\omega}{\Delta t}\\ instantaneous\ acceleration &: \alpha(t)=\frac{d\omega}{dt}\end{aligned}\tag{5.1.3}$$

The calculus that leads to the equations of motion works out exactly the same way as before (we have only changed the variable names), giving us:

$$\begin{aligned}\theta(t)=&\quad\theta_o+\omega_o t+\tfrac{1}{2}\alpha t^2\\ \omega(t)=&\quad\omega_o+\alpha t\\ \omega_f^2-\omega_o^2=&\quad 2\alpha\left(\Delta\theta\right)\\ \omega_{ave}=&\quad\frac{\omega_o+\omega_f}{2}\quad(if\ \alpha=constant)\end{aligned}\tag{5.1.4}$$

Note that like the case of one-dimensional linear motion, we need to define at the outset a "positive" direction, but for rotation, this means choosing clockwise or counterclockwise from a specific perspective.

Example 5.1.1

A bug stands on the outer edge of a turntable as it begins to spin, accelerating rotationally in the horizontal plane from rest at a constant rate. Find the rate of angular acceleration of the turntable in terms of its radius and the coefficient of static friction if the bug slides off it just as the turntable completes its third full rotation.

Solution

The bug remains on the edge of the turntable thanks to the static friction force, which keeps it going in a circle. When the rotational speed becomes so great that the maximum static friction force is insufficient to maintain this centripetal acceleration, the bug will slide off. The maximum static friction force is the coefficient of static friction multiplied by the normal force, and since the turntable is horizontal and not accelerating vertically, the normal force equals the weight of the bug. We therefore have:

$$\left.\begin{aligned}f_{max}&=ma_c=mr\omega^2\\ f_{max}&=\mu_s N\\ N&=mg\end{aligned}\right\}\quad\Rightarrow\quad\omega^2=\frac{\mu_s g}{r}$$

The "no time" kinematics equation for rotation relates the angular acceleration (which we are looking for), the starting rotational speed (which is zero here, as the turntable starts from rest), the final speed (the speed that causes the bug to lose its grip), and the angle through which the object has rotated (which in this case is 6π – three full rotations):

$$\alpha=\frac{\omega_f^2-\cancel{\omega_o^2}}{2\Delta\theta}=\frac{\frac{\mu_s g}{r}}{2\left(6\pi\right)}=\boxed{\frac{\mu_s g}{12\pi r}}$$

Directions of Rotational Kinematics Vectors

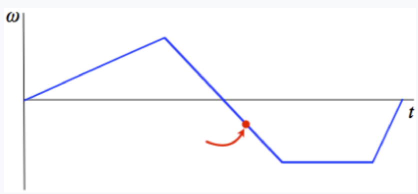
When we did all of this previously, we found it was easy to keep track of directions in one dimension, simply by checking the sign of the value, but when we extended to more dimensions, we needed to treat these quantities like vectors. How can we do this for rotational motion?

The answer comes from all the way back in Chapter 1 – the right hand rule! It goes like this: curl the fingers of your right hand (in their natural finger-curling manner) in the direction that the object is rotating, and your thumb points the direction of the vector. The direction is *perpendicular to the plane of rotation*.

This direction applies to all of the angular motion vectors – displacement, velocity, and acceleration. But be careful about the acceleration vector! Just as in the linear case, the acceleration vector points in the direction of the *changing* velocity vector, not the direction of the velocity vector itself. So if a rotating object is slowing down, the angular acceleration vector points in the opposite direction as the angular velocity vector.

Example 5.1.2

The graph below depicts the rotational velocity of a merry-go-round as a function of time, where the positive direction is defined to be downward (into the surface of the Earth). You are standing near the merry-go-round, watching children go by. At the point indicated in the graph, which of the following are you seeing?



a. *The kids closest to you are moving to the right and are speeding up.*
b. *The kids closest to you are moving to the right and are slowing down.*
c. *The kids closest to you are moving to the left and are speeding up.*
d. *The kids closest to you are moving to the left and are slowing down.*
e. *The kids closest to you are moving to the left, but their speed is not changing.*

Solution

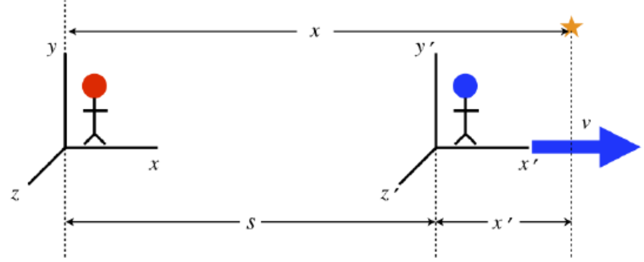
From the RHR, we determine that the positive rotational direction is clockwise as you look at the merry-go-round from above (the kids on the merry-go-round are wondering why you are apparently giving their ride a thumbs-down!). Looking at it from ground level, this means that rotation in a positive direction results in seeing the nearest kids go by from right-to-left. At the point in question, the sign of the rotational velocity is negative, which means the kids are going by left-to-right. A short time later, the rotational velocity will be more negative, which means they are speeding up. So the answer is (a).

5.2: Non-Inertial Frames

Linear Accelerated Frames

Let us return for a moment all the way back to Section 1.8, where we discussed relative motion, and in particular, the Galilean transformation equations. If you recall, this discussion was about comparing what one observer sees to what another observer sees, and the model we used simplified things by having the observers in linear relative motion at a constant speed along their common x-axis, with their clocks synchronized so that $t' = t = 0$ when their origins coincide:

Figure 5.2.1 – Relating Coordinates of Linear Reference Frames



$$\begin{aligned} t' &= t \\ x' &= x - vt \\ y' &= y \\ z' &= z \end{aligned} \tag{5.2.1}$$

These equations relate how one observer logs an event in terms of position and time to how the other person logs it. Suppose these two people are observing the motion of an object. Example 1.8.2 in that section shows how we can use the transformation equations to relate the velocities witnessed by the two observers of a common object. Suppose now we want to relate the accelerations witnessed by both. Clearly with v being a constant, its derivative is zero, which means that both observers agree upon the acceleration, even the component along the x-direction:

$$\frac{dx'}{dt'} = \frac{d\left(x - vt\right)}{dt'} = \frac{d\left(x - vt\right)}{dt} = \frac{dx}{dt} - v \quad \Rightarrow \quad \frac{d^2x'}{dt'^2} = \frac{d^2x}{dt^2} - \cancelto{0}{\frac{dv}{dt}} \tag{5.2.2}$$

Suppose, however, that the primed frame is *accelerating* relative to the unprimed frame. Then the separation of the frames (shown as an s in the diagram above) has a different time dependence, and we get different transformation equations. We once again simplify things again as much as possible by insisting that:

- the acceleration is a constant value a, along the line of motion in the $+x$-direction
- the relative velocities of the frames at the common starting time of $t' = t = 0$ is v_o in the $+x$-direction
- the origins coincide at the common starting time of $t' = t = 0$

This gives us the following transformation equations:

$$\begin{aligned} t' &= t \\ x' &= x - v_o t - \tfrac{1}{2}at^2 \\ y' &= y \\ z' &= z \end{aligned} \tag{5.2.3}$$

In this case, the two observers would *not* agree on the acceleration of the commonly-observed object, as it would differ in its x-component by an amount a:

$$\frac{d^2x'}{dt'^2} = \frac{d^2x}{dt^2} - \cancelto{0}{\frac{dv_o}{dt}} - a \tag{5.2.4}$$

Fictitious Forces

Now that we are talking about accelerations, the question of forces comes up. Consider an object that remains at rest in the unprimed frame. That observer must conclude, according to Newton's second law, that it is not experiencing a net force. Naturally this observer can do a careful analysis of all the forces (friction, normal, spring, etc.) that appear on the free-body diagram for that object to confirm that the forces all add up to zero.

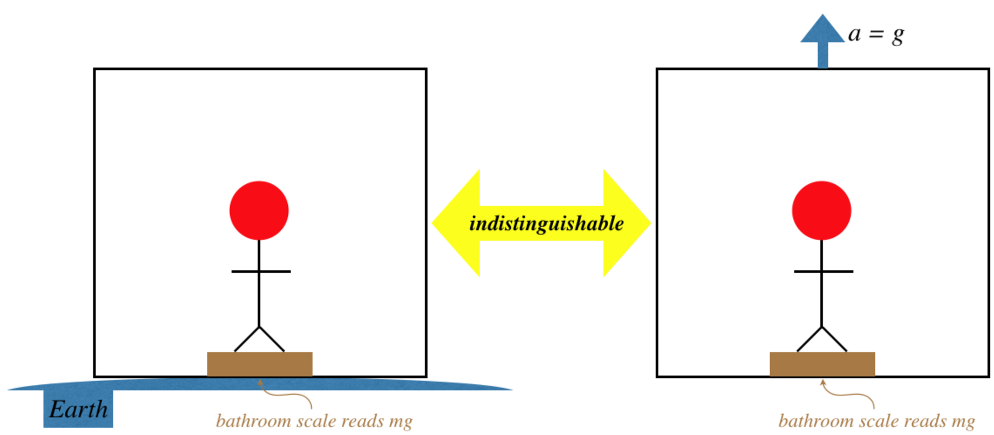
What about the other observer's analysis of Newton's second law? The primed observer does not see the object experiencing zero acceleration, as the transformation between their frame and the frame of the other observer demonstrates. This observer therefore concludes that there must be a net force on the object, to accelerate it. But the analysis performed by the unprimed observer who looked at all the physical forces will not come out different for the primed observer – the spring scale used to measure a normal force will not read something different for these two people. So how does this primed observer confirm Newton's second law? They must invent a new, "mystery force" that has no apparent source. We call such a force that comes about purely because the frame in which it is observed is accelerating *fictitious*. A frame in which there are no fictitious forces present is called an *inertial frame*, and frames that give rise to fictitious forces are called *non-inertial*.

One of the defining characteristics of a fictitious force is that it results in the same acceleration regardless of any physical features of the object, because it is the acceleration of the non-inertial frame that is actually witnessed. Given that the same acceleration is observed for all objects of all masses, the accelerated observer would have to conclude that the fictitious force is proportional to the mass of the object.

$$\left.\begin{array}{ll}\text{fictitious force } \mathcal{F}_1 \text{ on } m_1 \text{ and Newton's 2nd law:} & a=\dfrac{\mathcal{F}_1}{m_1} \\ \text{fictitious force } \mathcal{F}_2 \text{ on } m_2 \text{ and Newton's 2nd law:} & a=\dfrac{\mathcal{F}_2}{m_2}\end{array}\right\} \Rightarrow \quad \frac{\mathcal{F}_1}{m_1}=\frac{\mathcal{F}_2}{m_2} \tag{5.2.5}$$

But wait, there is a force we use all the time that satisfies this very property – gravity! The force of gravity is given by mg, and indeed every object in free fall is seen to accelerate exactly the same amount. So is gravity a fictitious force? We know there is a cause for this force – the earth exerts the force on all objects in proportion to their mass, so maybe gravity just coincidentally resembles a fictitious force. But in fact Einstein, in his theory of general relativity, proposed that gravity *is* essentially a fictitious force. He posited what he called the *equivalence principle*, which stated that an observer in a closed room in the absence of gravity that is accelerating in the "upward" direction at a rate of $9.8\frac{m}{s^2}$ can perform no experiment that will distinguish between that circumstance and the case of the identical closed room on the surface of the earth.

Figure 5.2.2 – Equivalence Principle

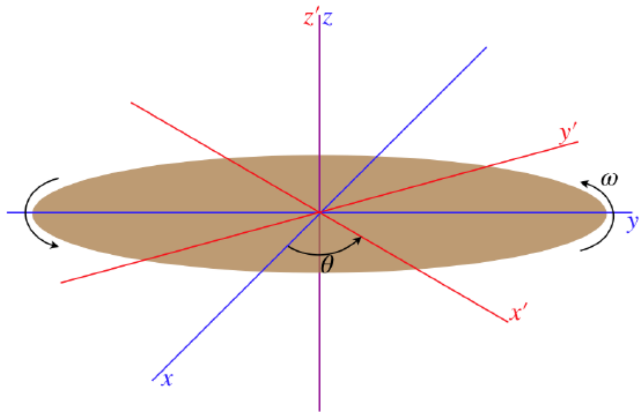


As simple as this equivalence might seem to be, it is nevertheless very complicated mathematically to relate a force that attracts objects toward a single point (the center of the planet) to the fictitious force seen in a linearly-accelerated frame, so we unfortunately have to end our discussion of general relativity here and leave it for future coursework in physics.

Rotating Frames

As we know, acceleration that involves speeding up or slowing down in a straight line is not the only variety. We now want to look at how to transform between an inertial frame and one that is rotating around a fixed point in the inertial frame. We will place the origin of the inertial frame on the axis of rotation of the rotating frame, with the axis perpendicular to the plane of rotation coinciding with the z-axis of the inertial frame. The red axes in the figure below are rotating at a constant rate while the blue axes remain stationary.

Figure 5.2.3 – Relating Coordinates of Rotational Reference Frames



Let's imagine an object at rest on the positive x'-axis at the moment when all three axes are aligned. The observer at rest in the stationary (unprimed) frame that is watching this object will see it moving in a circle, and at the moment the axes are aligned, it is moving in the $+y$-direction. If we call the distance from the z-axis r, then the speed of the object as seen by the observer in the unprimed frame is $\omega\, r$. Putting the magnitude and direction together gives us the velocity vector $\omega\, r\, \hat{j}$ for this object, as measured by the stationary observer.

We can write this velocity in terms of the position vector $\overrightarrow{r} = r\, \hat{i}$ and the rotation vector (which we get from the right-hand rule) $\overrightarrow{\omega} = \omega\, \hat{k}$ as follows:

$$\omega\, r\, \hat{j} = \overrightarrow{\omega} \times \overrightarrow{r} \tag{5.2.6}$$

This follows very simply, because the cross product of $\hat{k} \times \hat{i}$ is just $\hat{j}$.

The nice thing about this way of expressing the velocity is that it doesn't depend upon the specific circumstance of being on the x-axis. Wherever the object happens to be, as long as it has zero velocity measured in the rotating frame, this cross-product combination of the angular velocity vector and the position vector gives the proper magnitude and direction of the velocity vector seen by the stationary frame.

If an object instead has zero velocity in the stationary frame, then the observer in the primed frame will see the same thing, but in the opposite direction: $-\overrightarrow{\omega} \times \overrightarrow{r}$. Now suppose the object happens to have a non-zero velocity measured in the stationary frame. To get the velocity vector seen in the rotating frame, we simply add vectors according to Equation 1.8.3:

$$\overrightarrow{v}_{object\ rel\ to\ primed} = \overrightarrow{v}_{object\ rel\ to\ unprimed} + \overrightarrow{v}_{unprimed\ rel\ to\ primed} \quad \Rightarrow \quad \overrightarrow{v}' = \overrightarrow{v} - \overrightarrow{\omega} \times \overrightarrow{r} \tag{5.2.7}$$

This is the rotational version of the Galilean transformation equations.

Let's suppose we take a snapshot of the position of an object at an instant in time. The observers in the stationary and rotating frames will agree upon the position vector that locates the object. If their axes aren't aligned, then they won't agree upon the *components* of this vector, but provided their origins still coincide, they will naturally agree that the two position vectors have the same length and point in the same direction. They don't agree upon the velocity of the object, and since the velocity is the time rate of change of the position vector, *it must be their time derivatives that are different*. Writing the velocities in terms of the time derivatives of position vectors, we get the relationship between the time derivatives:

$$\left[\frac{d}{dt}\overrightarrow{r}\right]_{rotating} = \left[\frac{d}{dt}\overrightarrow{r}\right]_{stationary} - \overrightarrow{\omega} \times \overrightarrow{r} \quad \Rightarrow \quad \left[\frac{d}{dt}\right]_{stationary} = \left[\frac{d}{dt}\right]_{rotating} + \overrightarrow{\omega} \times \tag{5.2.8}$$

The "empty" cross-product of "$\overrightarrow{\omega} \times$" is not a typo. The idea is that you just put a vector on the right side of everything in this equation – the derivatives act on the vector, and the vector is the second half of the cross-product. In essence, we have "divided-out" the position vector from the previous equation. This gives us a translation between time derivatives in the two frames - any time derivative of a vector in the rotating frame is equivalent to a time derivative of the vector in the stationary frame minus the cross product of the rotation vector with that same vector.

Alert

Note that this rule only applies to ***time*** *derivatives, and only when they act on* ***vectors****. The former is because snapshots (holding time constant) render the two frames indistinguishable (they only differ in how they evolve through time). The latter is obvious, given that we have to take a cross product with the second term!*

Example 5.2.1

Given that an object can be at rest in one frame and moving in the other, observers in a stationary frame and a rotating frame will not in general agree upon the kinetic energy of an object. Consider a point mass moving directly away from the origin in the rotating frame. Show that the kinetic energy of this object measured in the stationary frame equals the kinetic energy in the rotating frame plus the rotational kinetic energy it gets from rotating along with the rotating frame.

Solution

We compute the kinetic energy using the dot product of the velocity vector with itself, so the kinetic energy in the stationary frame is:

$$KE_{stationary} = \frac{1}{2}mv^2 = \frac{1}{2}m\,\overrightarrow{v}\cdot\overrightarrow{v}$$

Plugging Equation 5.2.7 *into this gives:*

$$\begin{aligned} KE_{stationary} &= \frac{1}{2}m\left[\overrightarrow{v}' + \overrightarrow{\omega}\times\overrightarrow{r}\right]\cdot\left[\overrightarrow{v}' + \overrightarrow{\omega}\times\overrightarrow{r}\right] \\ &= \frac{1}{2}m\left[\overrightarrow{v}'\cdot\overrightarrow{v}' + 2\,\overrightarrow{v}'\cdot\left(\overrightarrow{\omega}\times\overrightarrow{r}\right) + \left(\overrightarrow{\omega}\times\overrightarrow{r}\right)^2\right] \end{aligned}$$

With the object moving directly radially in the rotating frame, it must be that $\overrightarrow{v}'$ *is parallel to* $\overrightarrow{r}$*, but the cross product* $\overrightarrow{\omega}\times\overrightarrow{r}$ *is necessarily* ***perpendicular*** *to* $\overrightarrow{r}$*. This means that the dot product in the middle term is between two perpendicular vectors, which gives zero. The last term is a dot product between two vectors that are the same, so it just gives the magnitude-squared of that vector. The vector* $\overrightarrow{r}$ *is perpendicular to* $\overrightarrow{\omega}$*, so the magnitude of the vector that comes from the cross product is:*

$$\left|\overrightarrow{\omega}\times\overrightarrow{r}\right| = \omega\, r\,\sin 90^o = \omega\, r$$

Putting this in above gives:

$$KE_{stationary} = \frac{1}{2}mv'^2 + \frac{1}{2}mr^2\omega^2$$

The point mass has a moment of inertia equal to mr^2 *relative to the axis of rotation, so the second term is precisely the rotational energy around that axis,* $\frac{1}{2}I\omega^2$*.*

More Fictitious Forces

Every point on the rotating frame is accelerating, so if we are observing from that frame, we should expect to witness mysterious accelerations of objects that are not accelerating in the unprimed frame. That is, the primed observer will not be able to pinpoint a true physical force causing it to accelerate, and will need to invent a fictitious force to explain this acceleration with Newton's second law. To determine this force, we first need to derive a translation between accelerations in the two frames. We get the acceleration by taking the derivative of the velocity, so we apply Equation 5.2.8 to the velocity vector in the stationary frame:

$$\left[\frac{d}{dt}\overrightarrow{v}\right]_{stationary} = \left[\frac{d}{dt}\overrightarrow{v}\right]_{rotating} + \overrightarrow{\omega}\times\overrightarrow{v} \tag{5.2.9}$$

The left side of this equation is the acceleration observed in the stationary frame, but we want it in terms of the acceleration in the rotating frame, so now we plug in Equation 5.2.7:

$$\vec{a} = \left[\frac{d}{dt}\left(\vec{v}' + \vec{\omega} \times \vec{r}\right)\right]_{rotating} + \vec{\omega} \times \left(\vec{v}' + \vec{\omega} \times \vec{r}\right) \tag{5.2.10}$$

Now apply the derivative, noting that ω is a constant and the derivative of the position vector in the stationary frame is the velocity in that frame:

$$\vec{a} = \vec{a}' + 2\vec{\omega} \times \vec{v}' + \vec{\omega} \times \left(\vec{\omega} \times \vec{r}\right) \tag{5.2.11}$$

Okay, suppose we have a set of real forces (tension, normal, etc.) acting on the object, resulting in a net force, and according to Newton's 2nd law, an acceleration. The net force that the observer measures for this object is therefore simply $m\vec{a}$. But this isn't the net force measured in the rotating frame. This observer sees the object affected by these forces *plus* two other fictitious forces that arise from the rotation of the frame from which this observer is making measurements. Specifically, the rotating observer measures a net force given by:

$$\vec{F}' = \vec{F} - 2m\,\vec{\omega} \times \vec{v}' - m\,\vec{\omega} \times \left(\vec{\omega} \times \vec{r}\right) \tag{5.2.12}$$

The first of the two fictitious force terms, which depends upon the velocity measured in the rotating frame, is called the coriolis force. The second, which doesn't require that the object be moving in the rotating frame, is called the centrifugal force.

Alert

It is a common mistake to confuse the terms "centrifugal" and "centripetal." There are several reasons for this, but the main reason is that when first learning physics, students think that it is natural to continue moving in a circle, and feel like there must be a force that "pulls things away from the center" when they turn corners. From the perspective of the frame that is accelerating, this is the fictitious centrifugal force. But then we are taught that the real force that is acting is causing us to change direction, and the acceleration caused by this force is centripetal. The thing to keep in mind is that "centripetal" refers to a specific kind of acceleration, caused by real forces – it is not a type of force at all. On the other hand, "centrifugal" does refer to a force, albeit a fictitious one that arises because the observer is in a rotating frame.

Note that if the object has zero velocity in the rotating frame, then the net force keeping it going in the circle produces an acceleration that is centripetal, which when multiplied by the mass yields a force vector that has the same magnitude as the centrifugal force vector in the rotating frame, and in the opposite direction.

Example 5.2.2

An object is at rest in a stationary frame.

a. *Is this object experiencing centripetal acceleration measured in the stationary frame?*
b. *This same object is now observed in a rotating frame with the same origin as the stationary frame, and the object is not positioned at this origin. Does an observer in this rotating frame measure a centrifugal force acting on the object?*
c. *Reconcile the answers to (a) and (b).*

Solution

a. *Of course not! The object is not moving at all, let alone in a circular fashion.*

b. *Yes! The frame is rotating, so $\vec{\omega}$ is not zero, and the object is not at the origin, so $\vec{r}$ is not zero. These vectors are at right angles to each other, so the centrifugal force $\vec{\omega} \times \vec{\omega} \times \vec{r}$ is not zero.*

c. *How can there be a centrifugal force but no centripetal acceleration? If this is confusing, it displays once again the confusion of these two terms. Let's think about what the observer in the rotating frame sees: They see the object traveling in a circle at a constant speed, in the direction opposite to their frame's rotation. So this observer sees not only a centrifugal force, but also a coriolis force. After working through all the right-hand rules, we find that the centrifugal force acts outward from the axis of rotation (as it always does), and in this particular case, the coriolis force acts inward toward the axis of rotation. Adding these two forces together gives:*

$$\vec{F}'_{net} = 2m\omega v' \text{ (inward)} + m\omega^2 r \text{ (outward)}$$

We know how the velocity measured by the rotating observer is related to the distance from the axis and the rotational speed of the frame, since the object is stationary in the rotating frame:

$$v' = r\omega$$

Putting this in above, we find that the net force measured in the rotating frame is:

$$\vec{F}'_{net} = 2mr\omega^2 \text{ (inward)} + mr\omega^2 \text{ (outward)} = mr\omega^2 \text{ (inward)}$$

This is exactly the net force needed to keep an object moving in a circle at a constant speed, which makes sense, because this is exactly what the rotating observer sees happening. The combination of the centrifugal force and coriolis force yielded a net force in the rotating frame that resulted in centripetal acceleration measured in that frame! How's that for confusing?

5.3: Rotational Inertia

Rotational Kinetic Energy and Rotational Inertia

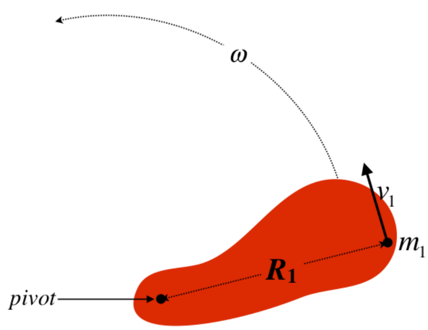
While our first approach to studying dynamics for linear motion was Newton's laws (forces cause accelerations), we will find it easier to examine rotational dynamics from a standpoint of energy first. Consider an object that is rotating around a stationary center of mass. Does such an object possess kinetic energy? We might be inclined to say that it does, but with the center of mass not moving, its momentum is zero, which would make the quantity $\frac{p^2}{2m}$ also equal to zero.

We must not be such slaves to memorized equations! This equation (by itself) *never* applied to a system of multiple particles, which can easily have a zero total momentum and yet still have a nonzero kinetic energy. Well, rigid objects are systems of multiple particles, and when they are rotating, all those particles (except those right at the pivot point) are moving, which means they all have kinetic energy. At any given moment, there are particles moving in opposite directions, and if the center of mass of the object is stationary, these opposite momenta (which are vectors) cancel, Their kinetic energies, on the other hand, are not vectors, and are all positive numbers, so they can never cancel out.

In some sense, the particles comprising a rotating object can be thought of as contributing to the "internal" energy of the object as we discussed back in Section 4.4. But doing this runs contrary to the main reason for the introduction of the mechanical/internal energy idea, which was to separate the kinetic energy of the system *that we can clearly see* from the kinetic energy that is *concealed from us* inside the confines of the system. We can clearly see rotational motion of an object, so we choose to include rotational kinetic energy in the category of "mechanical energy."

Okay, so a rotating object does possess kinetic energy. Our task now is to express that kinetic energy in terms of the rotation variables we have already defined, but all we know about kinetic energy is the linear version. In Figure 5.3.1 we consider the motion of a single particle within a rigid rotating object.

Figure 5.3.1 – Motion of a Single Particle in a Rotating Rigid Body



This is particle #1 – one of many within the rigid object. We can write down its kinetic energy, and in fact we can express it in terms of a rotational variable and the particle's distance from the pivot:

$$KE_1=\frac{1}{2}m_1v_1^2=\frac{1}{2}m_1\left(R_1\omega\right)^2=\frac{1}{2}m_1R_1^2\omega^2 \tag{5.3.1}$$

If we want the total kinetic energy of the object, we need to add up the kinetic energy of all the particles. Thanks to our definition of angular velocity, we can factor that part out of all the terms:

$$KE_{whole\ object}=\frac{1}{2}m_1v_1^2+\frac{1}{2}m_2v_2^2+\dots=\frac{1}{2}m_1R_1^2\omega^2+\frac{1}{2}m_2R_2^2\omega^2+\dots=\frac{1}{2}\left[m_1R_1^2+m_2R_2^2+\dots\right]\omega^2 \tag{5.3.2}$$

Notice that the quantity in brackets in the final equality is determined by the distribution of mass throughout the object. That is, it is an intrinsic property of the object, not dependent upon how it is moving. We generally abbreviate this quantity with an I, which gives us a familiar form for the kinetic energy formula:

$$KE_{rotating\ rigid\ body}=\frac{1}{2}I\omega^2,\qquad where:\ \ I\equiv m_1R_1^2+m_2R_2^2+\dots \tag{5.3.3}$$

This looks just like the linear kinetic energy formula, with the angular speed replacing the linear speed, and I replacing the mass. This quantity certainly contains some information about the mass of the object, but it is more complicated than just the mass, and is called the *rotational inertia*, or more commonly (and less descriptively), the *moment of inertia*. Notice that this "inertia" depends not only upon the amount of stuff (mass), but also where that mass is. This means that two different objects can actually weigh exactly the same amount, but when they are rotated at equal speeds, one of them has more KE than the other. As you might guess, this occurs when more of the mass is concentrated farther from the pivot for the former object than the latter.

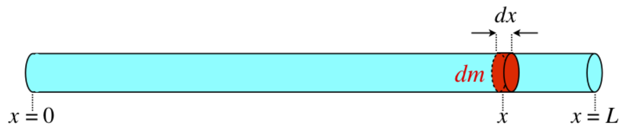
Alert

It is important to note that we will only be considering rotations around ***axes****, not points. In our two-dimensional figures, an axis that is perpendicular to the plane of the figure is indistinguishable from a single point, but we will not discuss motion that involves an object's motion changing its plane of rotation. So rotational inertia for three-dimensional objects involves the distances of the tiny masses from a common axis, not a common point.*

Calculating Rotational Inertia for Continuous Objects

Our task is to compute the rotational inertia, for which the formula in terms of masses and their positions is different from the one for center of mass (see Section 4.2), but the procedure is exactly the same. We start with the same picture (Figure 4.2.3, which is reproduced below), and convert the sums into integrals, as before.

Figure 5.3.2 – Setup Diagram for Computations Involving Mass Density of a Thin Rod



$$I = dm_1x_1^2 + dm_2x_2^2 + \cdots = \int\limits_{x=0}^{x=L} dm\ x^2 \tag{5.3.4}$$

Note that the rotational inertia is calculated around a specific pivot point, which we have chosen to be our origin for the calculation.

As before, we replace the dm with $\lambda\left(x\right)dx$, and we have our formula for the rotational inertia along the x-axis around the pivot point at the origin:

$$I = \int\limits_{x=0}^{x=L} \lambda\left(x\right)x^2dx \tag{5.3.5}$$

Let's return to the cases for which we computed the centers of mass in Section 4.2 – the uniform and non-uniform rod. Unlike the case of center of mass, where the answer is a location on the rod, the final answer for the rotational inertia will have units of $kg\cdot m^2$, and the formula for it will involve the total mass of the rod and its length. Also it is important to remember that while the center of mass is a location that doesn't depend upon where we put our coordinate system to calculate it, the rotational inertia is only defined relative to a specific pivot point.

A Uniform Rod of Mass M and Length L, Pivoted About an End

Plugging the constant λ into Equation 5.3.5 and performing the integral gives:

$$I = \int\limits_{x=0}^{x=L} \lambda x^2dx = \lambda\left[\frac{1}{3}x^3\right]_0^L = \frac{1}{3}\lambda L^3 \tag{5.3.6}$$

We are not finished yet, because this answer is not in terms of the rod's mass. Since this rod is uniform, the mass is simply the (constant) density multiplied by its length, which gives:

$$I = \frac{1}{3}\left(\frac{M}{L}\right)L^3 = \frac{1}{3}ML^2 \tag{5.3.7}$$

We will find that every rotational inertia we encounter has this basic form: A constant (usually written as a fraction) multiplied by the mass of the object and the square of some natural length dimension of the object. In this case it is the length of the rod, but it may also be something like the radius of a disk or sphere.

A Non-Uniform Rod of Mass M and Length L, Pivoted About Its Lighter End

Now we repeat the process for the non-uniform density function for which we computed the center of mass in Section 4.2:

$$\lambda(x) = \lambda_o\left(\frac{x}{L}+1\right) \tag{5.3.8}$$

Note that unlike the uniform case, the results should not come out the same for both ends of the rod, since more of the mass is concentrated near the end at $x = L$. We are calculating this rotational inertia about the lighter end, since all of the x values in the integral are measured from that end.

$$I = \int\limits_{x=0}^{x=L} \lambda_o\left(\frac{x}{L}+1\right)x^2dx = \lambda_o\left[\frac{1}{4L}x^4+\frac{1}{3}x^3\right]_0^L = \frac{7}{12}\lambda_o L^3 \tag{5.3.9}$$

We are not done yet, because we are given the mass of the rod, not the constant λ_o. We therefore need to compute the total mass in terms of this constant. We do this by integrating density function over the length of the rod:

$$M = \int\limits_{x=0}^{x=L} \lambda_o\left(\frac{x}{L}+1\right)dx = \lambda_o\left[\frac{1}{2L}x^2+x\right]_0^L = \frac{3}{2}\lambda_o L \quad\Rightarrow\quad \lambda_o = \frac{2}{3}\frac{M}{L} \tag{5.3.10}$$

Plugging this back in above gives our answer:

$$I = \frac{7}{12}\left(\frac{2}{3}\frac{M}{L}\right)L^3 = \frac{7}{18}ML^2 \tag{5.3.11}$$

Example 5.3.1

Find the rotational inertia of the non-uniform rod of mass M and length L whose mass density function is given by Equation 5.3.8, when rotated about its heavier end ($x = L$).

Solution

The difference between this calculation and the one above is that the variable x that appears in Equation 5.3.5 doesn't match the x that appears in the density formula. The density formula is referenced to our coordinate system, but the x in the rotational inertia integral represents the distance of each tiny piece of mass dm from the pivot point at $x = L$. So we need to make a change in the integral so that the x variable that appears in it matches the x in the density function. Making the substitution $x \rightarrow L - x$ (so $dx \rightarrow -dx$), into the integral does the trick, because then the integrand is zero at the pivot point ($x = L$) as it should be:

$$\begin{aligned} I_{heavy\ end} &= \int\limits_{x=L}^{x=0} dm\ (x-L)^2 = \int\limits_{x=L}^{x=0} \lambda(x)\,(x-L)^2\,(-dx) = \int\limits_{x=0}^{x=L} \lambda_o\left(\frac{x}{L}+1\right)(x-L)^2dx \\ &= \lambda_o \int\limits_{x=0}^{x=L}\left(\frac{x^3}{L}-x^2-xL+L^2\right)dx = \lambda_o\left[\frac{x^4}{4L}-\frac{x^3}{3}-\frac{x^2L}{2}+xL^2\right]_0^L \\ &= \tfrac{5}{12}\lambda_o L^3 \end{aligned}$$

We need to plug in for λ_o (which was computed above) to get our final answer:

$$I_{heavy\ end} = \frac{5}{12}\left(\frac{2}{3}\frac{M}{L}\right)L^3 = \boxed{\frac{5}{18}ML^2}$$

Principal Axes

It's clear that the choice of the pivot is important to the calculation of the rotational inertia, but so is the axis. Real objects are 3-dimensional, so they actually have 3 independent rotation axes, each of which has its own rotational inertia around it. These axes are called the *principal axes*. The origin of these axes is located at – what else? – the center of mass of the object. The principal axes are only easy to identify for objects with some degree of symmetry. Some objects are so symmetric that more than one set of axes will work. For example, a uniform sphere has so much symmetry that any set of three mutually perpendicular axes whose origin coincides with the center of the sphere will work, and of course the rotational inertias around all these axes are the same.

The reason it is natural to define the origin of the principal axes to be at the center of mass is that if an object is rotating freely in space with no forces on it, its axis of rotation *must* pass through its center of mass (though it doesn't need to be around one of the principal axes). This is actually surprisingly easy to prove. Suppose an object was rotating around an axis that does not pass through the center of mass. This would mean that the center of mass is moving in a circle around the axis of rotation. But circular motion is accelerated motion. According to Newton's second law, the center of mass cannot be accelerating if there are no forces on the object, which contradicts our assumption.

Computing Rotational Inertia Without Integration

Throughout our study of mechanics, our goal has been to develop shortcut tools to help us deal with physical systems in simpler ways. We developed work-energy so that we could solve problems that pay no attention to direction or time without slogging through Newton's laws (such as speed at a given height on a loop-de-loop). We developed impulse-momentum so that we could more easily solve problems involving systems in which the internal forces are complicated (such as collisions). Now we are developing a tools related to rigid body rotations so that we don't have to track the linear motions of all the particles in the system. With this very practical mindset, it is not surprising that physicists have developed tools for computing rotational inertia that avoid the ugliness of always having to perform integrals. The first such shortcut is simply a collection of rotational inertias that are associated with common symmetric geometries, such as rods, disks, and spheres. Our collection is given at the end of the section. There are two tools that we can combine with our collection of rotational inertias that will allow us to "bootstrap" our way to determining many more.

Additivity Around a Common Axis

Suppose we know the rotational inertias of two separate objects around a common axis. If these two objects are attached so that they rotate together rigidly around that common axis, then the rotational inertia of the combined object is simply the sum of their rotational inertias. This is evident from the formula for rotational inertia: Each object has its own sum of mx^2 terms, and when the objects are combined such that their x axes are common, then the new sum of mx^2 terms is simply the combination of the two individual sums. To summarize:

$$I=I_1+I_2 \tag{5.3.12}$$

Example 5.3.2

Use the additive property of rotational inertia and the result given by Equation 5.3.7 to find the rotational inertia of a uniform thin rod of mass M and length L about its center of mass.

Solution

We can treat a rod rotated around an axis through its center as if it is two separate half-rods of half the mass and half the length, attached at their ends. The axis that passes though the center of the rod passes through the ends of these two half-rods, and we know the rotational inertia of each half-rod. The additivity property then gives us the rotational inertia of the whole rod about its center:

$$I_{uniform\;thin\;rod\;about\;its\;center}=2I_{half-rod\;about\;end}=2\left[\frac{1}{3}\left(\frac{M}{2}\right)\left(\frac{L}{2}\right)^2\right]=\boxed{\frac{1}{12}ML^2}$$

Parallel Axis Theorem

As we have seen multiple times already, just changing the axis around which an object is rotated will result in a different rotational inertia. Suppose we calculate the rotational inertia of an object about an axis, then slide that axis in a parallel fashion on the object, and calculate the new rotational inertia, then do it over and over, recording the new values each time. One might ask, "Where is the axis (parallel to the original one) for which the rotational inertia is the *smallest*?" Is there any way to guess where this might be, and is it unique, or might there be multiple places where the rotational inertia hits a minimum?

To answer this question, let's look at a one-dimensional object that lies along the x-axis, and consider its rotational inertia around the y-axis. Writing it as a sum rather than an integral, it is:

$$I = m_1x_1^2 + m_2x_2^2 + \ldots \tag{5.3.13}$$

Now let's suppose we decide to change where we place the origin, moving it a distance $+x$ along the x-axis. When we do this, the distance from the axis to mass m_1 changes from x_1 to $x_1 - x$. Also, since the original axis went through the origin, this new axis is no longer the y-axis – now it intersects the x-axis at x. The new rotational inertia is, therefore:

$$I = m_1\left(x_1 - x\right)^2 + m_2\left(x_2 - x\right)^2 + \ldots \tag{5.3.14}$$

We can consider this to be a function of x. That is, this formula provides the rotational inertia of the object about the axis located at x. We can now answer our question about where the rotational inertia is a minimum by using calculus. The value of x for which the function $I\left(x\right)$ is a minimum satisfies:

$$0 = \frac{dI}{dx} = -2m_1\left(x_1 - x\right) - 2m_2\left(x_2 - x\right) + \ldots \tag{5.3.15}$$

Solving for x here provides a familiar result:

$$x = \frac{m_1x_1 + m_2x_2 + \ldots}{m_1 + m_2 + \ldots} \tag{5.3.16}$$

The rotational inertia of an object for all axes parallel to each other is a minimum for the axis that passes through the center of mass! Actually, this should not be too surprising. The rotational inertia of an object will be minimized around an axis that is as close as possible to as much of the object's mass as possible, and the center of mass is the "average location of mass," so it makes sense that this would be "as close to as much of the object's mass as possible."

Given this information, we can write the rotational inertia of an object around an axis parallel to an axis passing through the center of mass a positive-valued "adjustment" to the rotational inertia around the center of mass. It turns out (we will not prove it here) that this adjustment is quite simple – it is just the mass of the object multiplied by the square of the offset distance between the new axis and the axis through the center of mass. This is called the parallel axis theorem:

$$I_{new} = I_{cm} + Md^2, \tag{5.3.17}$$

where d is the distance separating the new axis and the center of mass.

Example 5.3.3

Use the parallel axis theorem and the result given by Equation 5.3.7 to find the rotational inertia of a uniform thin rod of mass M and length L about its center of mass.

Solution

The distance that the end of the rod is separated from the rod's center of mass is $d = L/2$. Plugging this into the parallel axis theorem gives our answer, which agrees with what we got in Example 5.2.2:

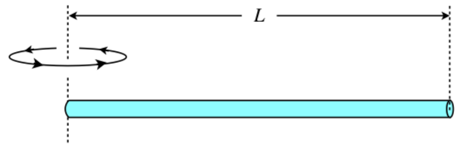
$$I_{new} = I_{cm} + Md^2 \quad \Rightarrow \quad I_{cm} = I_{new} - Md^2 = \frac{1}{3}ML^2 - M\left(\frac{L}{2}\right)^2 = \left(\frac{1}{3} - \frac{1}{4}\right)ML^2 = \boxed{\frac{1}{12}ML^2}$$

Rotational Inertias of Some Common Geometries

In all of the cases indicated below, the mass of the object is M, and the material making up the object has uniform density. The reader is encouraged (as an exercise) to navigate their way between various relations using the additivity and parallel axes theorem tools. [*Note: When it comes to rotating two-dimensional objects such as rings and disks, we will confine our studies to axes perpendicular to the two-dimensional planes in which these objects lie. For rotations around axes parallel to this plane, one would need yet another useful tool, known as the perpendicular axes theorem.*]

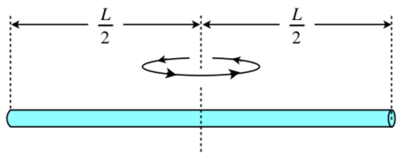
Thin Rods

Figure 5.3.3 – Thin Straight Rod Rotated About One end



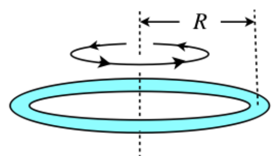
$$I = \frac{1}{3} M L^2 \tag{5.3.18}$$

Figure 5.3.4 – Thin Straight Rod Rotated About Center



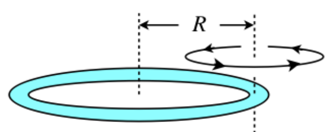
$$I = \frac{1}{12} M L^2 \tag{5.3.19}$$

Figure 5.3.5 – Thin Circular Ring (or Thin Cylindrical Shell) Rotated About Center



$$I = M R^2 \tag{5.3.20}$$

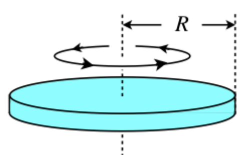
Figure 5.3.6 – Thin Circular Ring (or Thin Cylindrical Shell) Rotated About Edge



$$I = 2 M R^2 \tag{5.3.21}$$

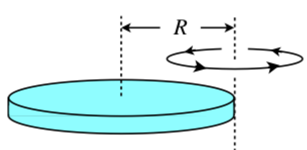
Disks (or Cylinders)

Figure 5.3.7 – Solid Disk (or Cylinder) Rotated About Center



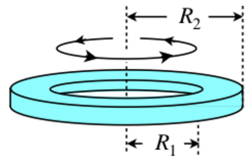
$$I = \frac{1}{2} M R^2 \tag{5.3.22}$$

Figure 5.3.8 – Solid Disk (or Cylinder) Rotated About Edge



$$I = \frac{3}{2} M R^2 \tag{5.3.23}$$

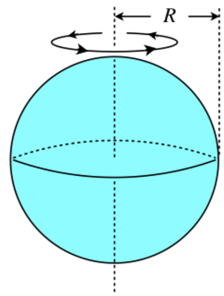
Figure 5.3.9 – Hollow Disk (or Cylinder) Rotated About Center



$$I=\frac{1}{2}M\left(R_1^2+R_2^2\right) \tag{5.3.24}$$

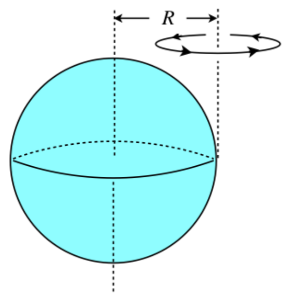
Spheres

Figure 5.3.10 – Solid Sphere Rotated About Center



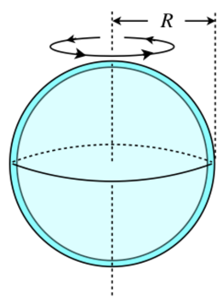
$$I=\frac{2}{5}MR^2 \tag{5.3.25}$$

Figure 5.3.11 – Solid Sphere Rotated About Edge



$$I=\frac{7}{5}MR^2 \tag{5.3.26}$$

Figure 5.3.12 – Thin Spherical Shell Rotated About Center



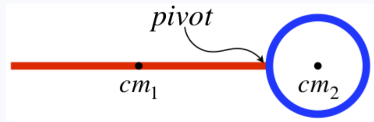
$$I=\frac{2}{3}MR^2 \tag{5.3.27}$$

Example 5.3.4

The frame of a badminton racquet is constructed from two identical thin aluminum rods of uniform density, mass M, and length L. One of the rods is bent into a circle and is welded to the end of the other rod. Find the rotational inertia of this racquet around the axis perpendicular to the plane of the circle that passes through the point where the loop connects to the rod.

Solution

Start with a diagram of the racquet:



The points labeled cm_1 and cm_2 are the centers of mass of the two rods, respectively. We have the rotational inertias of the two rods about their respective centers of mass, and using the fact that the length of the bent rod is its circumference gives:

$$I_{cm_1} = \frac{1}{12}ML^2$$

$$I_{cm_2} = MR^2 = M\left(\frac{L}{2\pi}\right)^2 = \frac{1}{4\pi^2}ML^2$$

Now we need to use our two tools. First, we need to change the axes for each of these rotational inertias from cm_1 and cm_2 to the pivot. We do this using the parallel axis theorem. The distance that the new axis is from cm_1 is \frac{L}{2}, and the distance between cm_2 and the pivot is $R = \frac{L}{2\pi}$, so:

$$I_1 = I_{cm_1} + Md_1^2 = \frac{1}{12}ML^2 + M\left(\frac{L}{2}\right)^2 = \frac{1}{3}ML^2$$

$$I_2 = I_{cm_2} + Md_2^2 = \frac{1}{4\pi^2}ML^2 + M\left(\frac{L}{2\pi}\right)^2 = \frac{1}{2\pi^2}ML^2$$

And now that we have the two rotational inertias about the same axis, we can use the additive property to get the total rotational inertia:

$$I = I_1 + I_2 = \boxed{\left(\frac{1}{3} + \frac{1}{2\pi^2}\right)ML^2}$$

5.4: Dynamics of Rotating Objects

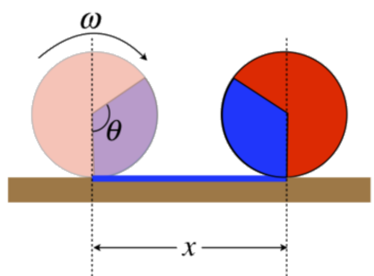
Rolling Without Slipping and Pulleys

A very large number of the mechanical energy conservation problems we will do involve the relation we discussed previously that relates rotational motion to linear motion. Specifically we will apply this to what is referred to as *rolling without slipping*, or *perfect rolling*. There are two reasons this is an important condition to understand:

- When two surfaces slip across each other, thermal energy is the result. So when an object rolls without slipping, there may be static friction present, but there is no kinetic friction, which means that no thermal energy is produced and mechanical energy is conserved.
- When a round object rolls perfectly, the distance it travels in a straight line is directly related to the angle through which it rotates.

We'll keep the first observation in mind for later, but right now let's focus on the second condition:

Figure 5.4.1 – Perfect Rolling



The linear distance traveled equals the arclength of the shaded region if the wheel is rolling without slipping, so we have:

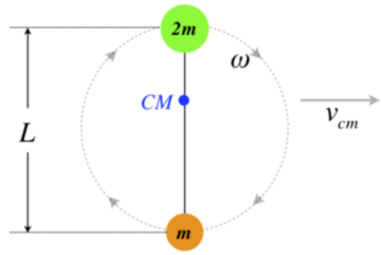
$$x = arclength = R\theta \quad \Rightarrow \quad v = \frac{dx}{dt} = R\frac{d\theta}{dt} = R\omega \tag{5.4.1}$$

Imagine now that instead of this being a wheel, it is a spool that is unwinding. Then the blue line represents string that is coming off the spool. We can therefore conclude also that the relation $v = R\omega$ also applies to the linear speed of a rope that is either unraveling from a rotating spool or passing over a turning pulley.

Total Kinetic Energy as a Sum of Linear and Rotational

It's time we considered the case of an object whose center of mass is moving while it rotates. Let's start with a simple case of two rocks of different masses attached by a string:

Figure 5.4.2 – Unbalanced Dumbell Spinning as It Moves



This system is rotating as its center of mass moves in a straight line (assume there is no gravity present). We are given its rotational speed ω and the velocity of its center of mass, and wish to answer the question, "How much kinetic energy does this system possess at the moment depicted in the diagram?"

We could easily answer this question if we knew the speeds of the two rocks, but we are not given those numbers. We have to extract them from what is given, and this requires some thought. We know three things that get us to this answer:

- The velocity of a rock relative to us equals its velocity relative to the center of mass, plus the velocity of the center of mass (see Section 1.8 for a refresher).
- The center of mass lies at the point two-thirds of the distance from m to $2m$.
- The rotational velocities of both rocks are the same, but the linear velocities relative to the center of mass depend upon their distances from the center of mass according to the usual $v = r\omega$.

Let us label the bottom rock as #1, and the top rock as #2. Putting the first and third conditions together first gives us:

$$v_1 = v_{cm} - r_1\omega \qquad v_2 = v_{cm} + r_2\omega \tag{5.4.2}$$

The sign of the second term in each equation is determined by whether the rotational motion adds to or takes away from the linear motion of the center of mass. Next we invoke the second condition. The fact that the center of mass is two-thirds of the distance from m to $2m$ means:

$$r_1 = \frac{2}{3}L \qquad r_2 = \frac{1}{3}L \tag{5.4.3}$$

Putting all of the above into the kinetic energy of the system gives an expression for the total kinetic energy in terms of the values given. Collecting terms proportional to the squares of center of mass velocity and angular velocity gives:

$$\begin{aligned} KE_{tot} &= KE_1 + KE_2 \\ &= \tfrac{1}{2}mv_1^2 + \tfrac{1}{2}(2m)\,v_2^2 \\ &= \tfrac{1}{2}m\left[v_{cm} - r_1\omega\right]^2 + \tfrac{1}{2}(2m)\left[v_{cm} + r_2\omega\right]^2 \\ &= \tfrac{1}{2}m\left[v_{cm} - \left(\tfrac{2}{3}L\right)\omega\right]^2 + \tfrac{1}{2}(2m)\left[v_{cm} + \left(\tfrac{1}{3}L\right)\omega\right]^2 \\ &= \tfrac{1}{2}(3m)\,v_{cm}^2 + \tfrac{1}{2}\left(\tfrac{2}{3}mL^2\right)\omega^2 \end{aligned} \tag{5.4.4}$$

The $3m$ in the first term is the total mass of the system, so the first term is the kinetic energy of system if was not spinning. That means that the second term is the amount of kinetic energy added to the system by virtue of its spinning. The part of the second term in parentheses looks suspiciously like a rotational inertia, and in fact it equals the rotational inertia of the system about its center of mass:

$$I_{cm} = m_1r_1^2 + m_2r_2^2 = (m)\left(\frac{2}{3}L\right)^2 + (2m)\left(\frac{1}{3}L\right)^2 = \frac{2}{3}mL^2 \tag{5.4.5}$$

This turns out to be a completely general rule for the kinetic energy of an object that is rotating as its center of mass moves:

$$KE_{tot} = KE_{lin} + KE_{rot} = \frac{1}{2}mv_{cm}^2 + \frac{1}{2}I_{cm}\omega^2 \tag{5.4.6}$$

Example 5.4.1

Show the same result (Equation 5.4.6) for two general point masses m_1 and m_2 separated by an unknown distance (call their distances from the center of mass r_1 and r_2), this time using the moment in time when m_1 is directly in front of m_2 (i.e. the line joining them is horizontal).

Solution

At the moment when the two masses form a horizontal line, their linear motions due to rotation are perpendicular to the center of mass motion. Determining their total speeds is therefore a simple application of the Pythagorean theorem, and the result follows surprisingly quickly:

$$\begin{aligned} v_1^2 &= v_{cm}^2 + (r_1\omega)^2 \\ v_2^2 &= v_{cm}^2 + (r_2\omega)^2 \end{aligned}$$

Now plug this into the kinetic energy for the system as the sum of the kinetic energies of the two masses:

$$\begin{aligned} KE_{tot} &= \tfrac{1}{2}m_1v_1^2 + \tfrac{1}{2}m_2v_2^2 \\ &= \tfrac{1}{2}m_1\left(v_cm^2 + r_1^2\omega^2\right) + \tfrac{1}{2}m_2\left(v_{cm}^2 + r_2^2\omega^2\right) \\ &= \tfrac{1}{2}(m_1 + m_2)\,v_{cm}^2 + \tfrac{1}{2}\left(m_1r_1^2 + m_2r_2^2\right)\omega^2 \end{aligned}$$

While the above equation is generally true for any object, if the object is rotating about a fixed point, the expression for total KE can be simpler to write. Specifically, it is what we have written before, in terms of the rotational inertia about the fixed point:

$$KE_{tot} = \frac{1}{2}I_{fixed\ point}\omega^2 \tag{5.4.7}$$

It's not hard to show that this is equivalent to Equation 5.4.6. Assuming the fixed point is not the center of mass (or the assertion is proved trivially), then let's call the distance from the center of mass to the fixed point "d." The center of mass is following a circular path of radius d around the fixed point, which means we can relate the linear velocity of the center of mass to its angular velocity around the fixed point:

$$v_{cm} = \omega d \tag{5.4.8}$$

Putting this into our center-of-mass energy equation gives:

$$KE_{tot} = \frac{1}{2}mv_{cm}^2 + \frac{1}{2}I_{cm}\omega^2 = \frac{1}{2}m(\omega d)^2 + \frac{1}{2}I_{cm}\omega^2 = \frac{1}{2}\underbrace{\left(md^2 + I_{cm}\right)}_{I_{fixed\ point}}\omega^2 \tag{5.4.9}$$

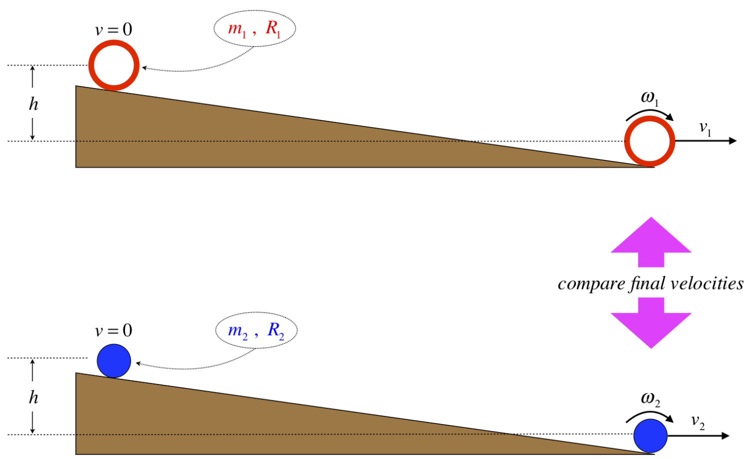
Where in the final step we employed the parallel-axis theorem.

Mechanical Energy Conservation with Perfect Rolling

Let's put together what we have concluded so far in this section. We begin by noting that two cylinders with equal masses do not possess the same rotational inertia about their central axes if one is hollow and the other is solid. Now imagine rolling both of these cylinders (without slipping) down an inclined plane. Can you guess which one would reach the bottom of the incline with the greater speed? The main point to be made here is that the energy that comes from gravitational PE goes into KE, but now the KE has two different forms: linear and rotational. The linear and angular speeds are directly related through the "no slipping" condition, so the energy will convert into the two forms of kinetic energy in a fixed ratio. We will soon see how the rotational inertia affects the ratio, but it seems clear that the hollow cylinder puts more of its energy into rotation (for the same velocity) than the solid cylinder. This would seem to indicate that the hollow will have the same kinetic energy as the solid cylinder only if it is turning (and therefore moving) more slowly.

It's easy to trick oneself in such situations, so let's solve the math carefully to be sure.

Figure 5.4.3 – Comparing Hollow and Solid Cylinder Rolling Dynamics



We will work both problems in parallel, to make the difference more evident. Start with mechanical energy conservation from the top of the plane to the bottom. We can invoke this because without slipping there is no rubbing, which means no mechanical energy is converted to thermal energy.

$$\Delta KE + \Delta PE_{grav} = 0 \quad \Rightarrow \quad KE_o + PE_o = KE_f + PE_f \tag{5.4.10}$$

If we choose the zero point of potential energy to be the bottom of the incline, the initial and final potential energies in both cases are mgh and zero, respectively. The initial kinetic energy is zero in both cases, and the final kinetic energy is the sum of the linear and rotational kinetic energies (Equation 5.4.6):

$$\begin{array}{rll} & hollow\;cylinder & solid\;cylinder \\ energy\;conservation: & m_1gh=\frac{1}{2}m_1v_{1f}^2+\frac{1}{2}I_1\omega_{1f}^2 & m_2gh=\frac{1}{2}m_2v_{2f}^2+\frac{1}{2}I_2\omega_{2f}^2 \\ perfect\;rolling\,(v=R\omega): & m_1gh=\frac{1}{2}m_1v_{1f}^2+\frac{1}{2}I_1\left(\dfrac{v_{1f}}{R_1}\right)^2 & m_2gh=\frac{1}{2}m_2v_{2f}^2+\frac{1}{2}I_2\left(\dfrac{v_{2f}}{R_2}\right)^2 \\ rotational\;inertia\;of\;cylinders: & m_1gh=\frac{1}{2}m_1v_{1f}^2+\frac{1}{2}\left(m_1R_1^2\right)\left(\dfrac{v_{1f}}{R_1}\right)^2 & m_1gh=\frac{1}{2}m_2v_{2f}^2+\frac{1}{2}\left(\frac{1}{2}m_2R_2^2\right)\left(\dfrac{v_{2f}}{R_2}\right)^2 \\ algebra: & v_{1f}=\sqrt{gh} & v_{2f}=\sqrt{\dfrac{4gh}{3}} \end{array} \tag{5.4.11}$$

So in fact the solid cylinder is moving faster than the hollow one, as we predicted. What is especially interesting is that with the perfect rolling condition in place, the masses and radii of the cylinders are irrelevant! We are used to final speeds of objects accelerated by gravity being independent of the mass, but here we see that when we impose perfect rolling, the radius also plays no role, but the distribution of the mass within the cylinder is all that matters.

Alert

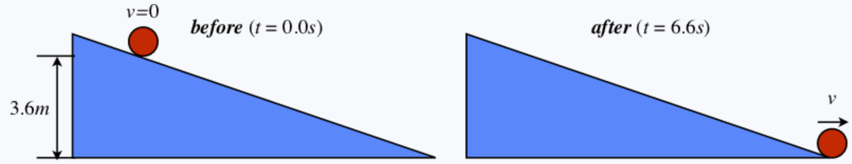
As we are discussing mechanical energy conservation again, it is a good time to remind ourselves that our conclusions only tell us how to compare speeds before and after – what goes on between these two moments and direction of motion are lost bits of information. This is as true now that rotation is involved as it was when it wasn't. For example, if we were to race the two cylinders down identical ramps, then naturally the solid cylinder would get to the bottom first, since they both start at rest and accelerate at constant rates. The object with the faster final speed must have taken less time to get to the bottom because it had a greater average velocity. The math shown above doesn't take into account the paths the two cylinders take, so if the ramps are not identical (but still result in the same height change), the conclusion about speeds at the bottom is the same as before, but the winner of the race may not be the solid cylinder!

Example 5.4.2

A solid uniform sphere starts from rest and rolls without slipping down a slope to a horizontal surface an elevation that is 3.6m below its starting point. It takes 6.6s for it to get down the slope. Find the angle the slope makes with the horizontal.

Solution

We can use energy conservation to most easily determine the final velocity of the sphere, but it is independent of the angle of the incline and the time elapsed. Still, it turns out to be a good starting point, so here is our before/after picture:



The starting KE is zero, and the starting PE is mgh, while the final KE has both a translational and a rotational part, and the final PE is zero, giving the following conservation equation:

$$mgh = \frac{1}{2}mv^2 + \frac{1}{2}I\omega^2$$

We are given that it is a uniform solid sphere, which gives us its rotational inertia in terms of its mass, and we are told that it rolls without slipping, which gives us a relationship between its linear velocity and its rotational velocity. Putting all this together gives us the final velocity:

$$\left.\begin{array}{l} I_{solid\ sphere} = \frac{2}{5}mR^2 \\ \omega = \dfrac{v}{R} \end{array}\right\} \quad \Rightarrow \quad mgh = \frac{1}{2}mv^2 + \frac{1}{2}\left(\frac{2}{5}mR^2\right)\left(\frac{v}{R}\right)^2 \quad \Rightarrow \quad v = \sqrt{\frac{10gh}{7}}$$

Okay, so how can we use this to find the angle? We are given the time it takes to accelerate from rest to this final speed, and the forces remain constant throughout, so the acceleration is also constant. We can therefore use kinematics to determine the distance the sphere travels down the incline. With constant acceleration starting from rest, the average velocity is just half the final velocity, and the distance covered is the average velocity multiplied by the time, so the distance the sphere rolls is:

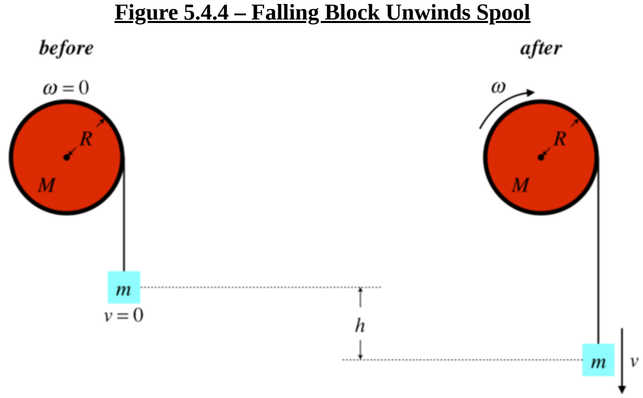
$$\Delta x = v_{ave}t = \left(\frac{0+v}{2}\right)t = \frac{1}{2}\sqrt{\frac{10gh}{7}}t$$

With the starting height and the distance along the plane known, we can use trigonometry to find the angle of the plane:

$$\theta = \sin^{-1}\left(\frac{h}{\Delta x}\right) = \sin^{-1}\left(\frac{2}{t}\sqrt{\frac{7h}{10g}}\right) = \boxed{8.8^o}$$

Another example that falls into this same category of mechanical energy conservation with perfect rolling is a falling mass unwinding a massive spool. Let's assume the spool is frictionless and is a uniform disk, and determine the speed of the falling block after it has dropped a known distance. We are also assuming – as always – that the string is massless, but we should also point out that it is very thin, so that its departure from the spool does not reduce the radius of the spool.

Figure 5.4.4 – Falling Block Unwinds Spool

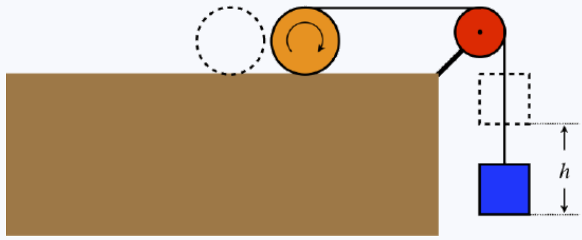


Once again, we can solve this using mechanical energy conservation, as there are no non-conservative forces present. What is new here is that some of the potential energy lost by the block as it drops goes into the rotational kinetic energy of the spool. The math is strikingly similar to the rolling cylinder case above:

$$\begin{aligned} energy\ conservation: &\quad mgh = \tfrac{1}{2}mv^2 + \tfrac{1}{2}I\omega^2 \\ perfect\ rolling\,(v = R\omega): &\quad mgh = \tfrac{1}{2}mv^2 + \tfrac{1}{2}I\left(\frac{v}{R}\right)^2 \\ rotational\ inertia\ of\ spool: &\quad mgh = \tfrac{1}{2}mv^2 + \tfrac{1}{2}\left(\tfrac{1}{2}MR^2\right)\left(\frac{v}{R}\right)^2 \\ algebra: &\quad v = \sqrt{\frac{4mgh}{2m+M}} \end{aligned} \tag{5.4.12}$$

Example 5.4.3

One end of a massless rope is wound around a uniform solid cylinder, while the other end passes over a massless, frictionless pulley and is attached to a hanging block with the same mass, as in the diagram below. The block is released from rest, pulling the cylinder along the horizontal surface such that it rolls without slipping. Find the speed of the block and the linear speed of the cylinder after the block has fallen a distance h.



Solution

The amount that the block falls equals the distance traveled by the cylinder plus the length of rope that unwinds from it. Since the cylinder rolls without slipping, the amount that unwinds is also equal to the distance it travels, so the sum of the distance traveled by the cylinder and the rope unwound is just double the distance that the cylinder travels. Therefore, the speed of the block is at all times twice the linear speed of the cylinder.

With no non-conservative forces present, the mechanical energy of the system is conserved, so:

$$mgh = \frac{1}{2}mv_{block}^2 + \frac{1}{2}mv_{cylinder}^2 + \frac{1}{2}I\omega^2$$

We know the rotational inertia of the cylinder in terms of its mass and radius, that the block moves twice as fast as the cylinder, and that the cylinder rolls without slipping. Putting all of these constraints into the equation above gives us our answer:

$$\left.\begin{aligned} I &= \tfrac{1}{2}mR^2 \\ v_{block} &= 2v_{cylinder} \\ R\omega &= v_{cylinder} \end{aligned}\right\} \Rightarrow \quad mgh = \frac{1}{2}m(2v_{cylinder})^2 + \frac{1}{2}mv_{cylinder}^2 + \frac{1}{2}\left(\frac{1}{2}mR^2\right)\omega^2 = \frac{11}{4}mv_{cylinder}^2 \quad \Rightarrow \quad \boxed{v_{cylinder} = \sqrt{\frac{4gh}{11}}}$$

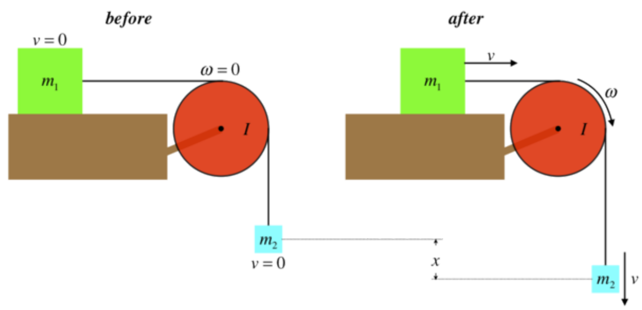
The speed of the block is twice this much:

$$\boxed{v_{block} = \sqrt{\frac{16gh}{11}}}$$

Massive Pulleys

The result for this example may remind you of an assumption we made long ago regarding pulleys. We have always assumed that they were frictionless and massless. We said that the result of these assumptions was that the tension for the rope was the same everywhere (namely on both sides of the pulley). We are now equipped to look at what happens if the pulley has mass. We'll do so with a simple model physical system. In Figure 5.4.5, the hanging block accelerates as it falls, linearly accelerating the block on the frictionless horizontal surface and rotationally accelerating the pulley in the process.

Figure 5.4.5 – Effect of a Massive Pulley on Rope Tension



We are interested in comparing the tension force by the rope on both sides of this pulley, so let's use the work-energy theorem, which takes into account the forces. Treating each block as a separate system, on which the tension in each end of the rope performs work (and gravity does work on block #2 as well), and noting that both move at the same speed at all times, we have:

$$\begin{array}{lcr} W_1=\Delta KE_1 & \Rightarrow & T_1\cdot x=\frac{1}{2}m_1v^2 \\ W_2=\Delta KE_2 & \Rightarrow & (m_2g-T_2)\cdot x=\frac{1}{2}m_2v^2 \end{array} \tag{5.4.13}$$

Let's compare the two tensions by computing the difference:

$$(T_1-T_2)\cdot x=\underbrace{\frac{1}{2}m_1v^2+\frac{1}{2}m_2v^2}_{\Delta KE_{blocks}}\underbrace{-m_2gx}_{\Delta PE_{grav}} \tag{5.4.14}$$

For the tensions to be equal, all of the gravitational potential energy lost by the falling block must go into the two blocks. But we now know that a massive pulley will have kinetic energy. Let's add the pulley's increase in kinetic energy to both sides of the equation, and invoke mechanical energy conservation:

$$(T_1-T_2)\cdot x+\left[\frac{1}{2}I\omega^2\right]=\Delta KE_{blocks}+[\Delta KE_{pulley}]+\Delta PE_{grav}=0 \quad\Rightarrow\quad T_2-T_1=\frac{\frac{1}{2}I\omega^2}{x} \tag{5.4.15}$$

The tensions can only be equal when the rotational inertia of the pulley is zero, which means it must be massless.

Digression: Energy Storage

One of the big issues today with green energy like solar and wind-generated electricity, is storage. The advantage to fossil fuel production of electricity is that we can produce it whenever we like, but for solar and wind power, we are at the mercy of when the sun shines or the wind blows. So storing the energy generated from these green sources is of paramount importance. Batteries are coming along, but they have their own environmental issues (lithium mining, waste when they degrade, etc.), so other means of storage are sought.

There are many ideas that have been put forth, such as using spare electricity to pump water above a dam so that it can be released when needed; pressurizing tanks of air with spare electricity, then allowing the pressurized air to drive a generator later; and using spare electricity to desalinate water, followed by using the osmotic pressure between the new fresh water and the salty water to drive a generator. But possibly the best idea (which has been around a long time) is to simply store the energy in the form of kinetic energy – spin a flywheel. A flywheel is just a disk created for the sole purpose of spinning so that it holds kinetic energy until it can be used later. The idea is for the spare electricity to get this thing spinning (with as little friction as possible), so that later when we need the energy back, the flywheel can be connected to a generator and the kinetic energy can be converted back into electrical energy. The beauty of this idea is in its simplicity – it is inexpensive and scalable. And reducing the friction to a very low value is something we can do quite well with today’s technology (think maglev and evacuated chambers). In order to be as efficient with our use of space as possible, and so that we don’t reach rotational speeds that are insanely high, we will of course want flywheels with very large rotational inertias.

Swinging Around Fixed Points

There is one other common physical situation involving mechanical energy conservation and rotation that needs to be addressed. If a rigid extended object is pivoted around a fixed point that is not the center of mass, and it is allowed to swing around that pivot under the influence of gravity, then how do we use mechanical energy conservation to describe its motion? Specifically, as the object swings, some points of the object may move upward (increasing gravitational potential energy), while others may swing downward (decreasing gravitational potential energy). How can we deal with the overall change in gravitational potential energy in such a case?

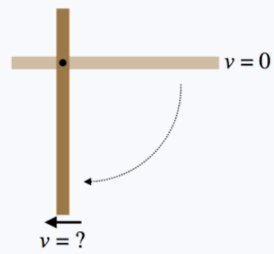
The answer will likely be unsurprising. Write the change of potential energy of the whole object as the sum of the potential energies of each tiny mass that makes up the object, and the result follows immediately:

$$\begin{aligned}\Delta U\,(whole\ object) &= \Delta U_1 + \Delta U_2 + \ldots \\ &= m_1 g \Delta y_1 + m_2 g \Delta y_2 + \ldots \\ &= \frac{Mg}{M}(m_1 \Delta y_1 + m_2 \Delta y_2 + \ldots) \\ &= Mg\frac{m_1 \Delta y_1 + m_2 \Delta y_2 + \ldots}{M} = Mg\Delta y_{cm}\end{aligned} \tag{5.4.16}$$

where $M = m_1 + m_2 + \ldots$ is the mass of the whole object and Δy_{cm} is the change in height of the center of mass of the object.

Example 5.4.4

A uniform rod of length 0.70m is pivoted around a point that is one-fourth of its length from one end, and is released from rest in a horizontal orientation. Find how fast the bottom end of the rod is moving when the rod reaches a vertical orientation.



Solution

The rod is uniform, so its center of mass is at its center, which means that during this swing it drops a distance of one fourth the length of the rod. All of the potential energy lost becomes kinetic energy, and all of the kinetic energy can be written as rotational if we use the rotational inertia around the fixed point, so:

$$Mg\left(\frac{L}{4}\right) = \frac{1}{2}I\omega^2$$

Now we need the rotational inertia of the rod about the fixed point. We can get this from the parallel axis theorem:

$$I_{pivot} = I_{cm} + Md^2 = \frac{1}{12}ML^2 + M\left(\frac{L}{4}\right)^2 = \frac{7}{48}ML^2$$

Plugging this in above gives us ω:

$$\frac{MgL}{4} = \frac{1}{2}\left(\frac{7}{48}ML^2\right)\omega^2 \quad \Rightarrow \quad \omega = \sqrt{\frac{24g}{7L}}$$

Now all we need to do is relate the rotational speed to the linear speed of the bottom end of the rod:

$$v = r\omega = \left(\frac{3}{4}L\right)\sqrt{\frac{24g}{7L}} = \sqrt{\frac{27}{14}gL} = \boxed{3.6\frac{m}{s}}$$

5.5: Torque

Rotational Newton's Second Law

As we saw for linear motion, we can only go so far with energy conservation. If we want to analyze aspects of motion such as elapsed time and direction of motion, we need more than mechanical energy conservation to work with. In the linear case, we found that this meant that we had to use Newton's Second Law. We now seek the rotational equivalent of that law.

The rotational equivalent of the Newton's Second Law must relate the reaction of the system (rotational acceleration) to an external influence (rotational force), with the degree of this effect being determined by an internal property of the system (rotational mass). That is, we need a rotational substitute for all of the participants of this formula:

$$\overrightarrow{a}_{cm} = \frac{\overrightarrow{F}_{net}}{m} \tag{5.5.1}$$

We already found a rotational version of acceleration in our discussion of rotational kinematics – it is the angular acceleration. We even defined a direction for this vector using the right-hand rule. The center of mass qualification in the case above is unneeded for the rotational case, because the angular acceleration is the same about every point on a rigid object.

We have also determined an appropriate candidate for the "rotational mass" – the rotational inertia. This is certainly a reasonable choice, for a couple of reasons. First, from our direct experience we know that it is easier to swing an object (e.g. a baseball bat) when holding the heavier end than when holding the lighter end, so the degree to which an extended object "resists" angular acceleration is determined by the distribution of mass. Second, if the physics is to remain consistent, why would the quantity that plays the role of mass in kinetic energy be different from the quantity that plays the role of mass for the second law?

With those two quantities established, we can now get a glimpse into what the "rotational force" is by examining the units:

$$[\alpha] = \frac{[rotational\ force]}{[I]} \quad \Rightarrow \quad [rotational\ force] = \left[\frac{rad}{s^2}\right]\left[kg \cdot m^2\right] = \frac{kg \cdot m^2}{s^2} \tag{5.5.2}$$

This is weird... These are units of energy! We'll need to chalk this up to coincidence, since clearly the vector quantity of rotational force cannot be a measure of energy. One way to see the difference is to remember the presence of radians in the numerator, even though they are not physical units. We will soon see the source of this coincidence, and it shouldn't take long before the apparent ambiguity between this quantity and energy fades away.

Alert

While the physical units are the same as energy, we ***never*** *refer to the SI units of this quantity as "joules." Using this term implies that we are talking about energy, which we are not. Generally we stick to "Newton-meters."*

We can't continue calling this vector "rotational force" forever, so we will henceforth refer to it by its proper name: *torque*. In keeping with our tradition of using greek variables for rotational quantities, we will represent torque with $\overrightarrow{\tau}$, giving as our rotational Newton's second law:

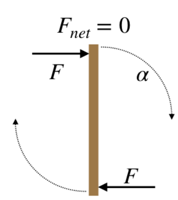
$$\overrightarrow{\alpha} = \frac{\overrightarrow{\tau}_{net}}{I} \tag{5.5.3}$$

Torque

In the cases of acceleration and inertia, we found a direct relationship between the linear and rotational quantities, so we would expect there to be a similar relationship between force and torque. Furthermore, since the linear/rotational bridge for acceleration and inertia both require a point of reference (the pivot), we would expect the same to be true for the bridge between force and torque.

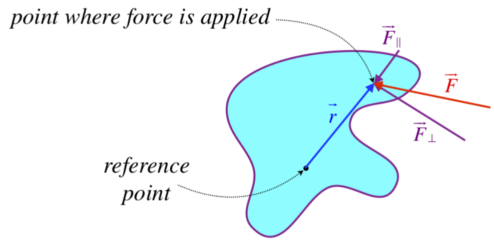
The first thing we notice is that an object can experience no net force and yet still experience a nonzero rotational acceleration:

Figure 5.5.1 – Zero Net Force Can Accelerate Rotationally



If the two forces shown in in Figure 5.5.1 are moved so that act at the same point on the object, then it's clear that they also cancel rotationally. So apparently the place *where* the force acts is important to computing torque. If we choose a reference point (we will refer to this as a "pivot" in cases when it happens to be a fixed point, but in general it does not), then the application point of a force can be described by a position vector $\overrightarrow{r}$ that points from the reference point to the point where the force is applied. But there is still more that we have to worry about here. If two forces with the same magnitudes as those in Figure 5.5.1 were applied at the same points on the bar, but were pointing vertically, then no angular acceleration would result. Let's put all this together...

Figure 5.5.2 – Parts of a Force that Cause Angular Acceleration



The force vector can be decomposed into two perpendicular vectors – one that is parallel to the position vector, and one perpendicular to it. When it comes to causing the object to accelerate its rotation around the pivot, it's clear that the part of the force that is parallel to the position vector $\overrightarrow{F}_\parallel$ will have no effect, while the perpendicular part of the force $\overrightarrow{F}_\perp$ will.

If we were to perform experiments to test the effects of various force magnitudes, we would find that the angular acceleration is proportional to the magnitude of the force – push twice as hard in the same direction as the same point on the object, and its angular acceleration is twice as great around the same pivot. If we were to perform further experiments to test the effects of applying the force at different positions, we would find that the angular acceleration is proportional to the magnitude of the position vector – extend the position vector in the same direction to twice its original length and apply the same force in the same direction, and the angular acceleration is once again twice as great around the same pivot. Mathematically, we express the results of these experiments this way:

$$\left|\overrightarrow{\tau}\right| \sim \left|\overrightarrow{r}\right| \left|\overrightarrow{F}\right| \tag{5.5.4}$$

Notice that the units of this product work out correctly, so all we need to do is incorporate the "only the perpendicular part of $\overrightarrow{F}$ has an effect" into the math. If we call the angle between the position vector and the force vector θ, then the perpendicular component is $F\sin\theta$. Assuming there are no other constants involved (and there aren't any), we get, for the magnitude of the torque:

$$\left|\overrightarrow{\tau}\right| = \left|\overrightarrow{r}\right| \left|\overrightarrow{F}\right| \sin\theta \tag{5.5.5}$$

This looks familiar – we actually saw something just like it, way back in Equation 1.2.9. Torque *is* a vector that is derived from the product of two other vectors. Is it possible that it is simply a cross-product of these two vectors? The magnitude works, but what about direction? In Figure 5.5.2, the force will accelerate the rotation counterclockwise, which means that according to the right-hand-rule, the acceleration vector points out of the page. If we perform a cross-product of the position vector (up to the right) and the force vector (up to the left), the right-hand-rule results in a vector that also points out of the page. We therefore write:

$$\overrightarrow{\tau} = \overrightarrow{r} \times \overrightarrow{F} \tag{5.5.6}$$

Example 5.5.1

A rigid object is pivoted around the origin. The force vector given below acts on this object at the position also indicated below. Find the torque vector exerted on the object due to this force.

$$\overrightarrow{F} = 1.5N\;\hat{i} + 0.80N\;\hat{j} - 2.4N\;\hat{k}, \qquad position:\;(x,\;y,\;z) = (3.0m,\;0.0m,\;-2.0m)$$

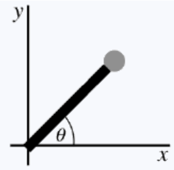
Solution

This is a straightforward calculation of a cross product:

$$\begin{aligned}\overrightarrow{\tau} &= \overrightarrow{r} \times \overrightarrow{F} \\ &= \left(3.0m\;\hat{i} + 0.0m\;\hat{j} - 2.0m\;\hat{k}\right) \times \left(1.5N\;\hat{i} + 0.80N\;\hat{j} - 2.4N\;\hat{k}\right) \\ &= \left[(3.0m)(0.80N) - (0.0m)(1.50N)\right]\hat{k} + \left[(0.0m)(-2.40N) - (-2.0m)(0.80N)\right]\hat{i} + \left[(-2.0m)(1.50N) - (3.0m)(-2.4N)\right]\hat{j} \\ &= \boxed{1.6Nm\;\hat{i} + 4.2Nm\;\hat{j} + 2.4Nm\;\hat{k}}\end{aligned}$$

Example 5.5.2

A small marble is attached to the end of a thin rigid rod of length L, whose other end is held fixed at the origin. The rod lies in the $x-y$ plane, and makes an angle θ up from the x-axis, as shown in the diagram.



The marble (but not the rod) is subjected to a force that gives rise to a potential energy field given by:

$$U(x,y)=\beta xy$$

Find the magnitude and direction of the torque exerted on the rod relative to the origin, in terms of L, θ, and β.

Solution

We can use the potential energy function to determine the force at every point in space:

$$\overrightarrow{F}=-\frac{\partial U}{\partial x}\,\hat{i}-\frac{\partial U}{\partial y}\,\hat{j}=-\beta\left(y\,\hat{i}+x\,\hat{j}\right)$$

The torque exerted relative to the origin at the point (x,y) is the cross-product of the position vector there and the force vector there:

$$\overrightarrow{\tau}=\overrightarrow{r}\times\overrightarrow{F}=\left(x\,\hat{i}+y\,\hat{j}\right)\times\left[-\beta\left(y\,\hat{i}+x\,\hat{j}\right)\right]=\beta\left(y^2-x^2\right)\hat{k}$$

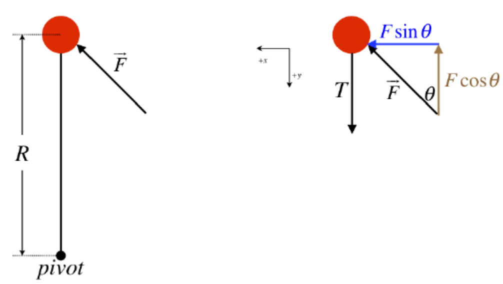
Now plug in for the coordinates of the marble in terms of L and θ:

$$\overrightarrow{\tau}=\beta L^2\left(\sin^2\theta-\cos^2\theta\right)\hat{k}=\boxed{-\beta L^2\cos 2\theta\,\hat{k}}$$

Linking Rotational and Linear

Let's do a sanity check on our definition of torque and it role in the rotational second law. We can do it very simply by choosing a single point mass tied to a string whose other end is held as a fixed pivot (we'll leave gravity out of this). We'll start with the linear version of Newton's second law, and translate it into the rotational version.

Figure 5.5.3 – A Simple System Solved Two Ways



The forces in the x and y directions provide two equations through Newton's second law:

$$a_x=\frac{\sum F_x}{m}\quad\Rightarrow\quad a_{\parallel}=\frac{F\sin\theta}{m}\tag{5.5.7}$$

$$a_y=\frac{\sum F_y}{m}\quad\Rightarrow\quad a_{\perp}=\frac{T-F\cos\theta}{m}\tag{5.5.8}$$

Now we translate to rotational motion by first converting the parallel part of the acceleration into angular acceleration:

$$a_{\parallel}=R\alpha\tag{5.5.9}$$

Then convert mass into rotational inertia:

$$m=\frac{I}{R^2}\tag{5.5.10}$$

Plugging Equation 5.5.9 and Equation 5.5.10 into Equation 5.5.7 gives:

$$R\alpha=\frac{F\sin\theta}{\frac{I}{R^2}}\quad\Rightarrow\quad\alpha=\frac{FR\sin\theta}{I}=\frac{\tau}{I}\tag{5.5.11}$$

One important thing to note here is that while the torque and rotational inertia depend upon the pivot point (i.e. they are different values if we use a new reference point), the translation between the angular acceleration and linear acceleration exactly balances this difference. For example, if we replace the pivot defined above with a new one that is a distance $2R$ from the object, all of the math works out exactly the same. That is, the torque is twice as great and the rotational inertia is four times as great, resulting in a rotational acceleration that is half as large as before, but when it is multiplied by twice the radius to get the linear acceleration, the same result occurs, as it must.

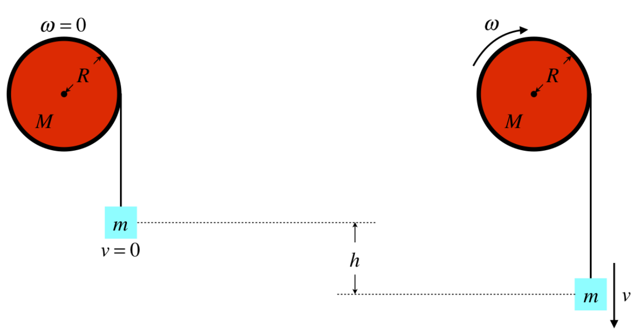
Solving Problems

Now we can do a whole set of problems involving torque causing rotational acceleration. There are many similarities with solving problems involving linear forces and accelerations, but here are some differences:

- Free-body diagrams now require that forces be placed appropriately on the objects, since torque depends upon force placement (no more using dots to represent the object).
- There usually is no need to resolve the torque vector into components. In fact, most problems can limit torque (and angular acceleration) to just "clockwise" and "counterclockwise" – the direction of the torque vector can be left until the end.
- One must either know or be able to calculate the rotational inertia of the object on which the torques acts.
- The perfect rolling condition is sometimes applied.

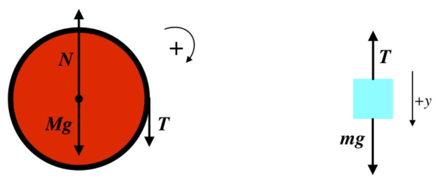
To get an idea of the process, we'll re-work the problem shown in Figure 5.5.4, this time using rotational second law instead of energy conservation:

Figure 5.5.4 – Falling Block Unwinds Spool (Redux)



Start with free-body diagrams:

Figure 5.4.5 – FBD's of Block and Spool



Next we need to right down the equations for Newton's second law for each object. The block is moving in a straight line, so we are already familiar with that one:

$$a_y = \frac{\sum F_y}{m} \quad \Rightarrow \quad a = \frac{mg - T}{m} = g - \frac{T}{m} \tag{5.5.12}$$

The spool is rotating, so we need to use the rotational version for it. Before we can sum the torques for the spool, we need to select a reference point, and its axle is a pretty obvious choice. The length of the position vector from this reference point to the where the gravity and normal forces act is zero, so those forces produce no torque around the axle (which makes sense – pushing on an axle should not cause something to spin around it). This leaves on the the tension force. It acts tangent to the spool, so this force is perpendicular to the position vector connecting the pivot to the point where the force acts, which makes the magnitude of torque it produces equal to simply the product of the tension and the radius of the spool. The direction of this torque is positive, since it causes a clockwise acceleration and our FBD defines that as the positive direction. As this is the only torque, it is the net torque, and we have:

$$\alpha = \frac{\tau_{net}}{I} \quad \Rightarrow \quad \alpha = \frac{T \cdot R}{I} \tag{5.5.13}$$

Now we have to incorporate our constraints (our "other information"). We know that the spool is a uniform solid disk with mass M, giving us its rotational inertia. Also, we know that the rate at which the string exits the spool is related to the rotation rate of the spool according to the usual "no slipping" condition, so we have an equation relating the block's linear acceleration a to the spool's angular acceleration α:

$$I = \frac{1}{2}MR^2, \qquad \alpha = \frac{a}{R} \tag{5.5.14}$$

Putting these constraints into Equation 5.4.13 and combining this with Equation 5.4.12 gives:

$$\left.\begin{array}{l} \dfrac{a}{R} = \dfrac{T \cdot R}{\frac{1}{2}MR^2} \quad \Rightarrow \quad T = \dfrac{Ma}{2} \\ a = g - \dfrac{T}{m} \end{array}\right\} \quad \Rightarrow \quad a = \frac{2m}{2m+M}g \tag{5.5.15}$$

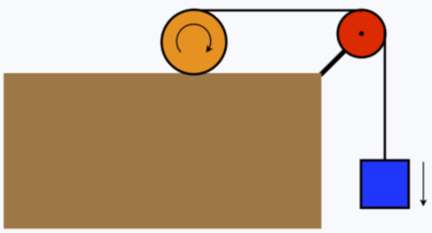
We see that the acceleration of the block is constant, so we can use a kinematics equation to determine the velocity after displacing a distance h from rest:

$$v_f^2 - v_o^2 = 2a\Delta y \quad \Rightarrow \quad v = \sqrt{\frac{4mgh}{2m+M}g} \tag{5.5.16}$$

This agrees with our previous answer.

Example 5.5.3

Let's take another look at Example 5.3.3 (see the diagram below). As the block falls, it pulls the uniform, solid cylinder (with a mass equal to that of the block), which rolls without slipping. If the horizontal surface was frictionless, then clearly this would be impossible. If it was very nearly frictionless, the disk would still slip. This means there must exist a minimum coefficient of static friction for which this physical situation can occur. Compute this minimum coefficient of static friction.



Solution

Let's get the constraints out of the way, as we already determined these in Example 5.3.3. The block moves twice as fast as the cylinder at all times, which means that it accelerates linearly at twice the rate of the cylinder as well:

$$v_{block} = 2v_{cylinder} \quad \Rightarrow \quad a_{block} = 2a_{cylinder} \equiv 2a$$

The rolling-without-slipping condition provided us with an equation relating the cylinder's linear velocity to its angular velocity. It naturally applies to the accelerations as well:

$$v_{cylinder} = R\omega \quad \Rightarrow \quad a_{cylinder} = R\alpha$$

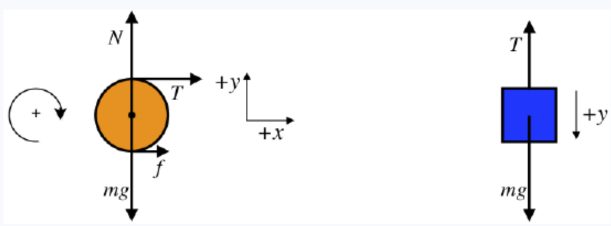
The cylinder is solid and uniform, so its rotational inertia is:

$$I = \frac{1}{2}mR^2$$

The only other constraint we have here is one we have encountered many times before – the cylinder will just barely not slip when the static friction force equals the coefficient of static friction multiplied by the normal force between the cylinder and the surface:

$$f = \mu_s N$$

Okay, now let's tackle the equations that come from Newton's second law. We of course start with force diagrams:



You might ask how we know that the friction force points in the direction indicated in the diagram. Technically, we don't yet know this, but we don't have to. If, in the course of our calculations, we find that the only way a solution can work out is if the value of f is negative, then the friction force must point the other way. We will see shortly that the direction on the diagram is in fact the only direction it can point.

There are three equations that come from Newton's second law for the cylinder (the horizontal and vertical linear net force equations, and the net torque equation), and there is one equation that comes out for the block:

$$\begin{array}{lll} & cylinder & block \\ x-direction: & a=\dfrac{T+f}{m} & \\ y-direction: & 0=N-mg & 2a=\dfrac{mg-T}{m} \\ torques: & \alpha=\dfrac{TR-fR}{I} & \end{array}$$

Plugging in for the rotational inertia and the angular acceleration gives:

$$\frac{a}{R}=\frac{TR-fR}{\frac{1}{2}mR^2} \quad \Rightarrow \quad \frac{a}{2}=\frac{T-f}{m}$$

Adding this equation to the x-direction equation for the cylinder gives:

$$\frac{3}{2}a=\frac{T+f}{m}+\frac{T-f}{m} \quad \Rightarrow \quad T=\frac{3}{4}ma$$

Now combine this result with the y-direction equation for the block to get:

$$2a=\frac{mg-\frac{3}{4}ma}{m} \quad \Rightarrow \quad a=\frac{4}{11}g$$

The y-direction equation for the cylinder and the friction/normal force constraint gives:

$$f=\mu_s mg$$

Solving for the friction force using the x-direction equation for the cylinder, and then plugging in what we found for the tension yields:

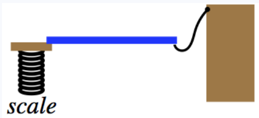
$$f=ma-T=ma-\frac{3}{4}ma=\frac{1}{4}ma$$

Combining these last three lines gives us our answer at last:

$$\mu_s mg=\frac{1}{4}ma=\frac{1}{4}m\left(\frac{4}{11}g\right) \quad \Rightarrow \quad \mu_s=\boxed{\frac{1}{11}}$$

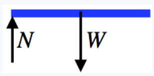
Example 5.5.4

A uniform rod weighing $32N$ is at rest horizontally, and is supported at one end by a scale. The support at the other end of the rod is suddenly removed, and the rod begins to fall & rotate. Find the force measured by the scale at the instant that the rod is released.



Solution

Start with a free-body diagram:



Naturally the normal force measured by the scale does not equal the weight of the rod, because the center of mass of the rod is accelerating. At the moment of release, the end of the rod on the scale is not free to move downward, so while this will not remain the case, at this moment that end of the bar is fixed, and the remainder of the bar experiences an angular acceleration around that point. We therefore have two sets of equations from Newton's second law – one for linear motion, and one for rotational motion. Setting downward and clockwise as the positive linear and rotational directions, and choosing the contact point with the scale as the reference, we have:

$$\begin{array}{lll} linear: & W-N=ma_{cm} & \Rightarrow \quad N=W-ma_{cm} \\ rotational: & W\left(\frac{1}{2}l\right)=I_{rod\ about\ end}\ \alpha=\frac{1}{3}ml^2\alpha & \Rightarrow \quad \alpha=\dfrac{3}{2ml}W \end{array}$$

The center of mass is half the length of the rod from the fixed point, so the linear acceleration of the center of mass is related to the angular acceleration according to:

$$a_{cm}=\frac{l}{2}\alpha$$

Plugging this back in above gives our answer:

$$N=W-m\left(\frac{l}{2}\alpha\right)=W-\frac{3}{4}W=\frac{1}{4}W=\boxed{8N}$$

Rotational Work

We have now discussed the rotational version of energy conservation and Newton's second law, so the link between these two topics – the work-energy theorem – should follow naturally. Rather than provide a derivation (which would really just resemble what we have done before for the linear case), we'll just write down the answer that makes sense from following our linear/rotational parallel.

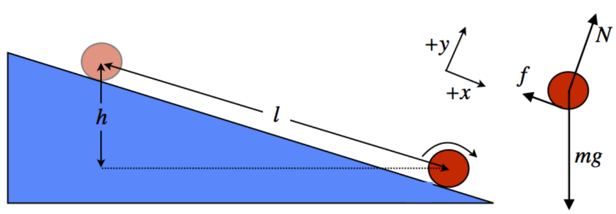
$$\begin{aligned} W_{A\rightarrow B}\,(linear) &= \int_A^B \overrightarrow{F}\cdot\overrightarrow{dl} &= \Delta KE &= \tfrac{1}{2}mv_B^2-\tfrac{1}{2}mv_A^2 \\ W_{A\rightarrow B}\,(rotational) &= \int_A^B \overrightarrow{\tau}\cdot\overrightarrow{d\theta} &= \Delta KE &= \tfrac{1}{2}I\omega_B^2-\tfrac{1}{2}I\omega_A^2 \end{aligned} \tag{5.5.17}$$

If we were so inclined, we could do the same unwinding-the-spool problem for a third time, this time with the rotational work-energy theorem. The approach looks slightly different, but when you actually sit down to do it, you see the same things come out of it as before. This time instead of relating the accelerations, we would relate the distance the mass drops to the angle the spool rotates.

Back when we discussed objects rolling down an inclined plane without slipping, we avoided talking about one potentially confusing point that we are now equipped to deal with. For a ball or cylinder to roll down, there has to be a friction force (otherwise it would merely slide). This friction force can only be static friction, because we are assuming there is no slipping, and we said that without any rubbing, the mechanical energy must be conserved. But this friction force acts *up the plane while the object moves down it*, which means that it does negative work on the object. This would seem to imply that mechanical energy should not be conserved, so how were we able to make the assumption that it is conserved?

The answer is, "Because the static friction force also does *positive* rotational work which adds energy to the object in rotational form, and this addition exactly balances the loss in linear form." This is not hard to prove. Start with a diagram and a FBD:

Figure 5.5.6 – Work Done on Cylinder by Static Friction as It Rolls Down Plane



Computing the work done by static friction for linear motion is very simple, since the friction force is constant and the motion is in a straight line:

$$W\left(x_1\rightarrow x_2\right)=\int_{x_1}^{x_2}\overrightarrow{F}\cdot\overrightarrow{dx}=-f\cdot l \tag{5.5.18}$$

As expected, this work takes energy out of the cylinder system. Next we compute the rotational work done on the cylinder. The torque is a constant equal to fR, and is acting in the same direction as the rotational displacement, so

$$W\left(\theta_1\rightarrow\theta_2\right)=\int_{\theta_1}^{\theta_2}\overrightarrow{\tau}\cdot\overrightarrow{d\theta}=+(fR)\cdot\theta \tag{5.5.19}$$

Putting these together gives us the total work done on the cylinder by the static friction force. Note that since it rolls without slipping, the linear distance it travels is related to the angle through which it rotates by the usual relation:

$$W_{static\ friction}=W\left(x_1\rightarrow x_2\right)+W\left(\theta_1\rightarrow\theta_2\right)=-f\cdot l+fR\theta=f\left(-l+R\theta\right)=0 \tag{5.5.20}$$

So we see that in fact the work done by static friction here only serves to convert linear kinetic energy into rotational kinetic energy, and our understanding of how thermal energy is generated remains intact.

Rotational Power

We spoke before about how sometimes we are interested in the rate at which work is done, calling this value “power.” Well, as with everything else we studied in linear motion, there is of course a rotational version:

$$P = \frac{dW}{dt} = \overrightarrow{\tau} \cdot \overrightarrow{\omega} \tag{5.5.21}$$

You sometimes hear the silly “debate” about torque vs. horsepower for car & truck engines. This should make it clear what the difference is. Power delivered to the wheels is directly related to torque exerted on them, but it is dependent upon how fast they are turning. Engines that can still produce a lot of torque at high speeds are powerful. To get an idea of why it might be hard to maintain torque at high speeds, imagine pedaling a bike downhill – when you get going fast enough, it’s difficult to push hard on (provide torque to) your pedals. So generally the effectiveness of an engine is defined by torque at low speeds and power at high speeds. If you want fast acceleration off the line or the ability to pull a stump out of the ground, you want torque. If you want to go fast or tow a heavy trailer up a hill at a steady speed, you want power.

5.6: Static Equilibrium

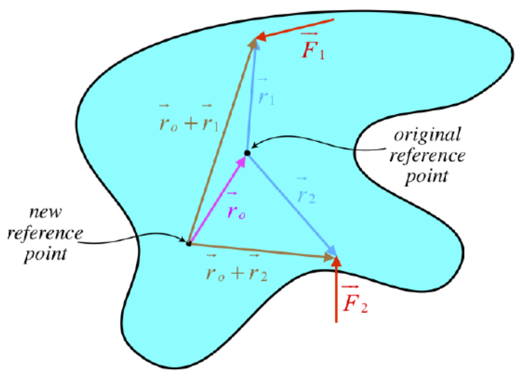
Pivots and Torque Reference Points

The definition of torque (Equation 5.5.6) includes the position vector $\overrightarrow{r}$, which points from a reference point to the point where the force is applied. When we are interested in how the torque is accelerating the object rotationally around a fixed point ("pivot"), it is convenient to choose the reference point to be that fixed point. This is because the forces applied at that fixed point (to keep it fixed) provide zero torque when referenced there, and those forces are generally not known. We explore here the effect of changing the reference point in the particular case when there is *no net force*, though perhaps there could be a net torque. The net torque around a given reference point is:

$$\overrightarrow{\tau}_{net} = \overrightarrow{r}_1 \times \overrightarrow{F}_1 + \overrightarrow{r}_2 \times \overrightarrow{F}_2 + \dots \tag{5.6.1}$$

The reference point is located at the tails of the $\overrightarrow{r}_i$ vectors, but suppose we want to change that reference point. We can do this by simply adding the same constant vector $\overrightarrow{r}_o$ to every position vector. This has the effect of shifting the reference point from the point of $\overrightarrow{r}_o$ to its tail, as shown in Figure 5.6.1 [*Note: The figure shows only two of the many forces applied.*]

Figure 5.6.1 – Changing the Reference Point



The net torque around this new reference point is:

$$\begin{aligned} \overrightarrow{\tau}_{net}(new) &= \left(\overrightarrow{r}_o + \overrightarrow{r}_1\right) \times \overrightarrow{F}_1 + \left(\overrightarrow{r}_o + \overrightarrow{r}_2\right) \times \overrightarrow{F}_2 + \dots \\ &= \left[\overrightarrow{r}_1 \times \overrightarrow{F}_1 + \overrightarrow{r}_2 \times \overrightarrow{F}_2 + \dots\right] + \overrightarrow{r}_o \times \left[\overrightarrow{F}_1 + \overrightarrow{F}_2 + \dots\right] \\ &= \overrightarrow{\tau}_{net}(original) + \overrightarrow{r}_o \times \overrightarrow{F}_{net} \end{aligned} \tag{5.6.2}$$

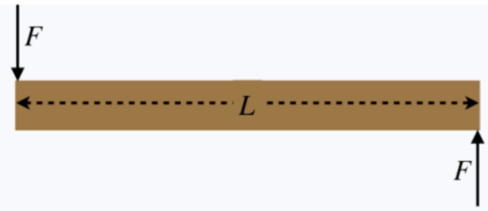
But we assumed that the net force was zero, so we get the remarkable result that the *net torque is the same around every reference point*!

Alert

As amazing as this result is, be careful not to mistake it for too general of a result. The net torque on an object by a collection of forces is only independent of the reference point if those forces result in zero net force.

Example 5.6.1

A non-uniform board of length L starts at rest and is pushed in opposite directions on both of its ends with forces of equal magnitude F at right angles to the board (see diagram). The forces continue to be applied at right angles with the same magnitude until the board has made a complete 360^o rotation, which occurs after a time T. Find the smallest possible rotational inertia for the board when it is rotated around an axis perpendicular to the page.



Solution

The smallest rotational inertia occurs when the axis of rotation is through the center of mass, so we are looking for I_{cm}, but we don't know where the center of mass is for this board, as it is non-uniform. However, in the physical situation given above, there is no net force on the board, so its center of mass, which started at rest, remains at rest, which means the board is rotating around its center of mass. If we can determine the torque about the center of mass and the angular acceleration of the board, then we can use the rotational second law to obtain the rotational inertia:

$$I = \frac{\tau_{net}}{\alpha}$$

We can get α from kinematics, since we know how far the board turns, how long it takes, and the fact that it started from rest:

$$\left.\begin{array}{l} \theta = \omega_o t + \frac{1}{2}\alpha t^2 \\ \theta = 2\pi \\ \omega_o = 0 \\ t = T \end{array}\right\} \;\Rightarrow\; \alpha = \frac{4\pi}{T^2}$$

Okay, so we know the acceleration of the board around its center of mass, and all we need is the net torque around that point, but we don't know where the center of mass is, so how are we supposed to continue? The net force is zero, which means we can choose any reference point, and it will give us the torque measured around every reference point, including the center of mass. Choosing the reference point to be one of the ends of the board, the torque due to the force on that end is zero, while the torque due to the force on the other end is simply FL, and the sum of these two torques is the net torque. Using this as the net torque around the center of mass gives us our answer:

$$\boxed{I = \frac{FLT^2}{4\pi}}$$

Note: One can pick a spot on the board and label it as the center of mass, calling the distance from one end x, which makes the distance from the other end equal to $L-x$. Then the torques can be written around the center of mass, and we'll find that the x's will cancel, giving the same result. But why go to all that trouble?

Static Equilibrium

We have spent a great deal of time studying motion in all its forms, but now we're going to step back and look at something called *static equilibrium*. Simply put, this means unmoving (static), and not about to move (equilibrium). This is a particularly important subject for engineers who aspire to build things that won't (easily) fall down. From Newton's laws for linear and rotational motion, we have two conditions for the equilibrium part of this condition:

- net force on object is zero
- net torque on object is zero

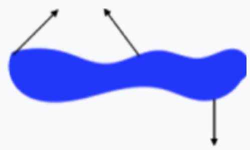
We are quite familiar with the net force part of this, but we need to do a bit of work on net torque. We know the formal definition of torque, but there is more we need to understand in order to apply this to static equilibrium problems. The first tool that we can immediately add to our toolbox for solving such problems is the result we got above. If the object is in static equilibrium, then it is experiencing zero net force, which means that no matter what reference point we choose, the net torque will be the same. But the net torque is *zero* for equilibrium, so we will have the following condition to work with:

For objects in static equilibrium, the net torque calculated around any reference point whatsoever is zero.

We will find the flexibility to choose any point we like as a reference to point to be very useful in what is to come.

Example 5.6.2

For the force diagram below, the force vectors are drawn in the proper locations on the object, and are pointing in the proper directions, but the lengths of the vectors are not to scale. Which of the following statements are true about the effects these forces can have on the motion of this object? Assume that none of the force magnitudes can be set to zero.



a. *The force magnitudes can be set so that the object will not accelerate rotationally, while at the same time its center of mass does not accelerate linearly.*
b. *There is no way to set the force magnitudes to prevent either linear or rotational acceleration.*
c. *The force magnitudes can be set so that either there is no linear acceleration of the object's center of mass, or there is no rotational acceleration of the object, but both cannot be achieved at the same time.*
d. *The force magnitudes can be set so that the object's center of mass will not accelerate linearly, but there is no way to prevent its rotational acceleration.*
e. *The force magnitudes can be set so that the object will not accelerate rotationally, but there is no way to prevent linear acceleration of its center of mass.*

Solution

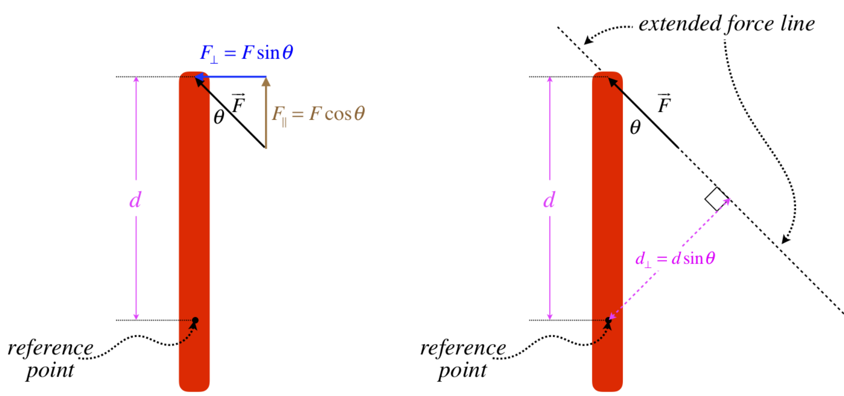
d

The two force vectors can be adjusted relative to each other so that their horizontal components cancel. Then both of their magnitudes can be adjusted in the same proportions so that the horizontal net force remains zero, while their combined vertical component of force cancels the other force vector. So zero net force is achievable. However, if we consider a reference point where the middle force acts on the object (giving that middle force zero contribution to torque), the torque of the other two forces will never cancel, no matter what adjustments are made to the force magnitudes. With no way to make the torque vanish, there is no way to prevent rotational acceleration.

Using Geometry to Determine Torque

Our definition of torque is all well-and-good, but in practice we rarely define a position vector and take a cross product. Instead, we tend to use the concept behind torque, and then some geometry. Figure 5.6.2 shows two ways to geometrically get to the same torque due to an applied force.

Figure 5.6.2 – Alternative Methods of Computing Torque



The left version consists of taking only the component of force that is perpendicular to the line joining the reference point and the point where the force is applied, giving the torque magnitude calculation:

$$\tau = F_\perp d = (F\sin\theta)\, d \tag{5.6.3}$$

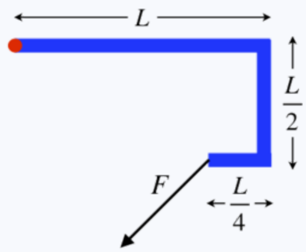
The right side of the figure shows another useful way to compute the same torque magnitude. Rather than finding the part of $\overrightarrow{F}$, it involves finding the perpendicular part of $\overrightarrow{r}$. This is done by extending the line of force and then geometrically determining the perpendiular distance from the reference point to that line. The result is the same as above:

$$\tau = Fd_{\perp} = F\left(d\sin\theta\right) \tag{5.6.4}$$

The perpendicular distance from the reference point to the line of force is often referred to as the *moment-arm*, or *lever-arm*. We will find this to often be the method of choice of computing torques when it comes to solving problems.

Example 5.6.3

What can you say about the torque applied to the object due to the force F about the red pivot in the diagram?



a. *it equals $\frac{1}{2}FL$*
b. *it equals $\frac{1}{4}FL$*
c. *it is greater than $\frac{1}{2}FL$*
d. *it is less than $\frac{1}{4}FL$, but greater than zero*
e. *net torques always sum to zero*

Solution

c

There are a couple of ways to answer this. This first is to extend the force line along F and look at the perpendicular distance from the pivot to that line (this is the moment arm). It should be clear from the geometry that this moment arm exceeds $\frac{1}{2}L$, which means the torque must be greater than $\frac{1}{2}FL$. Another way to see it is to break F into two separate vectors, point pointing left and the other pointing down. Both of these forces produce clockwise torques, and the horizontal force has a moment arm of $\frac{1}{2}L$, while the vertical force has a moment arm of $\frac{3}{4}L$. Since the sum of these two force components exceeds the magnitude of the original force, and since one of them has a moment arm larger than $\frac{1}{2}L$, then the combined torques must exceed $\frac{1}{2}FL$.

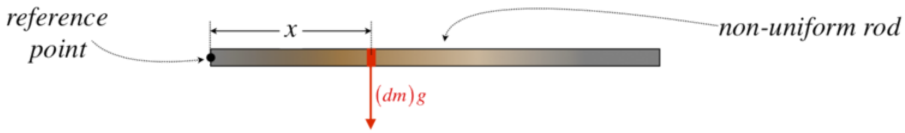
Center of Gravity

Up to now, whenever we have drawn a force diagram of an object, we have always placed the force vector for gravity at its center, while other forces are placed wherever they happen to act on the object. Gravity is somewhat special in that the force actually acts on every single atom in the object, but we can't draw all of those individual force vectors. Drawing it at the center of mass makes sense from the standpoint of Newton's second law, since if gravity is the only force, then it accelerates the object, and the part of the object that accelerates is the center of mass.

Wherever it happens to be appropriate to locate a single gravity force vector on a free-body diagram, it is called the object's *center of gravity*. We are currently dealing with torque, and the position at which a force acts has become quite important, so we need to examine more closely whether we can declare the center of mass of an object to be its center of gravity.

We choose as our test subject a horizontal non-uniform rod of length L, and select one of its ends as a reference point. The plan is to add up all of the infinitesimal torques that occur about this reference point due to gravity acting on every particle in the rod, and see if this total torque can be replaced by the entire gravity force acting at a single point (so that we can draw our free-body diagrams with only one gravity force vector!). An arbitrary piece of the rod will be a distance x from the reference point, and the torque exerted there will be the weight of that piece multiplied by x:

Figure 5.6.3 – Center of Gravity of a Non-Uniform Rod



$$d\tau = (dm\ g)\,x \quad \Rightarrow \quad \tau = \int_0^L dm\ gx = Mg\left[\frac{1}{M}\int_0^L dm\ x\right] = Mgx_{cm} \tag{5.6.5}$$

Sure enough, we get the same torque around the reference point if we put a single force vector with magnitude Mg (the object's full weight) acting at the object's center of mass.

Alert

It should be mentioned that there was a rather subtle assumption made in the above discussion – the gravity force is assumed to be the same at all points on the rod. If the gravity force can somehow vary from one end of the rod to the other, then the positions of these two centers will not coincide. If you are wondering how this can ever be the case, the answer is that the scale of distances must be very large, so that there are measurable differences in the gravity force from one point on the object to the other. This will not be an issue for our typically terrestrially-constrained studies, but can arise when talking about orbits of large bodies like moons.

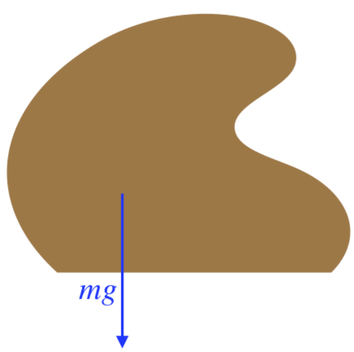
Note that like center of mass, the center of gravity of an object does not have to lie on the object. For example, a hoop's center of gravity is located in the empty space at its center. We now know how to locate the position of the gravity force on an object, and locating most other forces will be fairly intuitive (with one notable exception, which we will address next). This will enable us to use torque to analyze a whole range of real-world problems.

Placement of the Normal Force

Like the gravity force, the normal force can act at many places at once. When two surfaces come into contact, all of the particles at one surface repel all of the particles at the other. So once again we have the problem of where to draw a single force vector, this time for the normal force. The normal force is different from the gravity force, in one important way – it just *compensates* for other forces. That is, it adjusts according to other circumstances. Let's use what we know about static equilibrium to see how to place the normal force properly.

Consider the oddly-shaped object shown in Figure 5.6.4. We'll assume that the object sits at rest on a horizontal surface, the density of this object is not uniform, and that the center of gravity is at the position indicated in the diagram.

Figure 5.6.4 – Deducing the Normal Force Placement Balancing Only Gravity



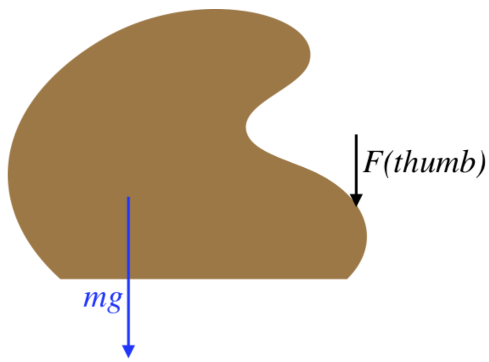
A (rather unsystematic) process for locating the position of the normal force goes like this:

1. Note that the object is in static equilibrium, which means that the normal force is equal to the weight (net force is zero), and that the net torque around any reference point we care to choose is also zero.
2. Try various positions for the normal force, and if we can prove that there must be a non-zero net torque around a reference point, then throw that position out.
3. Repeat step (2) for various positions until one is found that cannot be ruled-out.

For the object in Figure 5.6.4, we could try a normal force vector acting at the center of the base of the object. Then if we choose a reference point between the normal force vector and the weight vector, see see that those two forces must produce a non-zero counter-clockwise torque. We can similarly rule out any position to the right of the weight vector. If we try a position to the left of the weight vector, we get a similar result, this time with the torque being clockwise. We therefore conclude that for this case the normal force vector must be applied exactly where the weight vector intersects the base. No matter where we choose a reference point in that case, the two forces result in equal-and-opposite torques.

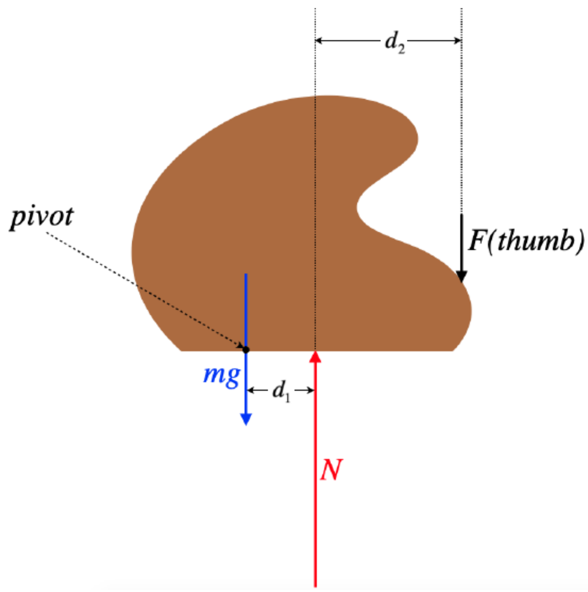
Let's complicate matters some by introducing a second force to our object – suppose we push down on the right side of the object with our thumb, as shown in Figure 5.6.5.

Figure 5.6.5a – Deducing the Normal Force Placement Balancing Two Forces



Let's try the same position for the normal force as before – in line with the gravity force. If we choose as a reference a point in line with these two forces, then they create no torque between the two of them, and the added force by the thumb creates a net clockwise torque. This isn't possible for an object in static equilibrium, so the normal force placement has moved from its original placement as a result of the added thumb force. It's easy to see that the normal force hasn't moved left, as placing the reference point at the normal force results in both the weight and the thumb force producing clockwise torques. So the normal force must move right, but how far? Perhaps it moves into line with the thumb force? No... We can choose the reference point to be in line with these two forces (so they both create zero torque), and the gravity force would yield a net counterclockwise torque.

Figure 5.6.6b – Deducing the Normal Force Placement Balancing Two Forces



So we conclude that the normal force must act somewhere between the gravity and thumb forces. If we know the magnitudes of these two forces, then we know the magnitude of the normal force (the net force is zero), and in fact we can also determine precisely where a single normal force is acting on the object. Calling the distance between the weight and normal force vector placements d_1 and the distance between the normal force and thumb force vector placements d_2, we can sum the torques around a reference point where the normal force acts (so it contributes no torque) to get:

$$0 = \tau_{net} = -mgd_1 + Fd_2 \quad \Rightarrow \quad \frac{d_1}{d_2} = \frac{F}{mg} \tag{5.6.6}$$

[*Note: In the torque sum, clockwise was chosen as the positive direction.*]

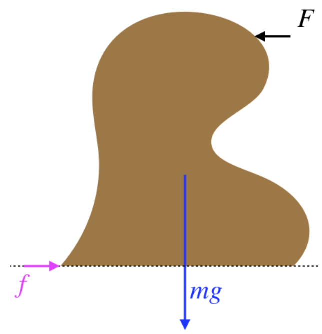
In the diagram the weight is shown to be greater than the thumb force, making the ratio less than 1, which means the placement of the normal force is closer to the placement of the weight vector than the thumb force vector. If the thumb pushes down more, then the normal force placement moves to the right. Note also that the same result arises if the reference point is chosen elsewhere. For example, if the reference point is chosen to be where the weight force acts, then the net torque equation gives zero contribution from the weight, and contributions from both the normal force (counterclockwise), and the thumb force (clockwise). The normal force can then be written in terms of the weight and thumb force (the net force is zero), giving:

$$0 = \tau_{net} = -Nd_1 + F\left(d_1 + d_2\right) \quad \Rightarrow \quad 0 = -\left(mg + \cancel{F}\right)d_1 + \cancel{Fd_1} + Fd_2 \quad \Rightarrow \quad \frac{d_1}{d_2} = \frac{F}{mg} \tag{5.6.7}$$

Conditions for Tipping

Let’s make a slight change to the situation just described. Suppose I push horizontally on the top of the object. What happens to the normal force position as the magnitude of the push increases? Assuming there is a static friction force to prevent the object from sliding, we have a free-body diagram (missing the normal force) that looks like this:

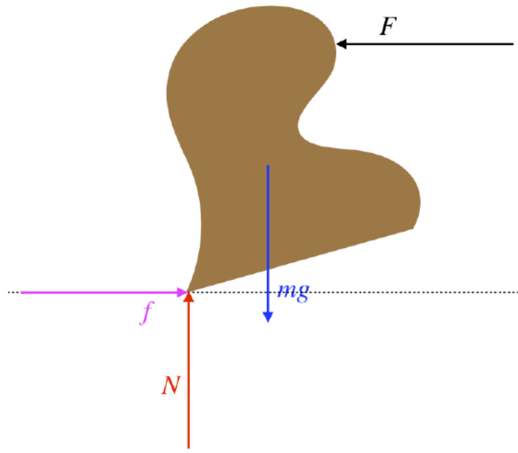
Figure 5.6.6c – Deducing the Normal Force Placement Balancing Two Forces



Choosing a pivot point at the intersection point of the gravity and friction forces, we see that the push force exerts a net counterclockwise torque. For the normal force to counteract this (and given that it must push straight up), we find that it must be placed to the left of the center of gravity.

Let's take a moment to consider the magnitudes of these forces. So long as the object doesn't slide, the static friction force must equal the push. The object doesn't accelerate up or down, so the normal force must have the same magnitude as the gravity force. Both of these conditions are important when we consider what happens when the push force is increased. The friction force also increases until it hits its maximum, at which point the object starts sliding. If we suppose the static friction force doesn't hit its maximum, how is the increased torque by the push compensated by the normal force, if it can't change magnitude? It must move left. But it can only move left for so long, and when it has gone as far as it can go, any added push results in angular acceleration – the object tips.

Figure 5.6.7 – Tipping



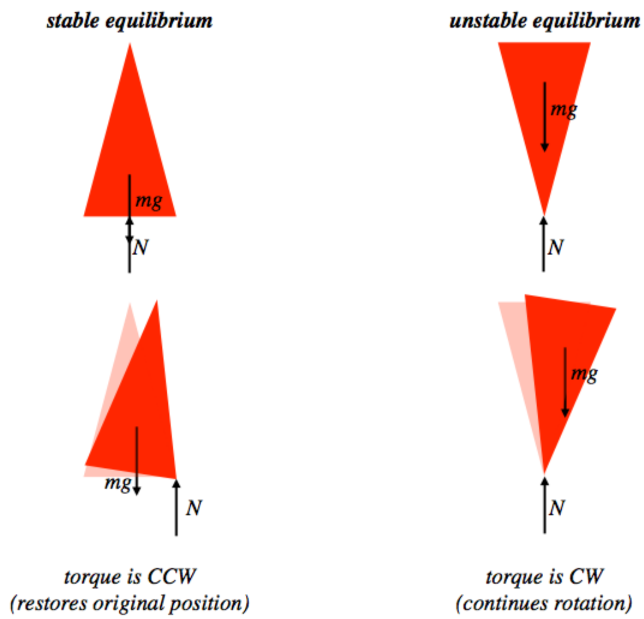
Suppose you want to push an object across the floor without tipping it over. To get it to slide, you have to push with a force at least equal to the static friction force, so to avoid tipping, this given amount of force needs to provide as little toque as possible – push at a point close to the bottom. With very little torque from the push force, the normal force can easily remain inside of the edge of the object, and the object won't tip before it slides.

Stable/Unstable Equilibrium

If the object is oddly-shaped and the only forces acting on it are gravity and the normal force, then this analysis gives us an answer as to whether the object falls over – if the normal force can be directly beneath the center of gravity, then it will stand up. By "can," of course we mean that some part of the base that is in contact with the surface (where the normal force acts) must be below the center of gravity.

In Section 3.7 we discussed stable and unstable equilibrium from the perspective of energy diagrams, and the concept of whether an equilibrium is stable or unstable was first addressed. The idea is that if the system is moved slightly from its equilibrium state, do the forces (or, in our current case, torques) act to return the system to equilibrium (stable), or to continue moving the system away from equilibrium (unstable). How these definitions apply to tipped objects is clear from the free-body diagrams, as shown in Figure 5.6.8. We can also define a *degree* of stability to a standing object. We define it as the angle through which we can rotate it such that if we let it go, restoring torques act to return it to its original position.

Figure 5.6.8 – Stable and Unstable Equilibrium



Note that this definition of stability matches with what we saw in energy diagrams. Recall that an equilibrium was a point where the potential energy function has zero slope, and the equilibrium is stable if the potential energy grows on both sides of the equilibrium, and is unstable if the potential energy falls off on both sides. Consider what happens to the gravitational potential energy of the object in both cases shown in Figure 5.6.8. In the stable case, tipping the object slightly *raises* the center of mass of the object (increasing its gravitational potential energy), while in the unstable case a slight tip *lowers* the center of gravity (decreasing its gravitational potential energy).

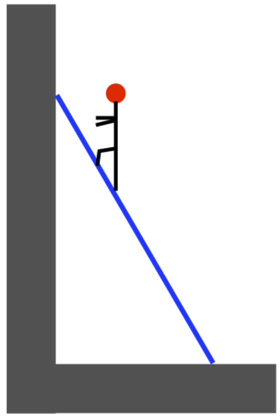
Problem-Solving

Problems involving static equilibrium can be approached in a very systematic way, the steps of which are outlined below:

1. Determine the object in static equilibrium you need to analyze and isolate it in a force diagram. This can sometimes be easier said than done. Sometimes the problem involves more than one extended object in contact with each other. In this case, determining the object you choose (or perhaps the combination of both objects) depends upon what you are asked to solve for (usually a force). You can't really go wrong here, though – if you choose an object that will not give you the answer you need, it should occur to you as you draw the force diagram. Also, you may find that a "wrong" choice of object may simply make your task a bit longer (more simultaneous equations) – annoying, but no real harm done.
2. Define a linear (x, y) coordinate system for force components, and a rotational coordinate system (positive rotation direction) for torques.
3. Extend each force vector with a dotted line as far as it will go on the page in both directions.
4. Choose a reference point. For now we won't worry about choosing a "good" one, choose any – but stick with it for the remaining steps. When you get better at these problems (which you can only achieve by doing them, especially if you do the same problem in multiple ways), you will get better at choosing convenient reference points. Please note that not all static equilibrium problems involve hinges or other "natural" pivots – The reference point doesn't need to be one of these!
5. Ignoring distractions like the object itself, use geometry to find the perpendicular distance from every force line to the pivot point (i.e. all the moment arms). Do not worry about what angle you use to find these (i.e. it doesn't have to be the angle from the torque equation $\tau = rf\sin\theta$) – just use geometry.
6. Multiply the moment arm by the magnitude of each force, and this is the magnitude of the torque due to each force.
7. Determine whether each torque is clockwise or counterclockwise, and give each the appropriate sign when summing the torques and setting that sum equal to zero for the equilibrium torque condition. Note that you could simply alternatively add up all the CW torques, place them on one side of the equation, and set them equal to the sum of the CCW torques on the other side of the equation. This is easier to implement, but loses the "flavor" of what equilibrium is (zero net torque), so I describe it both ways.
8. If you are lucky (or were clever at the outset), this equation may be all you need in order to find what you are looking for. If it isn't, you have two alternatives from here...
   - Write out the sum of the forces in the x and y directions (or just one of those directions, if that is all you need), and set the net forces equal to zero (another condition for equilibrium). These additional equations should be all you need to find what you are looking for.
   - Choose a new reference point and repeat the torque method described above. Recall that the torques should sum to zero around any point, so this is completely valid. The thing to keep in mind is that wherever you choose your reference point, if a force line goes through it, then that force won't appear in the torque equation because the moment arm for it is zero. Therefore you can choose your reference point at a spot through which lines for unknown forces pass, eliminating the need to eliminate them using simultaneous equations later. Whatever you do, don't choose a reference point that lies along a line of the force that you are actually looking for – it doesn't give you an equation that includes that force!

Let's look at some examples of this process in action with a couple of variants on a problem we'll call "Man Climbs Ladder"...

Figure 5.6.9 – Man Climbs Ladder



The following information is given:

- The uniform ladder makes a 60^o angle with the floor.
- The weight of the man is $700N$.
- The weight of the ladder is $150N$.
- The wall is frictionless, but the floor is not.

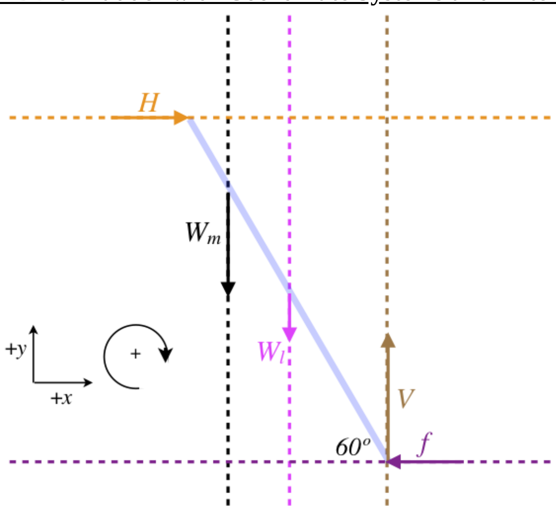
We will hold off for now on what is being asked, as this is sufficient for drawing a free-body diagram.

Alert

While it might seem unlikely, this is actually quite an effective approach to problem-solving in general, but it is especially true for static equilibrium problems: Start analyzing the situation without yet worrying about "where you are going." By removing the focus from a goal to fully understanding the situation, and getting the early steps out of the way, the solution tends to present itself more readily.

In this case, the object in equilibrium experiencing many different forces and torques is the ladder, so we isolate that. We will also include steps (2) and (3) given above (coordinate systems and extended force lines) for our diagram:

Figure 5.6.10 – FBD of Ladder with Coordinate Systems and Extended Force Lines



The vertical normal force by the floor has been labeled V, and the horizontal normal force by the floor is H. The wall is frictionless, so it provides no vertical component of force, and the friction force by the floor is called f. The weight of the ladder is W_l, and since the ladder is uniform, its center of gravity is its geometric center. The force of the man's foot on the ladder equals the weight of the man, and is called W_m. We don't yet know where this last force is placed on the ladder.

Version 1: Man is 75% of the Way Up, Find the Friction Force

We continue our problem solution with step (4) – picking a reference point. The point where the ladder meets the floor seems like a pretty natural place, so let's try that. We are not given the length of the ladder, so we will call it L. Now we need the moment arms (perpendicular distance from reference point to the line of force) for all the forces. Clearly it is zero for V and f. The right triangles formed by the force lines allow us to determine the moment arms for the other three forces easily:

$$\begin{array}{ll} \text{force} & \text{moment arm} \\ wall\;\;(H): & L\sin 60^o \\ man\;\;(W_m): & \frac{3}{4}L\cos 60^o \\ ladder\;weight\;\;(W_l): & \frac{1}{2}L\cos 60^o \end{array} \tag{5.6.8}$$

The force from the wall causes a clockwise (positive) torque, while the other two produce counterclockwise (negative) torques, so the torque equation for equilibrium is:

$$0=\tau_{net}=+H\left(L\sin 60^o\right)-W_m\left(\frac{3}{4}L\cos 60^o\right)-W_l\left(\frac{1}{2}L\cos 60^o\right)\;\;\Rightarrow\;\;H=346N \tag{5.6.9}$$

[*Note how the length of the ladder drops out of the equation.*]

This equation does not give us an answer for the friction force, but now that we have H, we can use the zero net force in the horizontal direction to get f immediately. The friction force is the only one opposing H, so it too equals 346N.

Before we move on to version #2, let's try using a different reference point, but this time give some thought to what reference point might save us a little bit of work. The overriding consideration is to pick a point that gives zero moment arm for the forces that we neither know, nor care to know. In this problem, those forces are H and V – every other force is either given (W_m and W_l) or desired (f). But those two force lines don't intersect on the ladder. ***It doesn't matter!*** The torques will sum to zero around any reference point whatsoever, and the intersection point of those two force lines works just fine. Calculating the non-zero moment arms from that point to the three force lines, we get:

$$\begin{array}{ll} \text{force} & \text{moment arm} \\ man\;\;(W_m): & \frac{3}{4}L\cos 60^o \\ ladder\;weight\;\;(W_l): & \frac{1}{2}L\cos 60^o \\ friction\;\;(f): & L\sin 60^o \end{array} \tag{5.6.10}$$

Putting these into the zero net torque equation gives us an equation with only f as an unknown, which gives us an instant solution:

$$0=\tau_{net}=-W_m\left(\frac{3}{4}L\cos 60^o\right)-W_l\left(\frac{1}{2}L\cos 60^o\right)+f\left(L\sin 60^o\right)\;\;\Rightarrow\;\;f=346N \tag{5.6.11}$$

Version 2: Coefficient of Static Friction with Floor is 0.4, Find Percentage of Ladder the Man Can Climb

This is a tougher version of the problem, because a constraint is thrown in. From Equation 5.6.11, it's clear that the higher the man climbs, the greater the static friction force needs to be to balance the torques. The ladder will start to slide when the friction force exceeds its maximum, so we are working with the constraint:

$$f_{max}=\mu_s V \tag{5.6.12}$$

The same free-body diagram applies, except that we now need to describe the distance up the ladder that the man has climbed when the maximum friction force is attained. We'll describe this distance in terms of the fraction of the total length of the ladder that the man is from the top, and we'll call the fraction β.

Okay, now we need to choose a reference point. There is no way to solve the whole problem with one equation, but we still don't care about the force from the wall (we are using both forces from the floor this time), so let's choose the point where the ladder touches the wall. The moment arms are:

$$\begin{array}{ll} \text{force} & \text{moment arm} \\ man\;\;(W_m): & \beta L\cos 60^o \\ ladder\;weight\;\;(W_l): & \frac{1}{2}L\cos 60^o \\ friction\;\;(f): & L\sin 60^o \\ floor\;\;(V): & L\cos 60^o \end{array} \tag{5.6.13}$$

The forces by the man, the weight of the ladder, and friction all cause positive torques, and the vertical force by the floor causes a negative torque, so the zero net torque equation is:

$$0=\tau_{net}=+W_m\left(\beta L\cos 60^o\right)+W_l\left(\frac{1}{2}L\cos 60^o\right)+f\left(L\sin 60^o\right)-V\left(L\cos 60^o\right) \tag{5.6.14}$$

The length of the ladder again cancels out, but still three unknowns remain in this equation. We can compute the vertical floor force from the zero net force in the y-direction:

$$0=-W_m-W_l+V \quad\Rightarrow\quad V=850N \tag{5.6.15}$$

We get the friction force from the constraint equation:

$$f_{max}=\mu_s V \quad\Rightarrow\quad f=340N \tag{5.6.16}$$

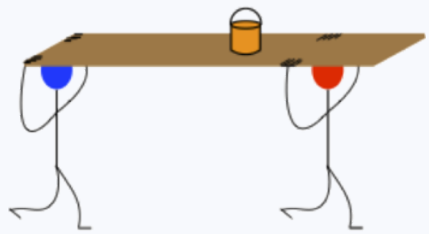
V and f can now be plugged into the torque equation to get our answer:

$$\beta=0.266 \tag{5.6.17}$$

This is the fraction of ladder from the top, so the man can climb 73.4% of the way up before the ladder starts to slide across the floor.

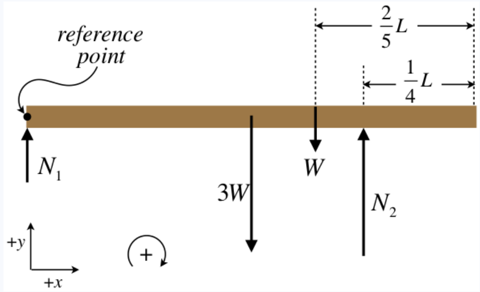
Example 5.6.4

Two painters carry a plank of plywood that they use for scaffolding over their heads on their way to the job site. The plank has a uniform mass distribution. On top it is a can of paint weighing one third as much as the plank. The painter in the rear is holding the plank at the very end and the painter in front is holding the plank one quarter of the the plank length from the front. The can of paint is two-fifths of the plank length from the front. Find the percentage of the total weight carried by the painter in front. Assume that the plank is horizontal as they carry it.



Solution

We start by identifying the object in equilibrium (the plank), and drawing a free-body diagram for it (we'll call the length of the plank L). We will choose the pivot to be the back of the plank, and will refer to the weights of the can of paint and plank as W and 3W, respectively. Also we have chosen an (x,y) coordinate system and the positive direction of rotation to be clockwise, as shown in the diagram.



Next apply the conditions of equilibrium. Clearly the x-direction forces are not meaningful, and the y-direction force equation and torque equations are:

$$\begin{aligned} &vertical\ forces: && 0=N_1-3W-W+N_2 \\ &torques: && 0=N_1(0)+3W\left(\tfrac{1}{2}L\right)+W\left(\tfrac{3}{5}L\right)-N_2\left(\tfrac{3}{4}L\right) \end{aligned}$$

The L's cancel out of the torque equation, resulting in a relation between the force exerted by the front painter and the weight of the can:

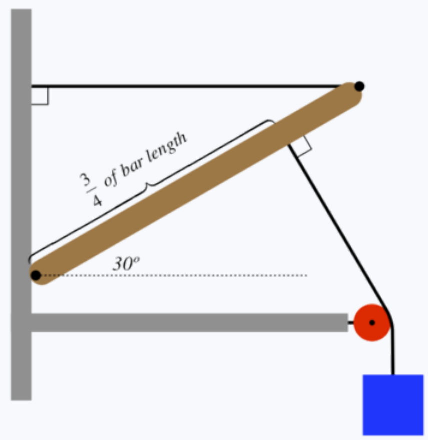
$$N_2=\frac{14}{5}W$$

The total weight carried by the two painters is found from the force equation (or from common sense), and equals $4W$. So the percentage of the total weight carried by the front painter is:

$$\frac{N_2}{4W}=\boxed{70\%}$$

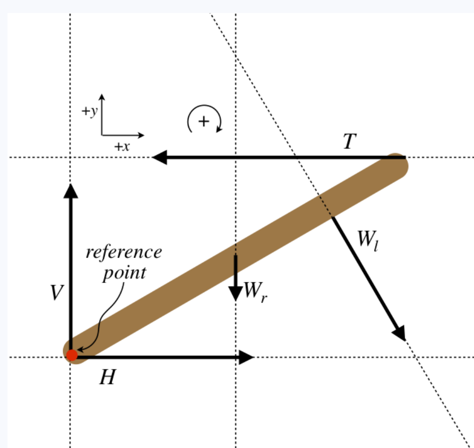
Example 5.6.5

A uniform rigid rod that weighs $400N$ is hinged on a vertical wall and connected to a support wire and a hanging load weighing $1250N$ as shown in the diagram below. Find the magnitude of the force exerted on the rod by the hinge.



Solution

We start with a free-body diagram of the rod, including a coordinate system, a positive direction of rotation, extended force lines, and a reference point:



Next we determine the moment arms. The hinge forces have zero moment arms, and in terms of the length of the rod (which we will call L), the other moment arms are:

force	moment arm
$weight\ of\ rod\ (W_r):$	$\frac{1}{2}L\cos 30^o$
$weight\ of\ load\ (W_l):$	$\frac{3}{4}L$
$tension\ (T):$	$L\sin 30^o$

Now for the zero-net-torque equation. The two weights cause clockwise torque, and the tension counterclockwise torque, so:

$$0=\tau_{net}=W_r\left(\frac{1}{2}L\cos 30^o\right)+W_l\left(\frac{3}{4}L\right)-T\left(L\sin 30^o\right)$$

The L cancels out of every term, and we can solve for the tension:

$$T=2221N$$

To get the horizontal and vertical forces by the hinge, we need the vertical and horizontal zero-net-force equations:

$$\begin{array}{ll} x-direction: & 0=H-T+W_l\sin 30^o \\ y-direction: & 0=V-W_r-W_l\cos 30^o \end{array}$$

Solve for H and V and combine them to get the magnitude of the force by the hinge:

$$\left.\begin{array}{l} H=1596N \\ V=1483N \end{array}\right\} \;\Rightarrow\; F_{hinge}=\sqrt{H^2+V^2}=\boxed{2179N}$$

CHAPTER OVERVIEW

6: Angular Momentum

Thumbnail: A gyroscope is a device used for measuring or maintaining orientation and angular velocity. It is a spinning wheel or disc in which the axis of rotation (spin axis) is free to assume any orientation by itself. When rotating, the orientation of this axis is unaffected by tilting or rotation of the mounting, according to the conservation of angular momentum. Image used with permission (Public Domain; LucasVB).

6.1: Linking Linear and Angular Momentum

Rotational Impulse-Momentum Theorem

By now we have a very good sense of how to develop the formalism for rotational motion in parallel with what we already know about linear motion. We turn now to momentum. Replacing the mass with rotational inertia and the linear velocity with angular velocity, we get:

$$\overrightarrow{p} \equiv m\overrightarrow{v} \quad\Longleftrightarrow\quad \overrightarrow{L} \equiv I\overrightarrow{\omega} \tag{6.1.1}$$

The vector L is called *angular momentum*, and it has units of:

$$[L] = \frac{kg\cdot m^2}{s} = J\cdot s$$

Continuing the parallel with the linear case, the momentum is relates to the force through the impulse-momentum theorem, which is:

$$\int_{t_A}^{t_B} \overrightarrow{F}_{net}dt = \Delta\overrightarrow{p}_{cm} \quad\Longleftrightarrow\quad \int_{t_A}^{t_B} \overrightarrow{\tau}_{net}dt = \Delta\overrightarrow{L} \tag{6.1.2}$$

While there is no need to append "*cm*" to the angular momentum as we do with the linear momentum, we do have to keep in mind that all of the quantities in the rotational case must be referenced to the same point. That is, the net torque requires a reference point, and the angular momentum contains a rotational inertia, which also requires a reference point.

Recall that the impulse-momentum theorem is just a repackaging of Newton's second law, and so it is with the rotational case, though there is a twist, as we will see shortly:

$$\overrightarrow{F}_{net} = \frac{d\overrightarrow{p}_{cm}}{dt} = \frac{d\left(m\overrightarrow{v}_{cm}\right)}{dt} = m\overrightarrow{a}_{cm} \quad\Longleftrightarrow\quad \overrightarrow{\tau}_{net} = \frac{d\overrightarrow{L}}{dt} = \frac{d\left(I\overrightarrow{\omega}\right)}{dt} = I\overrightarrow{\alpha} \tag{6.1.3}$$

Link Between Angular and Linear Momentum

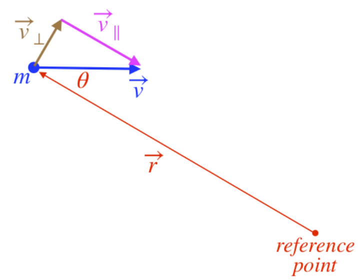
When there are several particles in a system, we find the momentum of the system by adding the momenta of the particles:

$$\overrightarrow{p}_{cm} = \overrightarrow{p_1} + \overrightarrow{p_2} + \dots \tag{6.1.4}$$

We have a definition for the angular momentum of a rigid object, but can we define the angular momentum of a single particle, and then add up all of the angular momenta of the particles to get the angular momentum of the system, in the same way that we do it for linear momentum? The answer is yes, but we have to be careful about our reference point. That is, to add the angular momentum of every particle together to get a total angular momentum, the individual angular momenta must be measured around the same reference.

So how do we define the angular momentum of an individual particle around a certain reference point? Let's look at a picture of the situation. The particle has a mass m, a velocity $\overrightarrow{v}$, and is located at a position $\overrightarrow{r}$ with the tail of that position vector at the reference point.

Figure 6.1.1a – Angular Momentum of a Point Particle

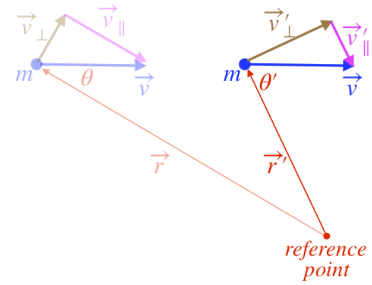


If this particle was a part of a rigid body rotating around the reference point, the parallel component of the velocity vector would be zero. So it makes sense to exclude that part of the velocity vector when defining the angular momentum of this particle. We know the rotational inertia of the point particle, and the relation between $v_\perp$ and ω, so we get for the magnitude of the angular momentum:

$$L_{single\ particle} = I\omega = \left[mr^2\right]\left[\frac{v_\perp}{r}\right] = mrv_\perp = mrv\sin\theta \tag{6.1.5}$$

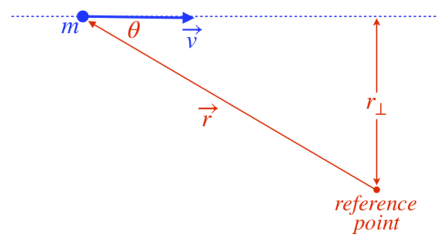
Suppose the particle continues moving free of any forces. What happens to its angular momentum? Let's look at what happens to the picture:

Figure 6.1.1b – Angular Momentum of a Point Particle



What a mess! The mass and velocity vector remain the same, but everything else changes. How can we determine what happens to the angular momentum? Well, have a look at Equation 6.1.2. With no force on the particle, there can't be any torque on the system, so the angular momentum must remain unchanged. It turns out there is a simpler way to look at the angular momentum, to see why this must be the case.

Figure 6.1.1c – Angular Momentum of a Point Particle



We can define the quantity r_{bot} in a manner similar to how we defined moment arm – it is the perpendicular distance from the reference point to the line along which the particle is moving. Doing this gives us an alternative way of writing the magnitude of the particle's angular momentum. Using the fact that $r_\perp = r\sin\theta$, we have:

$$L_{single\ particle} = mrv_\perp = mrv\sin\theta = mvr_\perp \tag{6.1.6}$$

Now it is quite easy to see that the angular momentum of the particle doesn't change while it moves – it keeps the same mass and speed, and stays on the same line, so $r_\perp$ doesn't change either.

Angular momentum is a vector, so what direction does it have here? Going back to the idea of this particle being part of a rigid object, it's clear that this object would be rotating clockwise around the reference, so from the right hand rule, the vector must point into the page. We would like a mathematical expression of this, and as with the case of torque, it comes from the cross product. The two vectors involved are the position vector and the velocity vector, and indeed we see that the following cross product results in the correct direction, and takes care of the $\sin\theta$ contribution as well:

$$\overrightarrow{L}_{single\ particle} = m\overrightarrow{r}\times\overrightarrow{v} = \overrightarrow{r}\times\overrightarrow{p} \tag{6.1.7}$$

This is a nice, compact expression of the relation between the linear momentum of a particle and its angular momentum around a reference point. To see this relation come full circle, imagine that a force is exerted on the particle. This would cause the momentum to change. It would also result in a torque on the system about the reference point, causing the angular momentum to change. Taking the derivative of Equation 6.1.7 with respect to time gives:

$$\overrightarrow{\tau} = \frac{d\overrightarrow{L}}{dt} = \frac{d}{dt}\left(\overrightarrow{r}\times\overrightarrow{p}\right) = \frac{d\overrightarrow{r}}{dt}\times\overrightarrow{p} + \overrightarrow{r}\times\frac{d\overrightarrow{p}}{dt} = \overrightarrow{v}\times\overrightarrow{p} + \overrightarrow{r}\times\overrightarrow{F} \tag{6.1.8}$$

The velocity vector is parallel to the momentum vector, so the cross product in the first term is zero, leaving us with a relation between torque and force that we have seen before (Equation 5.5.6).

Now that we can deal with the angular momentum of a single particle relative to some reference point, we can simply add the contributions of many such particles within a system, relative to the same reference point:

$$\overrightarrow{L}_{system} = \overrightarrow{r}_1\times\overrightarrow{p}_1 + \overrightarrow{r}_2\times\overrightarrow{p}_2 + \ldots \tag{6.1.9}$$

Note that these particles may be part of a rigid object, or may not be bound to each other at all. If they happen to be bound into a single rigid object rotating around a fixed point on the object, then the result is more easily expressed in terms of the rigid object's rotational inertia and angular velocity (Equation 6.1.1):

$$\begin{aligned}\vec{L}_{rigid\ object} &= \vec{r}_1 \times m_1 \vec{v}_1 + \vec{r}_2 \times m_2 \vec{v}_2 + \dots \\ &= [r_1\hat{r}_1] \times m_1 [v_1\hat{v}_1] + [r_2\hat{r}_2] \times m_2 [v_2\hat{v}_2] + \dots \\ &= [r_1\hat{r}_1] \times m_1 [r_1\omega\hat{v}_1] + [r_2\hat{r}_2] \times m_2 [r_2\omega\hat{v}_2] + \dots \\ &= m_1 r_1^2 [\omega\hat{r}_1 \times \hat{v}_1] + m_2 r_2^2 [\omega\hat{r}_2 \times \hat{v}_2] + \dots \\ &= m_1 r_1^2 [\omega\widehat{\omega}] + m_2 r_2^2 [\omega\widehat{\omega}] + \dots \\ &= I\vec{\omega}\end{aligned} \tag{6.1.10}$$

Consider next an extended object that is not rotating, but is moving in a straight line relative to some reference point. Despite the fact that it is not rotating, it can have angular momentum relative to that reference point. Writing the angular momentum of the whole object as a sum of the angular momenta of its particles, we get:

$$\vec{L}_{not-rotating\ extended\ object} = m_1 \vec{r}_1 \times \vec{v}_1 + m_2 \vec{r}_2 \times \vec{v}_2 + \dots \tag{6.1.11}$$

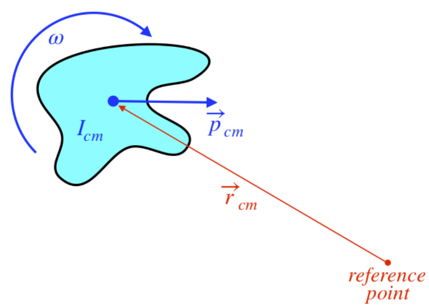
With the object not rotating and all the particles held rigidly in place, every particle has the same velocity, which equals the velocity of the object's center of mass, so this can be factored out of all the cross products, giving:

$$\vec{L}_{not-rotating\ extended\ object} = \left(m_1 \vec{r}_1 + m_2 \vec{r}_2 + \dots\right) \times \vec{v}_{cm} = \left(\frac{m_1 \vec{r}_1 + m_2 \vec{r}_2 + \dots}{M}\right) \times \left(M\vec{v}_{cm}\right) = \vec{r}_{cm} \times \vec{p}_{cm} \tag{6.1.12}$$

What this means is that an extended object moving in a straight line has the same angular momentum relative to a reference point as a point particle located at the object's center of mass, with the same mass and velocity.

If the extended object has both its center of mass moving at a constant velocity relative to the reference point and it is also rotating around an axis through its own center of mass, then things get complicated. We won't into the details of the most general case, but it is not unreasonable to consider the case of the linear velocity lying in the plane perpendicular to the rotation vector (e.g. an object moving within this screen while rotating around an axis perpendicular to this screen – see Figure 6.1.2).

Figure 6.1.2 – Total Angular Momentum



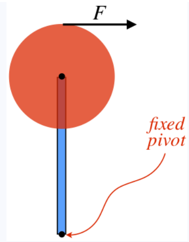
The total angular momentum comes out to be reminiscent of the parallel-axis theorem and of the kinetic energy being the sum of linear and rotational parts:

$$\vec{L}_{tot} = \vec{L}_{rotation\ around\ cm} + \vec{L}_{cm\ moving\ by\ reference\ point} = I_{cm}\vec{\omega} + \vec{r}_{cm} \times \vec{p}_{cm} \tag{6.1.13}$$

An interesting and important consequence of this is that an object that is only rotating around its center of mass (but not moving linearly) has the same angular momentum measured relative to *every* reference point.

Example 6.1.1

The center of a uniform solid disk is attached to a thin uniform rod of equal mass at its end. The radius of the disk is one-third the length of the rod. The rod is held fixed and is free to pivot about at its other end. A constant magnitude of force is exerted tangent to the edge of the disk at the point farthest from the pivot for a short time. There is no gravity present.



The system starts from rest, and after the short time, it has some angular momentum relative to the pivot, with the disk and rod each contributing a portion of it. Find the ratio of the angular momenta contributed by each part if...

a. *... the disk is held rigidly to the rod (i.e. it is not free to rotate around its center).*
b. *... the disk can rotate freely about its center.*

Solution

a. Even though the disk is held rigidly at it center by the rod, it still rotates relative to its center (if you don't see this, imagine painting an arrow on the side of the disk, and focus only on the disk as the whole object spins – clearly the orientation of the disk rotates). In fact, the disk rotates around its center at precisely the same rate as the whole object rotates (the aforementioned painted arrow takes the same amount of time to come all the way around as the object itself). We can compute the angular momentum contribution of the disk and rod in terms of this common angular speed. For the rod it is easy:

$$L_{rod} = I_{rod}\omega = \frac{1}{3}ml^2\omega$$

The disk requires a bit more effort, since its angular momentum consists of a part around its center of mass and a part from the linear motion going around the pivot:

$$L_{disk} = I_{disk}\omega = \frac{1}{2}mR^2\omega + mr_\perp v = \frac{1}{2}m\left(\frac{1}{3}l\right)^2 + ml\left(l\omega\right)\omega = \frac{19}{18}ml^2\omega$$

The ratio of the angular momenta is:

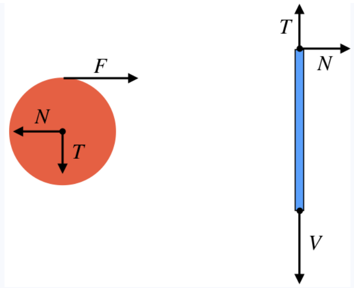
$$\frac{L_{rod}}{L_{disk}} = \boxed{\frac{6}{19}}$$

b. This is significantly more difficult than part (a), because the angular velocity of the disk is not equal to that of the rod. Hopefully it is clear that the rod must rotate. That is, just because the disk is free to rotate, it doesn't mean that the force will not affect the rod's motion. To see this, consider a FBD of the whole system. If F is the only force, then the disk's center of mass must be accelerating, which it is not doing if it is only spinning in place. If there is a balancing force to keep the disk's center of mass in place, then it can only come from the point of contact with the rod. But then the equal-and-opposite force on the end of the rod by the disk will make the rod rotate.

Okay, so we need the separate angular momenta (about the pivot) of the disk and rod to determine the ratio. The angular momentum of the disk is once again the sum of its angular momentum about its center and that which comes from its linear motion around the pivot (which is directly related to the rotational speed of the rod). Calling the rotational speed of the rod ω_1, and the rotational speed of the disk around its center ω_2, we have the following angular momenta for the two pieces:

$$\begin{aligned} L_{rod} &= \tfrac{1}{3}ml^2\omega_1 \\ L_{disk} &= \tfrac{1}{2}mR^2\omega_2 + ml^2\omega_1 = ml^2\left(\tfrac{1}{18}\omega_2 + \omega_1\right) \end{aligned}$$

Note that so far, this same analysis applies to part (a), with the difference being that in part (a) the disk is turning at the same rate as the rod: $\omega_2 = \omega_1 = \omega$. What we need for this case is a way to relate ω_1 and ω_2, which are no longer equal. Both come after a torque is applied for an equal time, but the net torque applied to the rod about the pivot (which determines ω_1) is different from the net torque applied to the disk about its center (which determines ω_2). To get a handle on this, a pair of FBDs are helpful:



A quick explanation of these FBDs: N is the normal force by the axle on the disk, reacting to the applied force, as discussed in the first paragraph above. T is the "tension" force keeping the disk moving in a circle. V is the vertical force by the pivot that makes sure there is a net force which keeps the center of mass of the rod moving in a circle. Neither T nor V play a role in either torque. We'll say that F and N act for a time Δt.

The net torque on the rod about the pivot is Nl. Multiply this by the time it acts (and remembering that it starts from rest), we have, from the impulse-momentum theorem:

$$Nl\Delta t = I_{rod\ about\ pivot}\ \omega_1 = \frac{1}{3}ml^2\omega_1 \quad \Rightarrow \quad N\Delta t = \frac{1}{3}ml\omega_1$$

The net torque on the disk about its center is FR, so:

$$FR\Delta t = I_{disk\ about\ center}\ \ \omega_2 = \frac{1}{2}mR^2\omega_2 \quad \Rightarrow \quad F\Delta t = \frac{1}{6}ml\omega_2$$

[Note: While it might appear as though this rotational impulse determines the rotational motion of the disk relative to the rod, it does not. The resulting motion is the total angular velocity (relative to the lab). If this force was zero, and the rod is made to turn without any torque on the disk (i.e. the force is applied to the rod instead of the disk), the disk would maintain its orientation relative to the lab as the rod rotates, turning relative to the rod in the opposite direction at the same rate that the rod rotates. In this case, ω_2 would be zero, which matches the zero value of F.]

We need one more equation, and it comes from the linear impulse-momentum theorem for the disk. From the FBD, we see that the net force on the disk is $F-N$, and this results in a change of (tangential) momentum of mv_{cm}. The final linear velocity of the disk's center of mass is directly related to ω_1 (it moves with the end of the rod) so:

$$(F-N)\,\Delta t = mv_{cm} = ml\omega_1$$

Putting these last three equations together gives a relationship between ω_1 and ω_2:

$$\frac{1}{6}ml\omega_2 - \frac{1}{3}ml\omega_1 = ml\omega_1 \quad \Rightarrow \quad \omega_2 = 8\omega_1$$

Finally, we put this into the angular momentum for the disk that we found earlier, and compute the ratio:

$$L_{disk} = ml^2\left(\frac{1}{18}[8\omega_1]+\omega_1\right) = \frac{13}{9}ml^2\omega_1 \quad \Rightarrow \quad \frac{L_{rod}}{L_{disk}} = \boxed{\frac{3}{13}}$$

It makes sense that more of the angular momentum goes into the disk in the case where it spins faster than the rod. If you really want to go down the rabbit hole, you can explore a comparison of the kinetic energy given to the system in the two cases. As we have seen before, just because the system has the same angular momentum in both of these cases (same torque on the system for the same period of time), it doesn't mean that they have the same kinetic energy when there are two components of the system that can distribute the momentum differently between them.

6.2: Effects of Torque

Gyroscopic Precession

Back in Section 1.6 and Section 1.7, we discussed circular motion at constant speed as motion that occurs because a net force pulling an object toward a central point causes the object's velocity vector to only change direction, and not magnitude. At the time, we hadn't yet discussed momentum, but clearly we can now replace "velocity vector" in the previous sentence with "momentum vector." We can write Newton's second law (Equation 4.1.4) in terms of the changing magnitude and direction of the momentum:

$$\overrightarrow{F}_{net}=\frac{d}{dt}\overrightarrow{p}=\frac{d}{dt}(p\hat{p})=\frac{dp}{dt}\hat{p}+p\frac{d\hat{p}}{dt} \tag{6.2.1}$$

Circular motion at a constant speed would exhibit no change in the magnitude of momentum – the first term in Equation 6.2.1 is zero – while all of the force would go into changing the direction of momentum. As we saw back in Section 1.6, the two terms in Equation 6.2.1 are always perpendicular to each other, which means that the net force on an object going in a circle at a constant speed is always perpendicular to the momentum vector.

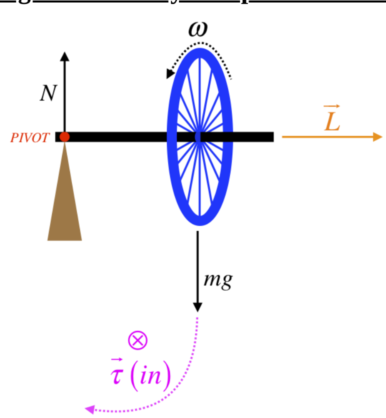
None of this is new to us, but as we have been doing for the last two chapters, we will now look at the rotational equivalent of this behavior. Switching Equation 6.2.1 to the rotational equivalent gives:

$$\overrightarrow{\tau}_{net}=\frac{d}{dt}\overrightarrow{L}=\frac{d}{dt}\left(L\widehat{L}\right)=\frac{dL}{dt}\widehat{L}+L\frac{d\widehat{L}}{dt} \tag{6.2.2}$$

We are already aware of how a net torque can change the magnitude of an object's angular momentum – speeding up and slowing down rotation is something we have already looked at in detail. But what if we insist that the magnitude remain constant (the object maintains the same rotational inertia and keeps spinning at a constant rate), while only the the direction of motion changes? That is, what if the first term in Equation 6.2.2 is zero, while the second term is not? How can we construct a physical system that behaves this way? Answering this last question will require quite a lot of facility with the right hand rule, but here goes...

We start with a rotating object. We'll use as our model a bike wheel turning around an axle. The angular momentum vector will point along the axis of the wheel according to the right hand rule. Now we need a net torque that points perpendicular to the angular momentum. We can achieve this by placing on end of the wheel's axle on a support and allowing the weight of the wheel to pull it down as the support pushes up.

Figure 6.2.1 – Gyroscopic Precession

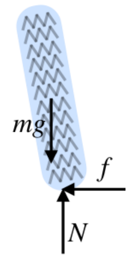


There is clearly a net torque around the pivot at the point of support, trying to turn the wheel clockwise in Figure 6.2.1. But the *direction of the torque vector*, according to the right hand rule, is into the page. This torque vector points in a direction that is perpendicular to the angular momentum vector, exactly as we required above.

The next question is, how does this system behave? In the case of the object going in a circle at a constant speed, the momentum vector of the object never changed length, it only rotated its direction. The way that it rotated was the change of the linear velocity vector pointed in the direction of the force (see Figure 1.7.1). To follow this same behavior, the point of the angular momentum vector must turn in the direction of the torque vector, which is into the page. That is, to behave Newton's second law for rotations, this wheel should not fall down, but instead should *rotate into the page*! In fact it does exactly this, rotating around the pivot in a counterclockwise direction as viewed from above. This phenomenon is known as *gyroscopic precession*.

While this is certainly a striking phenomenon to witness first-hand, it is quite ubiquitous in everyday life. Everyone who rides a bicycle knows that getting it to turn is a matter of *leaning*, not turning the handle bars. To see how this is an example of gyroscopic precession, consider the following force diagram of a tilted wheel:

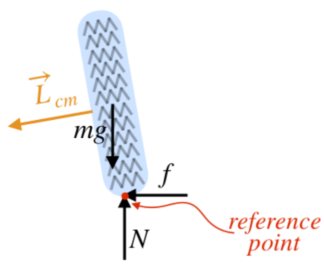
Figure 6.2.2 – Free-Body Diagram of a Tilted Wheel (Forces, End View)



The fact that there is only one force in the horizontal direction (the friction force) means that this wheel's center of mass must be accelerating to the left. If this wheel is stationary, then this is certainly happening as the wheel falls over to the left. On the other hand, if the wheel is rolling forward (into the page), and this tilt is a result of turning, then it is not falling over, but its center is mass is still accelerating to the left (centripetally, toward the center of the turn).

This explains how the tilt results in a change of the wheel's direction of motion, but not its change of *orientation*. That is, why does the wheel *turn* as its center of mass changes direction? For an explanation of this, we can look at the torques and the angular momentum vector. Again assuming that the wheel in the diagram is rolling into the page, the angular momentum vector measured around the center of mass points to the left (and slightly downward).

Figure 6.2.3 – Free-Body Diagram of a Tilted Wheel (Torques, Rear View)

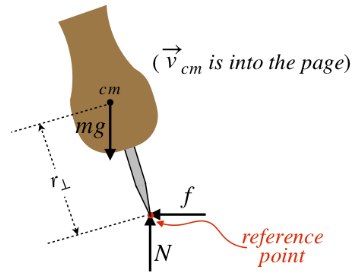


The direction of the net torque around the center of mass is difficult to determine directly. Gravity provides no torque, and the friction and normal forces give opposing torques. But if we choose a reference point where the wheel contacts the road, it is clear that about that point the torque is counterclockwise (only gravity contributes). If the forces act to torque the wheel counterclockwise around that reference point, clearly the net torque around the center of mass will be counterclockwise as well. [*Note: It won't be as large of a torque around the center of mass as there is around the road contact point. The weight and normal forces are equal (the wheel is not accelerating vertically), and the moment arm for gravity around the road contact is the same as the moment arm for the normal force around the center of mass, so those provide equal counterclockwise torques. But for the center of mass reference point there is an additional torque applied by the friction force, and it is clockwise, reducing the net counterclockwise torque.*]

A counterclockwise net torque gives (through the right hand rule) a torque vector that points out of the page. This causes a change in the angular momentum vector out of the page, according to the second law. For $\overrightarrow{L}_{cm}$ to gain a component pointing out of the page, the wheel has to turn left.

It turns out, however, that the change of angular momentum of the spinning wheel about its center is not the only way that this phenomenon is manifest here. In fact, it's not even the greatest contributor to turning-by-leaning. Consider ice skaters. They too change direction by leaning, but they don't have rotating wheels to perform this gyroscopic effect for them. In this case, the diagram is the same (to avoid sketching an entire ice skater, we'll draw just the ice skate):

Figure 6.2.4 – Free-Body Diagram of a Tilted Ice Skate (Rear View)



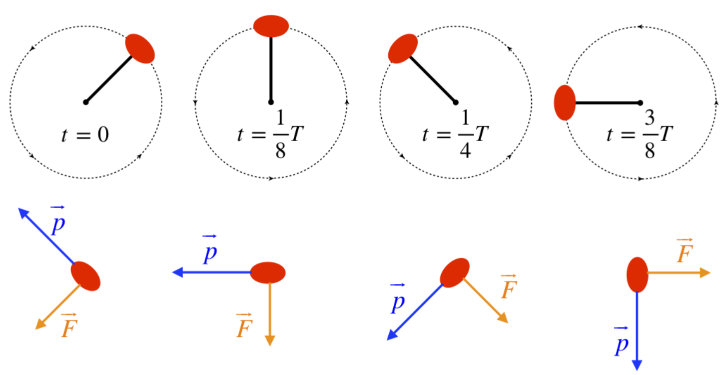
While there is no rotating wheel here, there *is* angular momentum relative to the reference point. The ice skate is moving into the page, and as we know, objects moving in straight lines do have angular momentum relative to reference points that don't lie along the line of linear motion. The magnitude of this angular momentum is $mv_{cm}r_\perp$ about the reference point indicated, but we are more interested in its direction. The position vector points from the reference point to the center of mass of the skate, and the momentum vector points into the page, so $\overrightarrow{L} = \overrightarrow{r} \times \overrightarrow{p}$ points to the left and slightly down, *just like it was for the bike wheel's rotation*. The result is the same – the torque about the reference point causes this angular momentum vector to turn in a direction that is out of the page, which means the skate turns left.

Note that the bicycle has this same thing going on – the center of mass of the bike + rider has an angular momentum relative to the point where the wheel contacts the ground, and leaning to one side or the other will turn the entire system, not just the wheel.

Let's return to the original example of the precession of a wheel pivoted about an end of its axle. The torque remains constant in magnitude, which means that the angular momentum vector changes at a constant rate – the wheel therefore precesses around the pivot at a constant rate. Let's see if we can determine the rate of this precession, i.e. its rotational velocity around the *vertical* axis (not around the axle of the wheel - we already know that). Rather than try to slog through all the vector calculus, let's do this by following our circular motion analogy.

Figure 6.2.5 depicts a stone tied to a string going in a counterclockwise circle at a constant speed. The four diagrams are snapshots of the motion at four different times, separated by one-eight of a rotational period. Below each diagram is a depiction of the net force vector on the stone and its linear momentum vectors at that moment in time (note these vectors are always perpendicular, as they should be).

<u>Figure 6.2.5 – Force and Momentum Vectors for Stone in Circular Motion</u>

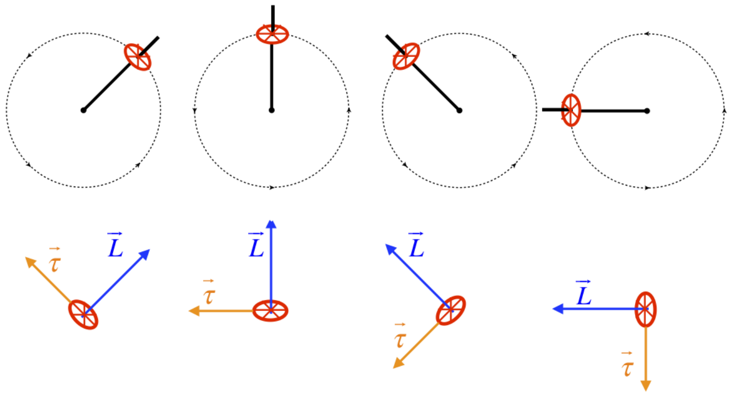


Let's compute the angular rate (which we will call Ω to avoid later confusion with the ω for the wheel) of the rock going around the circle *in terms of the magnitudes of* $\overrightarrow{F}$ and $\overrightarrow{p}$. The magnitude of the force is the mass of the rock multiplied by the centripetal acceleration:

$$F = ma_c = m\frac{v^2}{R} = mv\left(\frac{v}{R}\right) = p\Omega \quad \Rightarrow \quad \Omega = \frac{F}{p} \tag{6.2.3}$$

That's quite compact! Okay, let's follow precisely this path for the precessing wheel. We start with a similar diagram of a top view (gravity is acting into the page).

<u>Figure 6.2.6 – Torque and Angular Momentum Vectors for Precessing Wheel</u>



The main difference that jumps out between Figure 6.2.5 and Figure 6.2.6 is that for the stone, the momentum is tangent to the circle while the force is radial, while for the wheel, the angular momentum is radial and the torque is tangent. Other than that, however, these follow each other exactly – the force/torque is perpendicular to and "leads" the momentum, as they both revolve together. Given that they follow the same behavior and have exactly the same differential relationship (the force/torque is the time derivative of the momentum), it's perfectly reasonable to expect that the rotational frequency would have the same relationship. Namely:

$$\Omega = \frac{\tau}{L} \tag{6.2.4}$$

If you are not comfortable with this "derivation by analogy," then you can reach the same result quickly using Equation 6.2.2. The rate at which the wheel is spinning about its axle doesn't change, which means that the magnitude of its angular momentum remains constant, and the first term is zero. Solving the remainder of the equation for the rate at which the direction is changing requires only a division of both sides by the magnitude of the angular momentum, giving the same result as above.

We can write this in terms of the rotational inertia of the wheel about its axis I, the rotational speed of the wheel ω, the length of the axle l (the distance from the pivot to the center of mass of the wheel + axle), and the mass of the wheel M. The torque can be quickly calculated with a quick look at Figure 6.2.1, and plugging in for the angular momentum of the wheel, we get the precession speed:

$$\Omega = \frac{Mgl}{I\omega} \tag{6.2.5}$$

It should be noted that when the wheel precesses, the system is now also rotating it a horizontal plane, which gives it a component of angular momentum vector in the upward direction, and this angular momentum is also affected by the torque. This effect becomes progressively easier to ignore for a given setup as the wheel spins faster about its axis, because according to Equation 6.2.5, the faster the wheel spins about its axis, the slower it precesses, and the slower it precesses, the smaller the upward component of angular momentum. In any case, this secondary effect manifests itself as a bobbing up-and-down of the wheel as it precesses, and is known as *nutation*.

Example 6.2.2

A wheel whose axis is vertical (i.e. the plane of the wheel is parallel to the ground) rotates clockwise as viewed by someone looking down at it. If a small nudge is given to the top of the axis of this wheel toward the south, which way does the top of the axis move in its immediate response to this nudge?

Solution

Using the RHR, the clockwise rotation when looking down on it corresponds to an angular momentum vector that points downward. Imagine you are facing the spinning wheel from the north side (i.e. you are facing south), and push it at its top with the fingers of your right hand pointing upward. Your right hand will curl such that your thumb is pointing to the left, and that is the direction of the torque you are imparting on the wheel. The angular momentum vector that points down will change in the direction of the torque, so the axis of the wheel will tilt such that its bottom rotates in the direction of your thumb, which means the top of the axis will rotate the other way (to your right). If you are facing south, then a tilt to the right is toward the west.

Central Forces

Consider a flat disk rotating around its center. Every particle in this object is following a circular path, and so every particle is experiencing a net force. We might therefore ask how angular momentum can be conserved – with net forces on every particle, the forces on them are not cancelling-out. The answer is that net force is not the same thing as net torque. Measured from the fixed reference point, the force vector on every particle in the object points parallel to the position vector, which means the torque (the cross product of position and force) on every particle is zero. We can in fact elevate this idea to a very general rule, which first requires a definition:

Definition: Central Force

A central force is a force (which can act on many objects) that is directed directly toward or directly away from a single fixed point in space.

And from the preceding discussion, we saw that whenever the force vector is parallel to the position vector, if we choose the "source" of the force to be our reference point, we conclude:

Central forces do not exert torques (relative to the central point) on the objects they influence, and therefore angular momentum around the central reference point is conserved.

Example 6.2.3

Consider this position-dependent force:

$$\overrightarrow{F}(x,y) = \lambda\left[x\;\hat{i} + y\;\hat{j}\right]$$

A rock is tied to a $0.50m$ string whose other end is held fixed at the origin, and is then set into circular motion in the $x-y$ plane (there is no gravity present). While the rock is moving at a constant speed in a circle, the force described above is turned on. At one moment the rock (still under the influence of the string and the force above) is moving at a speed of $4.0\frac{m}{s}$, when suddenly the string breaks. A little while later, the rock crosses the $+x$-axis a distance $2.5m$ from the origin at a speed of $5.8\frac{m}{s}$. Find the angle the velocity vector of the rock makes with the x-axis at this time.

Solution

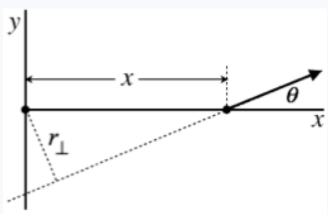
This force can be rewritten in terms of the position vector relative to the origin as:

$$\overrightarrow{F} = \lambda\overrightarrow{r}$$

Clearly this force points directly away from a single point (the origin), which makes it a central force. This means that it exerts no torque around the origin, which also means that it cannot change the angular momentum relative to this point. When the force is turned on, it therefore has no effect on the speed of the rock while the rock remains attached to the string. When the string breaks, the rock will neither follow a straight line, nor will it maintain a constant speed, but with only the central force acting on it, the angular momentum remains constant. We can easily compute the angular momentum it started with. It is:

$$L_o = mv_or_\perp = mv_oR$$

When the rock reaches the x-axis, its position on the axis and the angle its velocity vector makes with the axis are related to the perpendicular distance $r_\perp$, as can be seen in this diagram:



$$\sin\theta = \frac{r_\perp}{x}$$

The angular momentum at the x-axis is $mv_fr_\perp$, so setting this equal to the starting angular momentum, we can solve for θ:

$$L_o = L_f \quad \Rightarrow \quad mv_oR = mv_fr_\perp \quad \Rightarrow \quad \theta = \sin^{-1}\left[\frac{\cancel{m}v_oR}{\cancel{m}v_fx}\right] = \boxed{7.9^o}$$

6.3: Applications of Angular Momentum Conservation

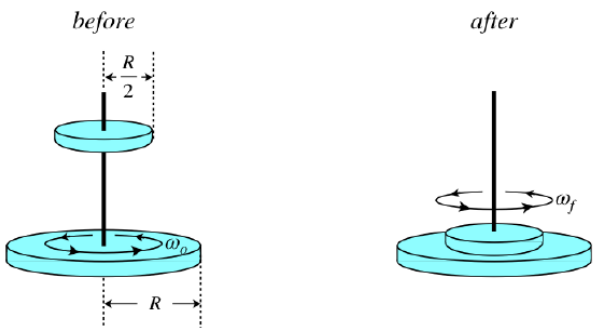
If we continue to follow the trail we blazed in linear motion, our next step is to consider what happens when we choose a system for which there are no external rotational impulses. For such a system, we can declare the angular momentum to be conserved before and after any event, however complicated the internal interactions might be.

In the linear case, we saw that the primary application of momentum conservation was related to collisions, because it was useful to ignore the complicated forces that come about between the colliding objects. What sorts of problems might angular momentum conservation be useful for solving? There are actually three basic varieties that commonly arise in classical mechanics, and we will look each one in turn.

Spinning Collisions

Two uniform solid disks with small holes in their centers, are threaded onto the same frictionless vertical cylindrical rod. One of the disks lies flat on a frictionless horizontal surface and is rotating at a speed ω_o around the rod, while the other disk is held at rest directly above it. Both disks are made from the same material, and have the same thickness, but the spinning disk has twice the radius of the stationary disk. The smaller disk is then dropped on top of the larger one, and after a short time the kinetic friction force between the two disks brings them both to the same rotational speed, which is a fraction of the larger disk's original speed. Find this fraction, and the fraction of the original kinetic energy still left the system afterward (it loses some from work done by kinetic friction).

Figure 6.3.1 – Rotating Disk Inelastic Collision



This is clearly the rotational version of a perfectly inelastic collision, as both of the objects end up moving together. We solve it in the same way that we solve the linear counterpart – by noting that the only torques involved are internal to the two-disk system, which means that the total angular momentum is the same before and after the collision.

$$\left.\begin{array}{ll} before: & L_{tot}=I_1\omega_o+I_2\left(0\right) \\ after: & L_{tot}=\left(I_1+I_2\right)\omega_f \end{array}\right\} \;\;\Rightarrow\;\; \omega_f=\frac{I_1}{I_1+I_2}\omega_o \tag{6.3.1}$$

Because it has the same thickness and is made from the same material, the ratio of the two disks' masses equals the ratio of their areas. With one-half the radius, the smaller disk therefore has one-fourth the mass, and we get:

$$I_1=\frac{1}{2}MR^2 \;\;\Rightarrow\;\; I_2=\frac{1}{2}\left(\frac{1}{4}M\right)\left(\frac{1}{2}R\right)^2=\frac{1}{16}I_1 \;\;\Rightarrow\;\; \omega_f=\frac{I_1}{I_1+\frac{1}{16}I_1}\omega_o=\boxed{\frac{16}{17}}\omega_o \tag{6.3.2}$$

Now for the fraction of kinetic energy leftover:

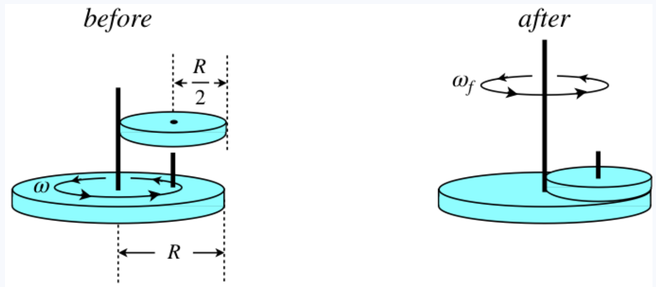
$$\left.\begin{array}{l} KE_o=\frac{1}{2}I_1\omega_o^2 \\ KE_f=\frac{1}{2}\left(I_1+I_2\right)\omega_f^2 \end{array}\right\} \;\;\Rightarrow\;\; \frac{KE_f}{KE_o}=\frac{\frac{1}{2}\left(I_1+I_2\right)\omega_f^2}{\frac{1}{2}I_1\omega_o^2}=\frac{\left(\frac{17}{16}I_1\right)\left(\frac{16}{17}\omega_o\right)^2}{I_1\omega_o^2}=\boxed{\frac{16}{17}} \tag{6.3.3}$$

We can actually achieve this last answer even more easily using $L_o=L_f$:

$$\frac{KE_f}{KE_o}=\frac{\frac{1}{2}I_f\omega_f^2}{\frac{1}{2}I_o\omega_o^2}=\frac{I_o}{I_f}\frac{\left(I_f\omega_f\right)^2}{\left(I_o\omega_o\right)^2}=\frac{I_o}{I_f}\frac{\cancel{L_f^2}}{\cancel{L_o^2}}=\frac{I_1}{I_1+I_2} \tag{6.3.4}$$

Example 6.3.1

Let's alter the sample problem above slightly. We'll use the same two disks as before, but rather than have them collide with their axes aligned, we'll drop the disk onto a post that is sticking out of the larger disk a distance of one-half the radius from the center. As before, the kinetic friction force will cause the disks to stop sliding across each other, and all the torques (from the friction, and from the force on the post when the small disk first lands on it) are internal to the system. Find the fractions of final rotational speed to the initial rotational speed and the fraction of the initial kinetic energy that is left at the end.



Solution

The angular momentum is again conserved, and the only thing that is different this time is the final rotational inertia. This time the rotational inertia afterward is greater (according to the parallel-axis theorem) by an amount md^2. Putting in $m=\frac{1}{4}M$ and $d=\frac{1}{2}R$, we get an added contribution to the final rotational inertia of $\frac{1}{16}MR^2$, which equals $\frac{1}{8}I_1$. So we get the same result as Equation 6.3.2, except that the denominator is larger by $\frac{1}{8}I_1$:

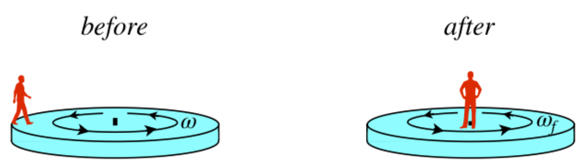
$$\omega_f=\frac{I_1}{I_1+\frac{3}{16}I_1}\omega_o=\boxed{\frac{16}{19}}\omega_o$$

As with the above example, the ratio of the kinetic energies comes out to the same as the ratio of the rotational speeds.

Changing Rotational Inertia

A child stands on the outer edge of a merry-go-round, which is spinning around a fixed axle on a horizontal frictionless surface. The merry-go-round is a solid, uniform disk with ten times the mass of the child, and is spinning at a rotational speed ω_o. The child then slowly walks to the center of the merry-go-round. What, if anything, happens to the rotational speed of the merry-go-round?

Figure 6.3.2 – System Changes Rotational Inertia While Rotating



Before we invoke angular momentum conservation and launch into the mathematics, it might help to think about this in a "less evolved" manner – let's think about the *internal* interactions in the child + merry-go-round system. When the child takes a step, toward the center, they are moving from a faster moving part of the merry-go-round to a slower part. This means that the merry-go-round will exert a static friction force on the feet of the child tangent to the circular motion, acting to slow them down. There is, of course, a Newton's third law pair friction force on the merry-go-round by the feet of the child in the opposite direction, which results in a torque that acts to speed up the merry-go-round's rotation. So we would expect the linear speed of the child to slow with every step, as the merry-go-round's rotational speed increases. The details of these changes are hard to work out using the details of the interaction, so now we turn to momentum conservation, which we know holds because the only forces/torques acting are internal to the system.

Calling the mass of the child m and the radius of the merry-go-round R, we can write down the angular momentum referenced at the axis of the merry-go-round before and after, and invoking angular momentum conservation makes the rest easy:

$$\left.\begin{array}{ll} before: & L_{tot}=\left[I_{child}+I_{mgr}\right]\omega_o=\left[mR^2+\frac{1}{2}(10m)R^2\right]\omega_o=6mR^2\omega_o \\ after: & L_{tot}=\left[I_{child}+I_{mgr}\right]\omega_o=\left[\;0\;+\frac{1}{2}(10m)R^2\right]\omega_o=5mR^2\omega_f \end{array}\right\}\;\Rightarrow\;\omega_f=\frac{6}{5}\omega_o \tag{6.3.5}$$

The rotation rate of the merry-go-round increases by 20%. It's interesting to note that there is no physical equivalent of this phenomenon in linear mechanics. That is, we don't see closed systems losing linear inertia (mass) and maintaining their momentum by compensating with a larger linear velocity.

It's also interesting to consider what happens to the kinetic energy of the system during this process. Like kinetic energy for linear motion, we can write it in terms of the momentum and inertia:

$$KE = \frac{p^2}{2m} \quad \Longleftrightarrow \quad KE = \frac{L^2}{2I} \tag{6.3.6}$$

Given that the angular momentum doesn't change, the kinetic energy goes up in the same proportion that the rotational inertia goes down. Where does this increase in kinetic energy come from? Where is work done? When the child just stands at the edge of the merry-go-round, the static friction force acts toward the center of rotation, but it does no work, because it is acting perpendicular to the direction of the child's motion. But as the child starts moving inward, this static friction *is* doing work. In the end, the kinetic energy of the merry-go-round equals its starting kinetic energy, plus the starting kinetic energy of the child, plus the work done by the static friction force.

It turns out that showing this for this case requires fancier integration to calculate the work than we want to do here, so let's try a simpler example. Let's let the mass of the merry-go-round be negligible compared to the mass of the child. Furthermore, we'll have the child walk halfway to the center of rotation (we can't let the child walk all the way in, or the massless merry-go-round will be spinning infinitely fast!).

First, let's compute the kinetic energy change of the system using angular momentum conservation (note that the merry-go-round doesn't contribute at all now, making things significantly easier):

$$L_{before} = L_{after} \quad \Rightarrow \quad I_o\omega_o = I_f\omega_f \quad \Rightarrow \quad mR^2\omega_o = m\left(\frac{R}{2}\right)^2\omega_f \quad \Rightarrow \quad \omega_f = 4\omega_o \tag{6.3.7}$$

As we saw above, the proportional increase in kinetic energy is the same as that of the rotational velocity, so the kinetic energy increase of the system is:

$$\Delta KE = KE_f - KE_o = 4KE_o - KE_o = \frac{3}{2}mR^2\omega_o^2 \tag{6.3.8}$$

Okay, now let's see if we can calculate the work done by the static friction force. The force that keeps the child going in a circle equals the mass of the child multiplied by the centripetal acceleration, so a force barely exceeding this amount will get the child moving inward. We don't want the child to accelerate appreciably in the radial direction (the child stops at the new radius and it doesn't matter how long it takes to get there), so we can use this as the force that is doing work. The only trouble is, this force changes as the child moves inward, because the rotational speed and distance from the center are changing all the way. We can determine how the rotational speed varies with the child's distance from the center using angular momentum conservation, which allows us to write the force as a function of r as follows:

$$\left.\begin{array}{lll} F = ma_c & \Rightarrow & F(r) = m[\omega(r)]^2 r \\ mr^2\omega(r) = mR^2\omega_o & \Rightarrow & \omega(r) = \dfrac{R^2}{r^2}\omega_o \end{array}\right\} \quad \Rightarrow \quad F(r) = \frac{mR^4\omega_o^2}{r^3} \tag{6.3.9}$$

Now we have only to do the work integral. The displacement is toward the center (r is getting smaller), so $dl = -dr$, and the force is in the direction of displacement, so $\vec{F}\cdot\vec{dl} = F\,dl = -F\,dr$. And the limits of integration are from $r = R$ to $r = \frac{R}{2}$:

$$W\left(R \rightarrow \frac{R}{2}\right) = \int_R^{\frac{R}{2}} -F(r)\,dr = -mR^4\omega_o^2\int_R^{\frac{R}{2}}\frac{dr}{r^3} = -mR^4\omega_o^2\left[-\frac{1}{2r^2}\right]_R^{\frac{R}{2}} = \frac{3}{2}mR^2\omega_o^2 \tag{6.3.10}$$

Comparing this with Equation 6.3.8, we see that the work done in moving the child inward is precisely equal to the change in the system's kinetic energy.

Example 6.3.2

A glass of ice water rests on the outer edge of a solid, uniform, rotating disk, which is spinning horizontally around its frictionless axle. The glass is a cylinder with a mass equal to one half the mass of the disk, and a radius that is one third the radius of the disk. Condensation at the bottom of the glass causes the coefficient of static friction on the top of the disk to go down, and the glass suddenly slides off. Describe what happens to the rotational speed of the disk as a result of the glass sliding off.

Solution

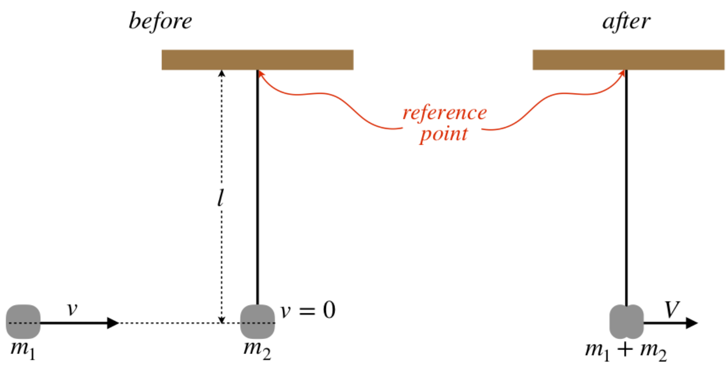
Don't be fooled by all the details given – the rotational speed of the disk doesn't change! When the static friction force goes away, the system continues with its angular momentum, but you can't erase the glass of water from the system just because it slid away. Yes, the rotational inertia of what is going around the axis has changed, but as the glass slides away, it still has angular momentum (it will be a combination of its rotation about its center and movement relative to the axle, see Equation 6.1.13 and Figure 6.1.2). In fact, the glass will continue to have the same angular momentum that it had right before it started sliding, since there is no net torque on it. If the whole system maintains its angular momentum, and the glass keeps the same angular momentum, then the disk must as well – it doesn't change speed at all.

Off-Center Collisions

Of all the problems that are solvable with angular momentum conservation, those that fall into the category of "off-center collisions" are the most interesting and complex. One reason is that unlike the cases of spinning collisions and changing rotational inertia, off-center collision problems often see cameo appearances from linear momentum conservation. Additionally, the fate of the system's mechanical energy becomes more interesting.

We begin with a problem that we are already familiar with from Section 4.6 – the ballistic pendulum. We were able to solve that problem by first solving the perfectly inelastic collision of the bullet & block to get their combined velocity, after which we used mechanical energy conservation to get the height to which the bullet & block swing. We will be more careful about extension in space (and the implications to rotational inertia) by replacing the bullet & block with two small clay balls that stick together. Also, we will not bother to look at the second half of the problem where the pendulum swings up, as the mechanical energy conservation portion of the problem is unchanged.

Figure 6.3.3 – Ballistic Pendulum with Two Small Clay Balls



If we choose the position where the string is attached to the ceiling as a reference point, we note that at the moment of the collision, the gravity and tension forces both act through the reference point, which means that there are no external torques on the system. The bullet and block exert torques on each other, but those are internal and cancel each other. Therefore, as an alternative to using linear momentum conservation, we can use angular momentum conservation.

Before the collision, the pendulum (the length of which we will call l) has no angular momentum relative to the reference point, but the bullet does, according to Equation 6.1.6:

$$L_{before} = m_1 v l \tag{6.3.11}$$

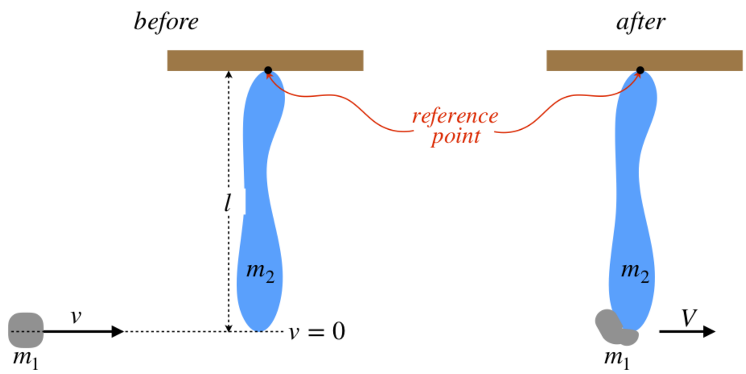
After the collision, the pendulum is rotating, and has a rotational inertia around the reference point, resulting in a final angular momentum of:

$$L_{after} = I\omega = \left[(m_1+m_2)\,l^2\right]\left[\frac{V}{l}\right] = (m_1+m_2)\,Vl \tag{6.3.12}$$

And setting the initial angular momentum equal to the final gives the same result as when we used linear momentum (Equation 4.6.1).

Using angular momentum conservation is no longer optional – it is a requirement – when the target is not just a small ball at the end of a string, but is an extended object with a rotational inertia.

Figure 6.3.4 – Ballistic Pendulum with One Clay Ball



We know that the rotational inertia for this target around the reference point is less than it was when the target was a clay ball, since some of its mass is closer to the reference point. We will write the rotational inertia as some unknown fraction β multiplied by the rotational inertia of a small ball at the end of a string:

$$I_{target} = \beta m_2 l^2 \tag{6.3.13}$$

For example, if this target is a uniform thin rod, then Equation 5.3.7 applies, and $\beta = \frac{1}{3}$, or if the target is a uniform disk or cylinder pivoted about an axis perpendicular to its flat side and about its edge, then Equation 5.3.23 applies, with $R = \frac{1}{2}l$, giving $\beta = \frac{3}{8}$, and so on.

Applying angular momentum as we did above, we can find the final speed of the blob of clay and/or the rotational speed of the pendulum. The initial angular momentum is the same as before, so:

$$L_{after} = I\omega = \left[m_1 l^2 + \beta m_2 l^2\right]\left[\frac{V}{l}\right] = (m_1+\beta m_2)\,Vl \quad \Rightarrow \quad V = \frac{m_1}{m_1+\beta m_2}v \quad \Rightarrow \quad \omega = \frac{V}{l} = \left(\frac{m_1}{m_1+\beta m_2}\right)\frac{v}{l} \tag{6.3.14}$$

The claim was made above that we no longer have the option of using linear momentum conservation for this problem. Before we see *why* this must be true, let's show that it is true for the specific case of a uniform thin rod that has the same mass as the clay ($m_1 = m_2$). If we use linear momentum conservation, then when the clay is stuck on the end of the rod, the center of mass velocity of the rod + clay system is:

$$m_1 v = (m_1+m_2)\,v_{cm} \quad \Rightarrow \quad v_{cm} = \frac{1}{2}v \tag{6.3.15}$$

The center of mass of the rod + clay system is halfway between the center of mass of the rod and the position of the clay, so it is a distance of $\frac{3}{4}l$ from the reference point. With a linear speed of $\frac{1}{2}v$, we get that the pendulum should have a rotational speed of:

$$\omega = \frac{\frac{1}{2}v}{\frac{3}{4}l} = \frac{2}{3}\frac{v}{l} \tag{6.3.16}$$

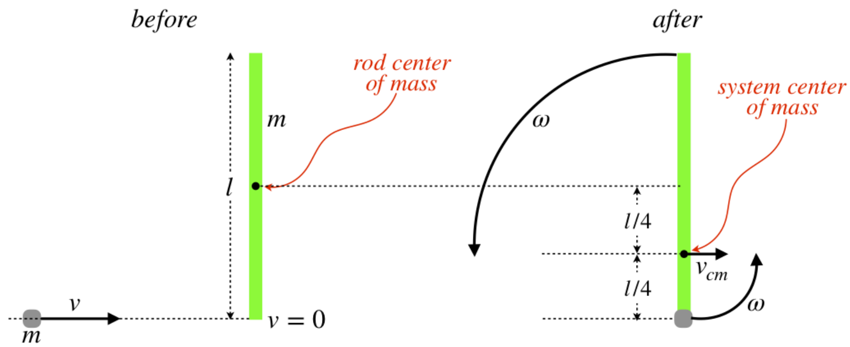
Let's check to see if this is right by plugging $\beta = \frac{1}{3}$ (for a thin rod rotated around its end) and $m_1 = m_2$ into Equation 6.3.14:

$$\omega = \left(\frac{m}{m+\frac{1}{3}m}\right)\frac{v}{l} = \frac{3}{4}\frac{v}{l} \tag{6.3.17}$$

So we see that using linear momentum conservation does not agree with using angular momentum conservation in this case. The reason is the presence of the pivot. The pivot will never exert a torque on the rod relative to the reference point, but it *will* exert a force on it, thereby ruining linear momentum conservation. But this brings up another puzzle: Whatever force the pivot exerts, it causes the speed of the center of mass to be *greater* after the collision than we found for conserved linear momentum, so the force on the rod by the pivot must be *forward*. That doesn't sound right – doesn't the pivot slow down the rod? To solve this puzzle, we get to look at yet another case – an off-center collision with no fixed pivot.

Let's do the same clay-hits-end-of-uniform-thin-rod-with-same-mass problem as above, this time free of any pivot (we'll also assume no gravity is present). First of all, we know that without a force coming from the pivot, the result we obtained in Equation 6.3.15 must be correct, as linear momentum must be conserved. Also we know that after the collision, with no forces on the clay + rod system, it must rotate around its center of mass. This calls for a fresh new diagram:

Figure 6.3.5 – Off-Center Perfectly Inelastic Collision



With the rod rotating counterclockwise, the bottom of the rod must be moving forward faster than the system's center of mass, while the top of the rod might actually be moving *backward*, depending upon the values of ω, l and v_{cm}. If this turns out to be the case, then it makes sense that a pivot could push the rod forward upon impact – the rod's rotation is fast enough compared to its linear motion that the top "tries" to move backward, but is prevented from doing so by a forward push from the pivot. Okay, let's see if this is the case mathematically.

At the moment of the impact, we have:

$$\begin{aligned} v_{bottom} &= v_{cm} + r\omega = \frac{1}{2}v + \frac{1}{4}l\omega \\ v_{top} &= v_{cm} - R\omega = \frac{1}{2}v - \frac{3}{4}l\omega \end{aligned} \tag{6.3.18}$$

Now we need to come up with ω. Even though linear momentum is conserved in this case, Equation 6.3.16 still isn't correct, as it assumes that the top end of the rod is held fixed. We need to use angular momentum conservation. Without a fixed pivot, what do we use as a reference point? The answer is *anywhere* – the angular momentum is conserved relative to every reference point! However, if we are carefree about this choice, we have to be extra careful when adding up the angular momentum after the collision. In the case of a fixed pivot, it was easy because we were able to use the rotational inertia around that point. When we have no fixed point on the object, we have to use Equation 6.1.13. So why not use the center of mass of the clay + rod system at the time of collision as the reference point, and get rid of that pesky second term from Equation 6.1.13?

$$\begin{aligned} L_{before} &= mvr_{\perp} \\ L_{after} &= I_{cm}\omega \end{aligned} \tag{6.3.19}$$

It's clear from the diagram that $r_{\perp}$ is $\frac{1}{4}l$, but we need to do a little bit of work to determine the rotational inertia of the system around its center of mass. This will be the sum of the rotational inertia of the point mass clay and the rotational inertia of the rod about its center (Equation 5.3.19), offset (using the parallel-axis theorem) by $d = \frac{1}{4}l$:

$$I_{cm} = I_{clay} + I_{rod} = \left[m\left(\frac{1}{4}l\right)^2\right] + \left[\frac{1}{12}ml^2 + m\left(\frac{1}{4}l\right)^2\right] = \frac{5}{24}ml^2 \tag{6.3.20}$$

Invoking angular momentum conservation and plugging in for $r_{\perp}$ and I_{cm} gives:

$$mvr_{\perp}=I_{cm}\omega \quad \Rightarrow \quad mv\left(\frac{1}{4}l\right)=\frac{5}{24}ml^2\omega \quad \Rightarrow \quad l\omega=\frac{6}{5}v \tag{6.3.21}$$

Plugging this back into Equation 6.3.18 confirms what we suspected, that without the top end fixed, its initial motion after the collision is *backward*, which is why the force by the pivot must be forward when it is attached:

$$\begin{aligned} v_{bottom} &= \frac{1}{2}v+\frac{1}{4}\left(\frac{6}{5}v\right)=\frac{4}{5}v \\ v_{top} &= \frac{1}{2}v-\frac{3}{4}\left(\frac{6}{5}v\right)=-\frac{2}{5}v \end{aligned} \tag{6.3.22}$$

This has been a long journey through off-center collisions, but we have one more stop – the fate of the system's mechanical energy. We explored perfectly inelastic head-on collisions in Section 4.5, and found a simple relation between the starting and ending kinetic energies of the system – Equation 4.5.7. Given that the final speed of the center of mass of the system has to be the same regardless of whether the collision is head-on or off-center, Equation 4.5.7 clearly cannot work for off-center collisions, as these result in rotations, and as we know, the total kinetic energy is the sum of linear and rotational parts. This means that perfectly inelastic collisions that occur off-center do not lose as much mechanical energy as perfectly inelastic head-on collisions.

This actually makes some intuitive sense. Let's take as an example a bullet digging into a block of wood. The bullet is subject to a non-conservative force that does enough work to slow the bullet to the same speed as the region of the block of wood it is entering (i.e. the bullet stops inside the block). Now let's assume that the force exerted on the bullet by the wood is the same wherever it enters the wood (it is something like "$\mu_k N$," where the normal force is the wood squeezing the bullet). Whether the bullet enters the block at its center of mass or at its edge, the center of mass of the block reaches the same final speed – we'll call the moment when this final speed is reached "t_o." If the bullet hits the center of mass, at t_o the bullet will have slowed to the same speed as the final speed of the center of mass. If the bullet hits the outer edge of the block and makes it spin, then the bullet is not slowed as much at t_o, because the edge of the block is moving faster than the final speed of center of mass of the block. If the bullet isn't slowed as much when it hits the edge, then not as much work is done on it (smaller change in its kinetic energy) by the non-conservative force, and less mechanical energy is converted to thermal.

Let's compute the fraction of kinetic energy that remains for the case above and compare it to the result if the collision occurs at the center of mass.

$$\begin{aligned} collides\ with\ center: \quad \frac{KE_f}{KE_o} &= \frac{m_1}{m_1+m_2} &&= \frac{m}{m+m} &&= \tfrac{1}{2} \\ collides\ with\ end: \quad \frac{KE_f}{KE_o} &= \frac{\frac{1}{2}(2m)\,v_{cm}^2+\frac{1}{2}I_{cm}\omega^2}{\frac{1}{2}mv^2} &&= \frac{m\left(\frac{1}{2}v\right)^2+\frac{1}{2}\left(\frac{5}{24}ml^2\right)\left(\frac{6}{5}\frac{v}{l}\right)^2}{\frac{1}{2}mv^2} &&= \tfrac{4}{5} \end{aligned} \tag{6.3.23}$$

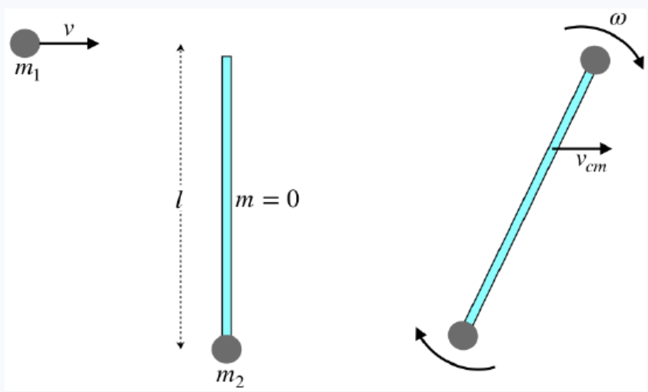
As you can see, less energy is lost when the clay sticks to the end and spins the rod than when it hits the center and doesn't spin it. Here is a nice demonstration of this phenomenon. First the puzzle:

And now the experimental evidence:

Both blocks rise to the same height, because their upward linear velocities start off the same, due to conservation of linear momentum, which is identical for both blocks, independent of where the bullet strikes. Our analysis above resolves the "puzzle" of the difference in mechanical energies of the two systems.

Example 6.3.3

A massless magnetic rod has a small steel ball (which does have mass, but a negligible radius) attached to one end, and is at rest. Another small steel ball approaches the open end of this rod at a right-angle, and when it reaches the end of the rod, sticks to it. The dumbbell-looking combination continues forward, spinning as it goes (see the diagram). Show the surprising result that no kinetic energy is lost in this collision. The diagram provides labeling of quantities that you can use – you cannot make any assumptions about the relative values of m_1 *and* m_2.



Solution

We are showing that kinetic energy is conserved, and the only principles that we can use are linear and angular momentum conservation. Let's start with linear momentum conservation. We have done this a hundred times – the incoming momentum equals the outgoing:

$$m_1 v = (m_1 + m_2) v_{cm} \quad \Rightarrow \quad v_{cm} = \frac{m_1}{m_1 + m_2} v$$

Now for conservation of angular momentum. Let's use the center of mass at the time of collision as the reference point. So we need to determine the perpendicular distance of the incoming ball from the center of mass. This is a straightfoward calculation (e.g. use the incoming ball as the origin), which gives:

$$r_\perp = r_1 = \frac{m_2}{m_1 + m_2} l$$

For later reference, we also have for the distance of the other ball from the center of mass:

$$r_2 = \frac{m_1}{m_1 + m_2} l$$

With this we get the starting angular momentum, and with the rotational inertia of the dumbbell about the center of mass, we get an equation resulting from angular momentum conservation:

$$\left.\begin{array}{l} L_o = m_1 v r_\perp = \dfrac{m_1 m_2}{m_1+m_2} vl \\ L_f = I_{dumbbell}\,\omega = \left(m_1 r_1^2 + m_2 r_2^2\right)\omega = \dfrac{m_1 m_2}{m_1+m_2} l^2 \omega \end{array}\right\} \quad \Rightarrow \quad L_o = L_f \quad \Rightarrow \quad \omega = \frac{v}{l}$$

Now all we have to do is construct the final kinetic energy:

$$KE_f = \frac{1}{2}(m_1+m_2)\,v_{cm}^2 + \frac{1}{2}I\omega^2 = \frac{1}{2}(m_1+m_2)\left[\frac{m_1}{m_1+m_2}v\right]^2 + \frac{1}{2}\left[\frac{m_1 m_2}{m_1+m_2}l^2\right]\left[\frac{v}{l}\right]^2 = \frac{1}{2}m_1 v^2 = KE_o$$

CHAPTER OVERVIEW

7: Gravitation

Thumbnail: A simple swinging pendulum. Image used with permission (Public domain; Lucas V. Barbosa (Kieff)).

7.1: Universal Gravitation

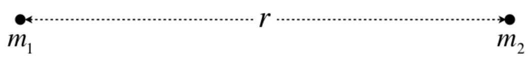
Newton Again

We return to a topic we have discussed only in the simplest of terms, but which has a great deal more depth. Of Newton's many achievements, one of his greatest has to have been the amazing realization that the gravity force is not simply a terrestrial phenomenon. Until he came along, people thought that objects "naturally" fall when they are near the Earth, and that heavenly bodies "naturally" do their little dance. To make the connection that the motions of planets could be explained using the very same paradigm that explains why things fall to Earth is truly a great achievement in human thought. Newton (apocryphally after seeing an apple fall from a tree) called this his *law of universal gravitation*, with emphasis on "universal," as it points out that the law applies both on Earth and in the heavens.

The key to Newton's idea is that the gravitational force actually exists between two objects and depends upon the masses of each and their separation in space. The Earth is no more special than the apple – both attract each other with equal force (which we know from the third law already), and the magnitude of that force depends upon their masses and their separations.

This actually does not fit well with our current understanding of the gravity force. In particular, we have been saying that the force equals mg, even as the height (distance from the Earth's surface) changes, so how is this dependent upon separation? First of all, it turns out that it is not the separation of the outer surfaces of the objects that matters, but rather their centers. In fact, it is even more complicated than that, so to simplify it, let's first just assume that the two gravitating objects are very small (point masses), so that their separation is well-defined:

Figure 7.1.1 – Two Point Masses Separated by *r*



In this case, the basis for Newton's law of universal gravitation can be described as follows:

- the force is exclusively attractive – experimentally, we only see gravity act as a "pull."
- the strength of the force grows linearly with the amount of each mass – experimentally, we find that the force doubles when we double either of the two masses involved, triples when either mass is tripled, and so on.
- the strength of the force varies in inverse proportion with the square of the separation – experimentally, we find that doubling the separation of the two objects reduces the force by a factor of four, tripling the separation reduces the force by a factor of nine, and so on.

Assuming these are the only factors that come into play for gravity (for example, the relative motions of the two objects doesn't affect the force), then we can write a proportionality for the magnitude of the gravity force between two point masses:

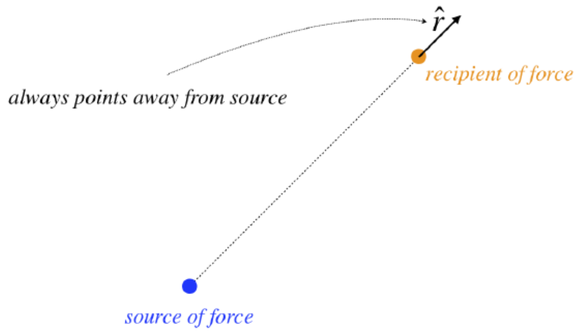
$$F_{gravity} \sim \frac{m_1 m_2}{r^2} \tag{7.1.1}$$

This satisfies all the criteria given above. All that remains is to turn it into an equality by inserting a multiplicative constant that turns it into units of force, with the correct observed magnitude:

$$F_{gravity} = \frac{G m_1 m_2}{r^2}, \qquad G \equiv 6.67 \times 10^{-11} \frac{Nm^2}{kg^2} \tag{7.1.2}$$

With the strength of the force, and the knowledge that it is attractive in nature, we have Newton's law of gravity. As usual, we would like to write this in a compact way that included the direction – as a vector equation. To do this, we temporarily discard the "equal partner" view, and treat one of the point masses as the source of the force (the object that the force is "by"), and the other as the recipient (the object the force is "on"). As an object's motion is determined by the forces on it, we treat the source of the force as the "origin," and define the position vector as pointing from the origin to the object on which the force acts. Therefore the unit vector of the position vector at the affected object always points away from the source of the gravity force.

Figure 7.1.2 – Defining Position Unit Vector for Gravity



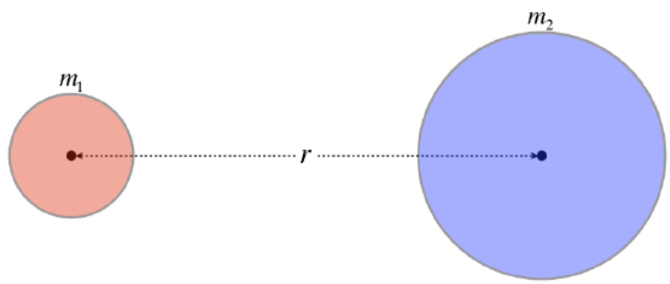
With this defined, we see that the attractive force on the recipient points in the $-\hat{r}$ direction, giving us a nice, compact vector equation for Newton's law of gravity:

$$\overrightarrow{F}_{gravity} = \frac{Gm_1m_2}{r^2}(-\hat{r}) \tag{7.1.3}$$

Spherical Bodies

Now of course we really aren't especially interested in gravity between point masses, when everything we see has some extension in space. So really what we have to do is treat two bodies as collections of point particles, all of which are attracting every other point particle. But this is quite cumbersome, and leads to all sorts of integral calculus. For our purposes, we will simply state the result that in the case of spherically-symmetric objects, they can be treated (in terms of gravitation) as if all of their mass were concentrated at their centers.

Figure 7.1.3 – Spheres Gravitate Like Point Masses at Their Centers



[*Note: This "spherical symmetry" does not require that the density of the spheres be uniform – the density can still vary radially. So the spheres can (for example) be more dense near their centers than near their surfaces, but the density cannot vary with the polar or azimuthal angles. That is, sampling the density throughout the sphere must reveal the same density everywhere that the distance from the sphere's center is held constant.*]

This turns out to be a convenient consequence of the inverse-square law, as you will no doubt examine in greater detail in more advanced math & physics classes. This result is something we will exploit greatly (at least as an approximation), since planets and stars are very close to being spherical.

Gravity at the Earth's Surface

Imagine a very small object (which can be effectively treated as a point object) was pulled toward a large spherical body, and stopped when it reached its surface. In that case, the gravitational force would be calculated using the radius of the large spherical object. Now we'll let that large spherical object be the Earth, and let the small object be a human (you).

Figure 7.1.4 – Universal Gravitation at the Earth's Surface

[*Note that the separation R_E points to the center of the Earth, not to a point on the Earth's surface just off the coast of Florida.]*

We can use now Newton's law of gravity to compute the force exerted on you. All we need is the radius of the Earth ($6.38 \times 10^6 m$), and the mass of the Earth ($5.98 \times 10^{24} kg$):

$$your\ weight = \frac{Gm_1m_2}{R_E^2} = \frac{\left(6.67 \times 10^{-11} \frac{Nm^2}{kg^2}\right)\left(5.98 \times 10^{24} kg\right) m}{\left(6.38 \times 10^6 m\right)^2} = \left(9.80 \frac{m}{s^2}\right) m = mg \tag{7.1.4}$$

So now you know where our constant g comes from. Now you might be concerned that our projectile calculations have not been accurate, because g is only correct at the surface of the earth, and projectiles might go quite high. Let's look at an example – how much does the gravitational force decrease when we go high up in the sky in a commercial airline? Commercial flights typically fly at an altitude of about $10,000m$ (about $33,000 ft$), so making the adjustment to the gravitational force gives:

$$your\ weight\ in\ airplane = \frac{\left(6.67 \times 10^{-11} \frac{Nm^2}{kg^2}\right)\left(5.98 \times 10^{24} kg\right) m}{\left(6.38 \times 10^6 m + 1 \times 10^4 m\right)^2} = \left(9.77 \frac{m}{s^2}\right) m \tag{7.1.5}$$

It's hardly noticeable. If you weighed yourself on earth and were 150 lbs, then in the plane the scale would read 149.5lbs. Okay, so let's go to a place where we know the distance makes a big difference – all the way into outer space to the international space station (ISS). The altitude in this case is about $400,000m$

$$your\ weight\ in\ ISS = \frac{\left(6.67 \times 10^{-11} \frac{Nm^2}{kg^2}\right)\left(5.98 \times 10^{24} kg\right) m}{\left(6.38 \times 10^6 m + 4 \times 10^5 m\right)^2} = \left(8.68 \frac{m}{s^2}\right) m \tag{7.1.6}$$

Wait just a minute… How can those people be floating around their space station if they have only lost about 11% of their weight?

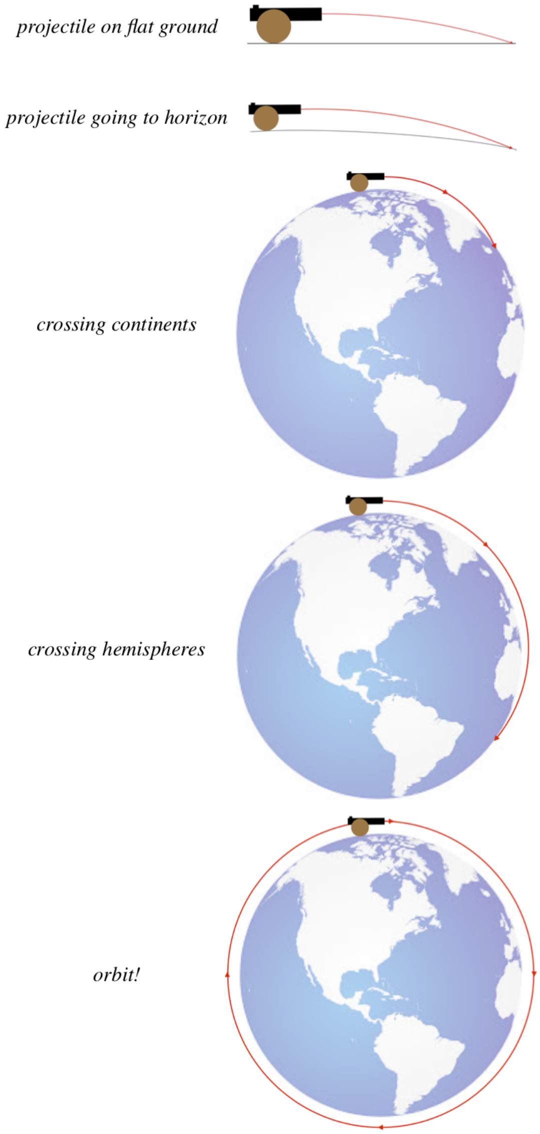
Free-Fall and Orbits

Suppose you are at the top floor of a skyscraper in an elevator when suddenly, tragically, the cable breaks. Assuming you could see past what I can only assume would be your abject terror, what would you see going on around you? The other screaming people around you, the hat on your head, and the penny that was on the floor would all be accelerating at the same rate, g. Since nothing is accelerating faster than anything else, if you hold out your pencil and release it, it doesn't drop to the floor of the elevator – it just "floats" there in front of you. In effect, the entire contents of the elevator is experiencing zero gravity.

Even conceding that being in a room in free-fall is equivalent to zero gravity, the space station is not plummeting to Earth, so how does it apply? Well, we know from our study of projectile motion that the horizontal motion of an object doesn't take away from the fact that it is in free-fall vertically. Indeed, there exist companies that fly planes in parabolic projectile trajectories so that the passengers can experience weightlessness for a couple minutes (before they have to pull out of the dive). So in fact if our elevator were a projectile, we would have the same zero gravity experience. Newton knew this, and came up with the following incredibly clever thought experiment:

If we fire a cannon horizontally, the cannonball follows the usual parabolic path, landing some distance away. If we increase the muzzle velocity, it goes farther before landing. Increase the muzzle velocity even more, and the landing point approaches the horizon. As we keep going this way, the projectile "falls over" the curvature of the Earth, and when the speed is finally fast enough, it never actually lands! Orbits are just the most extreme case of projectile motion.

Figure 7.1.5 – Newton's Cannon



So if being inside a container in free-fall is equivalent to being weightless, then a mouse inside a hollow cannonball fired by Newton's cannon would conclude that there is "no gravity," because the orbiting cannonball is a projectile in free-fall at all times. The astronauts on the space station experience weightlessness not because the Earth’s gravitational influence is zero out there, but because that influence is the same on everything in the station, and everything is therefore in free-fall at the same rate. Indeed, if there was no gravity out there, then there would be no force to keep the space station moving in a circle, and it would fly away from Earth!

To understand how important this argument was in the context of his time, Newton used it to explain how a single phenomenon (gravity) could explain both terrestrial (projectile) motion and heavenly (orbital) motion at the same time. What is more, he backed it all up with mathematics! His law of universal gravitation predicted to very high precision the motions of the planets, even as it predicts the motions of cannonballs.

7.2: Kepler's Laws

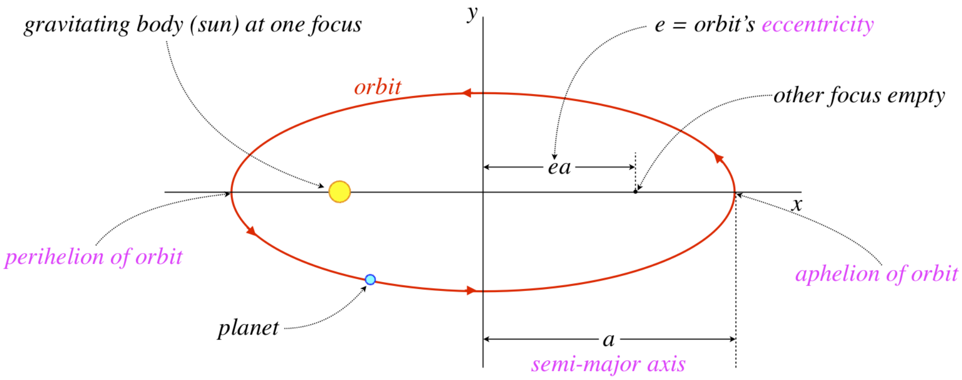
More than 20 years before Newton was born, a fellow named Johannes Kepler took a shot at explaining the orbits of the planets. He too posited that physical laws might be able to explain the motions, but didn't possess the tools (mathematical and physical) at Newton's disposal decades later (though admittedly, Newton did develop these tools for himself). Instead, what Kepler had were the remarkably detailed and accurate measurements of planetary motions made by an astronomer named Tycho Brahe, which he used to look for patterns in the motions. Amazingly, he found that the planets indeed moved with mathematical precision, and published his three laws of celestial motion, all of which are in exact accordance with the law of universal gravitation. While reading about his three laws, consider what a monumental accomplishment this was. Tycho Brahe's data detailed the motions of the planets *as he viewed them from Earth* (which itself is orbiting the sun).

Kepler's First Law

Kepler's First Law: The paths of bodies trapped in orbits form closed ellipses, with the gravitating body at one of the foci.

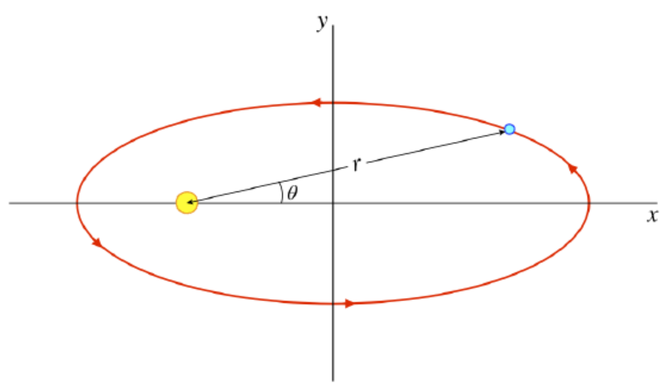
The many elements of an ellipse and how an orbit fits into the picture are expressed in Figure 7.2.1.

Figure 7.2.1 – Elliptical Orbit



There are many ways to describe such an orbit mathematically. A common way is to write the distance between the two bodies as a function of the angle that the line between them makes with the major (longer) axis:

Figure 7.2.2 – Polar Coordinate Description of Ellipse



The formula for the ellipse in these coordinates is:

$$r(\theta)=\frac{a\left(1-e^2\right)}{1-e\cos\theta} \tag{7.2.1}$$

[*Note: It is also possible to measure the angle in the opposite direction (with $\theta=0$ corresponding to the perihelion), in which case the denominator is a sum rather than a difference.*]

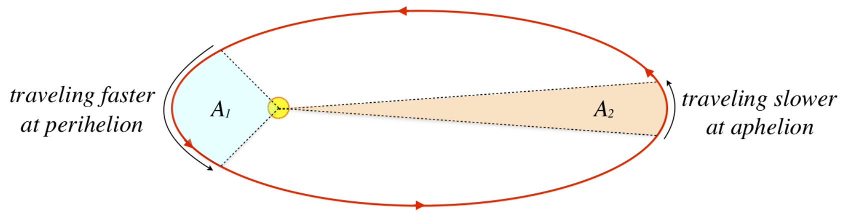
Circles are special cases of ellipses (eccentricity equal to zero), so naturally circular orbits are possible. Notice that if we plug in $e=0$ above, we get the simple orbit equation $r=a$. With a great deal of mathematics (first surmounted by Isaac Newton), one can can show that in fact this is a natural consequence of the inverse-square force law we already stated for gravitation. While we won't go quite so far as to perform this derivation, below we will make a closer examination of features of the ellipse (which we have expressed as a purely mathematical object here) in terms of physical quantities.

Kepler's Second Law

Kepler's Second Law: An orbiting planet sweeps out equal areas in equal times during its orbit.

Kepler noticed that the planet moved faster when it was near the perihelion than when it was near the aphelion, and through painstaking examination of the data determined that in fact the amount of area the orbit sweeps out in a given period of time is the same everywhere in the orbit.

Figure 7.2.3 – Equal Areas in Equal Times



Note that while the diagram compares areas swept out through the perihelion and aphelion, the result applies to any part of the orbit – if we wait the same period of time, the area swept out will be the same: $A_1 = A_2$. We will soon look into the physical aspects of gravitational orbits that lead to this result, but again one can't help but marvel at what it must have taken to derive this remarkable discovery from the raw data.

Kepler's Third Law

Kepler's Third Law: For every object orbiting the same gravitational source, the ratio of the cube of the semi-major axis of the orbital ellipse and the square of the orbital period is the same constant: $\frac{a^3}{T^2} = constant$.

While the first law makes a general statement about all gravitational orbits, and the second law relates two different parts of a specific gravitational orbit, the third law gives a way of comparing different orbits of the same gravitating object (in Kepler's case, this gravitating object was the sun). Of the three laws, this one has the greatest practical value, because it means that even without knowing anything about the law of gravitation, one can make a prediction about one orbiting body based on observations of another orbiting body, if both are going around the same gravitating object. We will see what this mysterious constant is in terms of physical details, but to solve certain problems, knowing the actual value of the constant isn't necessary.

Example 7.2.1

An astronomer notices that an asteroid is positioned such that the Earth is directly between it and the Sun. It has a roughly circular orbit (like the Earth), which is in the same plane and in the same direction as the Earth. This asteroid is far away from the Earth – about 8 times farther from the Earth than the distance separating the Earth and the Sun. These are all approximations, but about how long will this astronomer have to wait see these three bodies reach these same positions?

Solution

For a circular orbit, the semi-major axis of the orbit is simply the radius of the orbit, and since the asteroid and Earth are both orbiting the same gravitational source (the sun), then Kepler's third law results in the same constant for both:

$$\frac{R^3_{earth}}{T^2_{earth}} = constant = \frac{R^3_{asteroid}}{T^2_{asteroid}} \quad \Rightarrow \quad T_{asteroid} = \left(\frac{R_{asteroid}}{R_{earth}}\right)^{\frac{3}{2}} T_{earth}$$

The period of the Earth's orbit is exactly 1 year, and the asteroid is 9 times farther from the sun as the Earth, so the time it will take the asteroid to come back to the same spot will be:

$$T_{asteroid} = 27 years$$

Of course, in 27 years, the Earth will also be in the spot where it started, so this is our answer.

Note that if the question asked for the time that elapses before the sun, Earth, and asteroid are all aligned again (which is different from reaching their original positions), the answer is different: When they realign for the first time, the Earth will have completed one orbit plus a bit more, while the asteroid will have completed a fraction of an orbit. Let's call the angle that the asteroid moves through in that first year $\Delta\theta$. The Earth catches up to it after a full revolution, so the angle the Earth moves through is:

$$\theta_{earth}=2\pi+\Delta\theta$$

The Earth is moving 27 times as fast as the asteroid, so in this equal time frame the Earth has moved through an angle 27 times as great as the asteroid, which gives:

$$2\pi+\Delta\theta=27\Delta\theta \quad\Rightarrow\quad \Delta\theta=\frac{2\pi}{26}$$

That is, the asteroid completes 1/26th of its orbit at the point when the Earth catches up to it. The asteroid's orbit takes 27 years, so the first alignment occurs at:

$$\Delta T=\frac{27years}{26}=1year,14days$$

A nice application of Kepler's 3rd law involves man-made satellites that orbit the Earth. Telecommunications satellites we like to remain at a single position in the sky, so that we don't have to turn our satellite dishes to find them – we just point them in the right direction and leave it. To accomplish this, we need two things: The satellite has to be directly above the equator, and it has to be orbiting the Earth in the same direction that the Earth is rotating, with an orbital period of exactly 1 day. This is known as a geostationary orbit. We can use Kepler's 3rd law to determine how high off the Earth's surface this satellite needs to be.

Reconciling Kepler's Laws with Universal Gravitation

There are a couple of things we can say about the physics of gravitational orbits. First, gravity is a conservative force, which means that the mechanical energy of the system is conserved. We don't yet know how to describe the potential energy due to the gravitational force (hopefully it is clear that our old "$U(y)=mgy+U_o$" treatment is no longer adequate, since this results in a constant force), but we will look at this in Section 7.3. The point is that the mechanical energy is a "constant of the motion," which we can use, for example, to describe the speed of the orbiting body as a function of the distance from the gravitational source.

Consider the Earth + sun system. The fraction of this system's mass that belongs to the Earth is about 3×10^{-6}, which means that the distance from the center of the sun to the center of mass of the system is this fraction multiplied by the (on average) 93 million miles separating the two bodies. The distance from the center of the sun to the center of mass of the system is therefore:

$$r_{cm}=\left(3\times10^{-6}\right)\left(93\times10^{6}\;miles\right)=279\;miles \tag{7.2.2}$$

The radius of the sun is about 432,000 miles, so the center of mass of the system lies less than one tenth of one percent of the sun's radius from the sun's center. It's therefore a pretty good approximation to treat the center of the sun as a fixed point. [*Note: Even the center of mass of the Jupiter + sun system barely lies outside the sun's radius, even though Jupiter is much more massive than Earth, and is much farther away.*] The gravitational force is directed at this fixed point, so it constitutes a central force. As we found in Section 6.2, central forces have the property of conserving the angular momentum of the system (since they produce no torque). We therefore conclude that like the mechanical energy, the angular momentum of an orbiting body is also a constant of the motion, when the gravitating body is significantly more massive than the orbiting body. [*Note: The angular momentum of the whole system is always conserved, even when the two masses are comparable, but in that case we can't treat one object as orbiting another stationary one, which means we have to consider the motion of both objects, complicating our picture.*]

Kepler's First Law

Showing that elliptical orbits are a direct result of the law of universal gravitation is a mathematical exercise that is somewhat beyond the scope of this work. While this derivation could nevertheless be included here, the value of doing so is minimal, and will therefore be left for the reader to explore in an upper-division treatment of classical mechanics. Instead, we will look at only small – but very instructive – aspects of Kepler's first law.

Never mind that we have no reason to expect that orbits will be elliptical... Why would we even expect them to be *closed*? That is, it certainly isn't clear that when the polar angle θ changes by 2π, that the orbiting body's distance from the gravitating body will be the same. It turns out, however, that this element of gravitational orbits is not hard to demonstrate. Start by applying Newton's second law to the orbiting body on which a net force due to gravity is acting:

$$m\frac{d\overrightarrow{v}}{dt}=-\frac{GMm}{r^2}\hat{r} \tag{7.2.3}$$

We can rewrite this in terms of how the velocity changes with the angle θ by using the chain rule:

$$m\frac{d\overrightarrow{v}}{d\theta}\frac{d\theta}{dt}=-\frac{GMm}{r^2}\hat{r} \tag{7.2.4}$$

The angular momentum of the orbiting body can be written in terms of its mass, its angular velocity $\omega=\dfrac{d\theta}{dt}$, and its distance from the fixed point:

$$L=mr^2\omega \tag{7.2.5}$$

This angular momentum is a constant of the motion (i.e. it is conserved throughout the orbit), so we find that the rate at which the velocity vector changes with respect to θ is a vector with a constant magnitude:

$$\frac{d\overrightarrow{v}}{d\theta}=-\frac{GMm}{L}\hat{r} \tag{7.2.6}$$

We can write the position unit vector in terms of the angle relative to a cartesian coordinate system, as we did in Equation 1.6.11:

$$\frac{d\overrightarrow{v}}{d\theta}=-\frac{GMm}{L}\left(\cos\theta\;\hat{i}\;+\sin\theta\;\hat{j}\right) \tag{7.2.7}$$

Integrating over the angle gives the velocity vector as a function of θ (and an undetermined constant of integration $\overrightarrow{v}_o$):

$$\overrightarrow{v}(\theta)=-\frac{GMm}{L}\left(\sin\theta\;\hat{i}-\cos\theta\;\hat{j}\right)+\overrightarrow{v}_o \tag{7.2.8}$$

From this result we can conclude that the magnitude and direction of the velocity return to the same value every time θ changes by 2π. This means that the kinetic energy returns to its same value periodically as well. But the mechanical energy of this system is conserved, so the potential energy also returns to its value with the same periodicity. But (as we will see in the next section), the potential energy is defined by the separation of the two masses, so this separation also returns every time θ changes by 2π. Well, if every time the angle changes by 2π the orbiting body is the same distance away from the gravitating body, is moving at the same speed, and is moving in the same direction, then clearly its motion is being repeated – the orbit is closed.

Notice that if dependence on r in the law of gravitation was anything other than inverse-square, then the r's would not cancel as they did in reaching Equation 7.2.6 (the angular momentum would still be the same constant of the motion, as the force would still be central), which would give an equation for the velocity that depends on both r and θ. This ruins the argument above, and the orbit would not be closed.

Digression: Orbits Are Not Quite Closed After All

As amazing as Newton's accomplishment was with his theory of gravity, roughly 230 years later a fellow named Albert Einstein provided an improved theory, called the General Theory of Relativity. It had been known for some time that observations of the orbit of Mercury indicated that its orbit was in fact not closed. For a long time it was thought that the discrepancy was the result of an unseen planet or other gravitating body that was pulling Mercury off course, but Einstein's theory showed that the inverse-square theory of Newton, while a very good approximation, is not quite right, and his new theory predicted Mercury's motion perfectly.

While we are skipping the mathematics detailing how we get to the equation of the ellipse, we can still extract some information from Equation 7.2.8 that will be useful to us later. To simplify the discussion that follows, we will assume that we already know that the elliptical orbit is the result.

We haven't yet defined the *orientation* of the (x,y) coordinate system used in Equation 7.2.8 – so far we have only required that the origin be at the gravitating body. Let's choose the x-axis to lie along the major axis such that the point of maximum separation (aphelion) lies on the positive side of the x-axis (giving us Figure 7.2.2). With these axes, it is clear that at the point $\theta=0$, the velocity of the orbiting body is in the $+\hat{j}$ direction. Looking at Equation 7.2.8, we see that this means that the constant vector $\overrightarrow{v}_o$ must point parallel to the minor axis, otherwise it would give the body a component of velocity along the x direction. But which way does this constant vector point, $+\hat{j}$ or $-\hat{j}$?

Consider the velocity of the orbiting body at the perihelion ($\theta = \pi$). In this case, the object is now moving in the $-\hat{j}$ direction, but because it is closer to the reference point at the gravitating body, it must be moving faster to conserve angular momentum. For the constant vector $\overrightarrow{v}_o$ to make the orbiting body faster at ($\theta = \pi$) than at ($\theta = 0$), we must have:

$$\overrightarrow{v}_o = v_o\left(-\hat{j}\right) \quad \Rightarrow \quad \begin{cases} \theta = 0 \quad (aphelion) \quad v_{min} = \dfrac{GMm}{L} - v_o \\ \theta = \pi \quad (perihelion) \quad v_{max} = \dfrac{GMm}{L} + v_o \end{cases} \tag{7.2.9}$$

We can relate the maximum and minimum speeds of the orbit using angular momentum conservation. Looking at Figure 7.2.1, we see that the value of $r_\perp$ for the aphelion is: $a + ea = a(1+e)$. At the perihelion: $r_\perp = a(1-e)$. Setting equal the angular momenta at these two positions in the orbit gives:

$$L_{aphelion} = L_{perihelion} \quad \Rightarrow \quad mv_{min}\left[a(1+e)\right] = mv_{max}\left[a(1-e)\right] \quad \Rightarrow \quad v_{min} = v_{max}\left(\frac{1-e}{1+e}\right) \tag{7.2.10}$$

Using this result and Equations 7.2.9, we can eliminate the troublesome v_o, and get:

$$\begin{aligned} v_{max} &= (1+e)\frac{GMm}{L} \\ v_{min} &= (1-e)\frac{GMm}{L} \end{aligned} \tag{7.2.11}$$

It will also be useful to have an expression for the angular momentum in terms of the masses, eccentricity, and major axis. To get this, multiply the first of the Equations 7.2.11 by the orbiting body's mass and $r_\perp$ [for v_{max} this is: $a(1-e)$] to get:

$$L = mv_{max}r_\perp = m\left[(1+e)\frac{GMm}{L}\right]\left[a(1-e)\right] \quad \Rightarrow \quad L^2 = a\left(1-e^2\right)GMm^2 \tag{7.2.12}$$

Putting this back into Equation 7.2.1 simplifies it a bit in terms of physical constants, putting it in a form that will be useful later:

$$r(\theta) = \left(\frac{L^2}{GMm^2}\right)\frac{1}{1-e\cos\theta} \tag{7.2.13}$$

One last thing to note before moving on to Kepler's second law. Looking at Equation 7.2.8, we see that if v_o happens to equal zero, then the velocity vector has a constant magnitude, and its direction is always perpendicular to $\hat{r}$ (which can be confirmed quickly by performing a dot product). So a non-zero constant of integration is responsible for making an otherwise circular orbit eccentric.

Example 7.2.2

Show that the distance of closest approach (the perihelion distance) is given by:

$$r_{min} = \frac{L^2}{GMm^2}\left(\frac{1}{1+e}\right)$$

Do this in two different ways:

a. *Using calculus and Equation 7.2.13. [Note: There are two angles that result in extrema, but only one gives a minimum.]*
b. *Using the equation for angular momentum at r_{min} and one of the Equations 7.2.11.*

Solution

a. We seek the minimum value of $r(\theta)$, so we start by finding the value of θ where this minimum occurs:

$$0 = \left(\frac{L^2}{GMm^2}\right)\frac{d}{d\theta}\left(\frac{1}{1-e\cos\theta}\right) \quad \Rightarrow \quad 0 = \frac{\sin\theta}{1-e\cos\theta} \quad \Rightarrow \quad \theta = 0 \;\; or \;\; \pi$$

Note that $\theta = 0$ minimizes the denominator, so it gives a maximum for r, not a minimum. On the other hand, $\theta = \pi$ maximizes the denominator, and therefore gives a minimum. This result makes sense when we look at how θ is defined in Figure 7.2.2. Plugging $\theta = \pi$ into Equation 7.2.13 gives the desired answer.

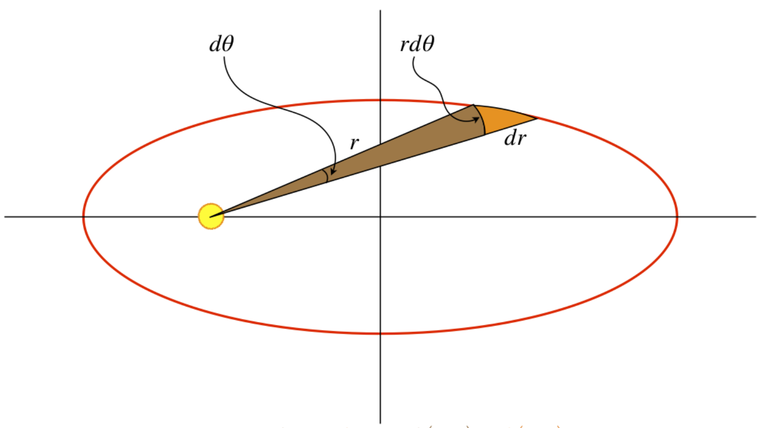
b. The angular momentum at the point closest approach will involve the maximum speed, since it is conserved throughout the orbit, so using the expression for the maximum speed in Equations 7.2.11 we get:

$$L = mv_{max}r_{min} \quad \Rightarrow \quad r_{min} = \frac{L}{mv_{max}} = \frac{L}{m\left((1+e)\dfrac{GMm}{L}\right)} = \frac{L^2}{GMm^2}\left(\frac{1}{1+e}\right)$$

Kepler's Second Law

We look next at Kepler's equal-areas-swept-out-in-equal-times law. The geometry of measuring areas swept out of ellipses is impossibly difficult to do mathematically, but an *infinitesimal* amount of area swept out in an infinitesimal time period is something we can do. Figure 7.2.4 shows how we can mathematically describe the area swept out in an infinitesimal period of time. The amount swept out can be broken into two triangles. Okay, so the sides of the triangles are curved, but when the curves are infinitesimal in length, the amount they differ from straight lines is insignificant. The area of the two triangles are color-coded in the diagram. We notice that while both have infinitesimal areas, the orange triangle includes a product of *two* infinitesimal quantities. When the area is then divided by a small time span and the limit is taken as the time span goes to zero, the ratio of $d\theta$ and dt approaches a finite value (specifically, the angular velocity ω at that moment in time), but in the second term there is still another infinitesimal dr that goes to zero in the limit. In other words, the orange triangle contributes nothing to the area swept out in the infinitesimal time span dt. With a little mathematical manipulation, we see that the rate at which area is swept out is the angular momentum of the orbiting body divided by twice its mass. We know that both the angular momentum and the mass remain constant for the orbit, so the rate at which area is swept out also remains constant. Kepler's second law is equivalent to angular momentum conservation.

Figure 7.2.4 – Kepler's Second Law Expresses Angular Momentum Conservation



$$dA = dA_1 + dA_2 = \tfrac{1}{2}b_1h_1 + \tfrac{1}{2}b_2h_2 = \tfrac{1}{2}\left(rd\theta\right)r + \tfrac{1}{2}\left(rd\theta\right)dr$$

$$\frac{dA}{dt} = \underbrace{\frac{1}{2}r^2\frac{d\theta}{dt}}_{finite} + \underbrace{\frac{1}{2}r\cancel{dr}\frac{d\theta}{dt}}_{infinitesimal} = \frac{1}{2}r^2\omega = \frac{mr^2\omega}{2m} = \frac{L}{2m} = constant$$

Kepler's Third Law

It is quite straightforward to show that Kepler's third law holds for circular orbits, so let's do that first. The speed is constant for the entire orbit, it equals the circumference of the orbit divided by the time it takes for a full orbit:

$$v = \frac{2\pi R}{T} \tag{7.2.14}$$

We can use the fact that the gravitational force is causing centripetal acceleration to get the following expression for the square of the constant speed of the orbiting body:

$$\frac{GMm}{R^2} = m\frac{v^2}{R} \quad \Rightarrow \quad v^2 = \frac{GM}{R} \tag{7.2.15}$$

Plugging Equation 7.2.14 into Equation 7.2.15 and doing some algebra gives Kepler's third law, with the semi-major axis equaling the radius of the circular orbit (zero eccentricity):

$$\left(\frac{2\pi R}{T}\right)^2 = \frac{GM}{R} \quad \Rightarrow \quad \frac{R^3}{T^2} = \frac{GM}{4\pi^2} = constant \tag{7.2.16}$$

This gives us not only that the ratio is a constant, but specifically what the constant is. As we can now confirm, this constant depends only upon the mass of the gravitating body.

It's quite remarkable that this law holds equally well for elliptical orbits, where R is replaced by a. We can show this by starting with a result we found from Kepler's second law. The rate at which area is swept out is constant, so the total area of the ellipse is this rate multiplied by the time of a full orbit. Reviewing our conic sections, we plug in the area of an ellipse, and get:

$$area\ of\ ellipse = \pi\ ab = \frac{dA}{dt}T = \frac{L}{2m}T \tag{7.2.17}$$

The quantity b is the length of the semi-minor axis, which is related to the length of the semi-major axis in terms of the eccentricity:

$$b^2 = a^2\left(1-e^2\right) \tag{7.2.18}$$

Squaring Equation 7.2.17 and eliminating b^2 using Equation 7.2.18 gives us:

$$\pi^2a^4\left(1-e^2\right) = \frac{L^2T^2}{4m^2} \tag{7.2.19}$$

We had the good foresight to derive Equation 7.2.12, so plugging L^2 from that equation into Equation 7.2.19 gives us our answer:

$$\pi^2a^4\left(1-e^2\right) = \left[a\left(1-e^2\right)GMm^2\right]\frac{T^2}{4m^2} \quad \Rightarrow \quad \frac{a^3}{T^2} = \frac{GM}{4\pi^2} \tag{7.2.20}$$

Example 7.2.3

Attractive central forces that are not inverse-square do not produce closed orbits except when the orbit happens to be circular. Whenever there is a closed orbit, a result like Kepler's third law (which relates the orbit period to the radius of the circular orbit) will be the result. Derive the effective Kepler third law for an attractive central force that varies as the inverse-cube of the separation.

Solution

Following the process used above for a gravitational circular orbit, we get:

$$\left.\begin{array}{l} \dfrac{k}{R^3} = m\dfrac{v^2}{R} \\ v = \dfrac{2\pi R}{T} \end{array}\right\} \quad \Rightarrow \quad \boxed{\frac{R^4}{T^2} = \frac{k}{4\pi^2 m} = constant}$$

7.3: Energy in Gravitational Systems

Gravitational Potential Energy

We showed in Section 3.2 that our terrestrial model of gravity is a conservative force, and it certainly seems reasonable to assume that universal gravitation is as well, but really we should check to see if this is the case. In Section 3.6 we outlined a procedure for determining whether a force is conservative or not – basically it consists of trying to construct a potential energy function whose gradient equals the force, and if we succeed, then the force is conservative. If we can show that it is impossible to do this, then the force is non-conservative. Let's see what happens when we bring this to bear on gravitation.

Let's start by writing the gravity force in cartesian coordinates:

$$\overrightarrow{F}\left(x,y,z\right)=-\frac{GMm}{r^2}\hat{r}=-\frac{GMm}{r^3}\overrightarrow{r}=-\frac{GMm}{\left(x^2+y^2+z^2\right)^{\frac{3}{2}}}\left(x\hat{i}+y\hat{j}+z\hat{k}\right) \tag{7.3.1}$$

Consider next the following partial derivative:

$$\frac{\partial}{\partial x}\left(x^2+y^2+z^2\right)^{-\frac{1}{2}}=-\frac{1}{2}\left(x^2+y^2+z^2\right)^{-\frac{3}{2}}\left(2x\right)=\frac{-x}{\left(x^2+y^2+z^2\right)^{\frac{3}{2}}} \tag{7.3.2}$$

This is precisely the x-component of the gravitational force. Obviously partial derivatives with respect to y and z yield similar results – the y and z components of the gravitational force. This means we can immediately define a potential energy function whose negative gradient is the force:

$$U\left(x,y,z\right)=\frac{-GMm}{\left(x^2+y^2+z^2\right)^{\frac{1}{2}}}+U_o \quad\Rightarrow\quad U\left(r\right)=\frac{-GMm}{r}+U_o \tag{7.3.3}$$

We could have saved ourselves a lot of trouble if we happened to know a useful fact from vector calculus: The gradient of a function that is purely a function of r can be written as:

$$\overrightarrow{\nabla}U\left(r\right)=\frac{d}{dr}U\left(r\right)\hat{r} \tag{7.3.4}$$

So:

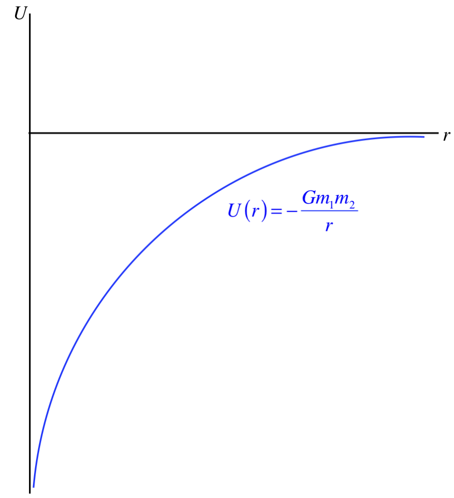
$$\overrightarrow{F}=-\overrightarrow{\nabla}U\left(r\right)=-\frac{d}{dr}\left[\frac{-GMm}{r}+U_o\right]\hat{r}=-\frac{GMm}{r^2}\hat{r} \tag{7.3.5}$$

As with other potential energy functions for another physical system that we have discussed (the intermolecular forces described by the Lennard-Jones potential), we typically choose our arbitrary constant of integration such that the potential energy falls off to zero at infinity. In the case of our gravitational potential energy, this gives what we will use for the potential energy function henceforth:

$$U_{grav}\left(r\right)=\frac{-GMm}{r} \tag{7.3.6}$$

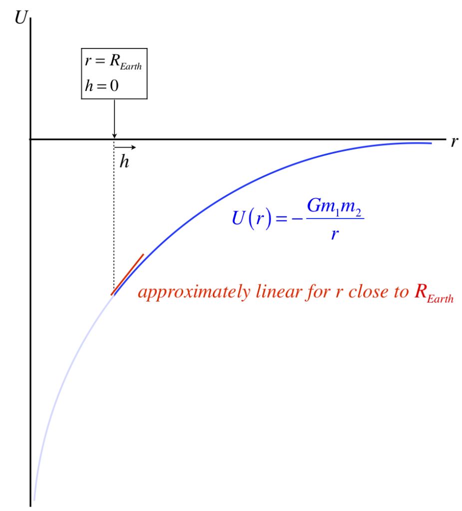
A graph of this function looks like this:

Figure 7.3.1 – Gravitational Potential



This function is significantly different from the "mgy" that we have been using up to now. To reconcile these two models, we need to make a note of the restriction of our terrestrial model of gravity – it holds for a region very close to the surface of the Earth, $r = R_{Earth}$. Calling "h" the height of the object from the ground, the gravitational potential function above is cut-off at ground level. If we restrict ourselves to r values close to this point, the curve above is very close to a straight line:

Figure 7.3.2 – Gravitational Potential Near Earth's Surface



The negative slope of the potential energy curve is the force, so the slope of the straight line approximating the curve near $r = R_{Earth}$ is the constant force in that region. Taking the negative derivative gives Newton's gravitational force, and when evaluated at $r = R_{Earth}$, this force comes out to be mg, as we saw in Equation 7.1.4.

Bound and Unbound Gravitational Systems

With a graph of the potential energy which goes to zero at infinity, we are naturally drawn again to energy diagrams. There is a problem with jumping straight into this, however. Energy diagrams require 1-dimensional motion, and while $U(r)$ is a function of a single variable, the motion is not 1-dimensional. To see where things break down, consider a closed elliptical orbit. As the planet moves toward the perihelion, the value of r gets smaller. If Figure 7.3.1 were the potential curve for an energy diagram, then when the planet is moving toward smaller values of r, it would keep speeding up indefinitely as it approaches $r = 0$. Okay, so in practice it would hit the surface of the sun before it could "accelerate indefinitely," but still this does not represent orbital motion. In short, there is no part of this potential energy graph that represents the turnaround point that is the perihelion.

Fortunately, we have a nice trick to take care of this shortcoming. When we draw energy diagrams, the kinetic energy comes from the motion along the direction parallel to the one dimension. In this case of measuring energy along the radius, the kinetic energy for the energy diagram can only come from the part of the velocity that is *radial*. The total kinetic energy is the sum of the radial and tangential parts:

$$KE_{tot} = KE_{radial} + KE_{tangential} \tag{7.3.7}$$

The tangential speed multiplied by the mass and the distance from the center is the angular momentum, which is a constant of the motion, so we have:

$$\left. \begin{array}{l} KE_{tangential} = \frac{1}{2}mv^2_{tangential} \\ L = mv_{tangential}\,r \end{array} \right\} \quad \Rightarrow \quad KE_{tangential} = \frac{L^2}{2mr^2} \tag{7.3.8}$$

If we now construct the total energy of the system, we have:

$$E_{tot} = KE_{radial} + KE_{tangential} + U = \frac{1}{2}mv^2_{radial} + \frac{L^2}{2mr^2} - \frac{GMm}{r} \tag{7.3.9}$$

Note that this can now be *treated* as 1-dimensional system, by combining the last two terms into a single function of r that we call an *effective potential*:

$$U_{eff}\left(r\right) = \frac{L^2}{2mr^2} - \frac{GMm}{r} \tag{7.3.10}$$

Graphing this allows us to work exactly as before, though we have to keep in mind that whatever we determine the kinetic energy to be is really only a fraction of the kinetic energy. So, for example, at the turnaround points (where we have said the kinetic energy is zero), the kinetic energy is in fact the tangential term. This makes sense for closed orbits, since the orbiting body never actually stops moving entirely – it only stops moving *radially*. The graph of this effective potential has a rather familiar shape, and including the total system energy makes it an energy diagram.

Figure 7.3.3 – Energy Diagram of a Gravitating System

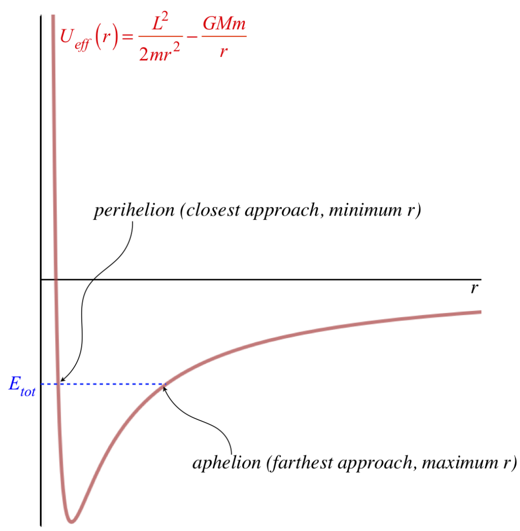


Figure 7.3.3 labels the turnaround points as the perihelion and aphelion, but there is much more we can extract from this diagram:

- **circular motion**

If the total energy lies at the bottom of the dip, then according to what we know about about energy diagrams, the kinetic energy is zero. But in this case, it means that the contribution to the kinetic energy by the radial component of velocity is zero. In other words, the orbit is circular. This fits with the fact that there is only one turnaround point, making the perihelion and aphelion the same distance.

- **semi-major axis**

The semi-major axis is the average of the perihelion and aphelion distances, so it is the value of r halfway between the two turnaround points on the graph.

- **eccentricity**

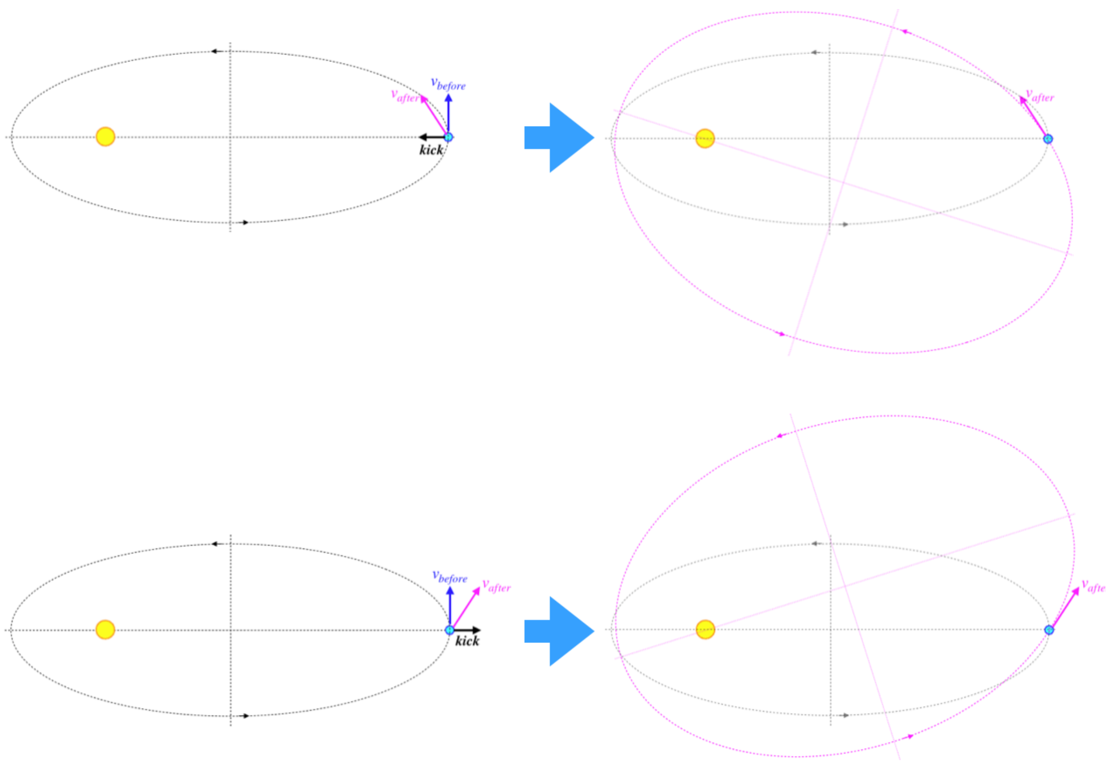
Looking at Figure 7.2.1, we can write the perihelion and aphelion distances in terms of the eccentricity and the semi-major axis, which we can then invert to get the eccentricity in terms of the perihelion and aphelion distances and the semi-major axis. Doing this gives:

$$e = \frac{r_{max} - r_{min}}{2a} \tag{7.3.11}$$

Looking at Figure 7.3.3, we see that raising the total energy (but keeping it negative, and doing it without increasing the angular momentum, which would change the graph) increases the length of the semi-major axis, but much more of this change comes from the increase of the aphelion distance than from the decrease of the perihelion distance (especially close to the horizontal axis). This results in an increase of eccentricity. How does one increase the total energy without increasing the angular momentum? By giving the orbiting body a "kick" that points radially (inward or outward). This exerts no torque (the force is parallel to the position vector), so the angular momentum doesn't change, but the push increases the total energy, because it adds a radial component of velocity without changing the tangential component.

When a radial kick is given, a polar angle of zero in the same coordinate system will no longer correspond to the aphelion of the orbit – the orbital ellipse will be rotated. If the kick is outward, then at the position of the kick the orbiting body is still getting farther from the gravitating body, which means it is heading for the new aphelion, and the orbit has rotated in the direction of the orbit. That is, if the orbit was clockwise, then the major axis rotates clockwise. For a radially inward kick, the orbiting body is now getting closer to the gravitating body, which means it is coming from the aphelion, so the major has rotated the direction opposite to the orbit direction.

Figure 7.3.4 – Effect of Radial "Kicks" to Orbits



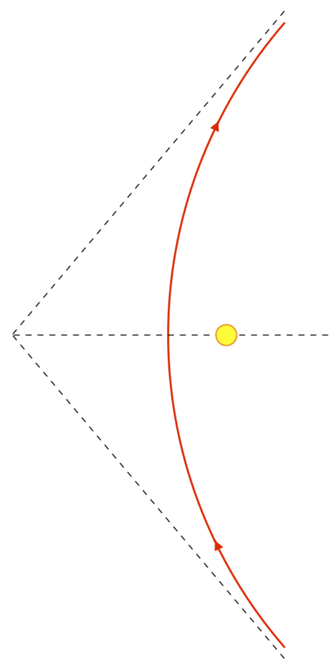
A tangential kick in the direction of motion would also raise the total energy, but it would also increase the angular momentum, increasing the positive term in the effective potential. This has the effect of raising the bottom of the curve, and if this is done properly (at the aphelion), the bottom of the curve comes up faster than the energy line goes up, bringing the eccentricity down. The eccentricity can even be brought all the way down to zero (i.e. a circular orbit) in this way (see Example 7.3.1).

- **hyperbolic trajectory – unbounded**

If the total energy is greater than zero, the orbiting body is not "bound" to the gravitating body. That is, after the orbiting body makes its closest approach, it zooms away, and although it slows down as it departs, it never stops moving away from the gravitating body. The height of the total energy line above the horizontal axis represents the finite kinetic energy of the orbiting body when it gets very far away, since the potential vanishes there. While the tangential part of the kinetic energy also goes to zero very far away, the angular momentum does not vanish – it remains conserved. With a finite speed and a non-zero angular momentum, the motion of the orbiting body must be asymptotically approaching a line that passes by the gravitating body, and we get what is called a *hyperbolic trajectory*. [*The word "trajectory" is generally preferred over "orbit," because the latter typically implies that the affected body is trapped by the gravitating body.*]

The orbit equation that we found for the elliptical orbit (Equation 7.2.13) still works, but for the hyperbolic trajectory the eccentricity is greater than 1.

Figure 7.3.5 – Hyperbolic Trajectory



- **parabolic trajectory – barely (un)bounded**

If the total energy of the system equals exactly zero, then the orbiting body just barely runs out of velocity as it reaches infinity. It technically is neither bound nor unbound – it is at the borderline between the two. In this case, the eccentricity equals exactly 1.

Example 7.3.1

An orbiting body is in an orbit where it is four times farther from the gravitating body at its aphelion than at its perihelion. By what percentage must its speed be increased at the aphelion to make its orbit circular?

Solution

Since orbits are closed, if we instantaneously speed up the orbiting body at its aphelion, it will return to that same point with the same speed and moving in the same direction (i.e. its motion will still be perpendicular to the radius). That means that when the orbiting body returns, it will either be once again at its aphelion (if the speed was not increased substantially), or it will be at its perihelion (if the speed was increased a great deal). There is also an "in-between" increase whereby the perihelion and aphelion are the same – a circular orbit. We can compute the required increase in speed by comparing the speed at the aphelion of an eccentric orbit to the speed of a circular orbit whose radius is the same as the semi-major axis of the eccentric orbit. Starting with Equation 7.2.11, we have:

$$v_{min} = (1-e)\frac{GMm}{L} = (1-e)\frac{GMm}{mv_{min}r_{max}} = \left(\frac{1-e}{1+e}\right)\frac{GM}{v_{min}a} \quad \Rightarrow \quad v_{min} = \sqrt{\left(\frac{1-e}{1+e}\right)\frac{GM}{a}}$$

For a circular orbit, the eccentricity is zero, which makes the ratio of the minimum velocity of the eccentric and circular orbits:

$$\frac{v_{circular}}{v_{min}}=\frac{\sqrt{\frac{GM}{R}}}{\sqrt{\left(\frac{1-e}{1+e}\right)\frac{GM}{a}}}=\sqrt{\left(\frac{1+e}{1-e}\right)\frac{a}{R}}$$

The new circular orbit must have a radius equal to the aphelion distance of the previous orbit, so:

$$\frac{v_{circular}}{v_{min}}=\sqrt{\left(\frac{1+e}{1-e}\right)\frac{a}{(1+e)\,a}}=\sqrt{\frac{1}{1-e}}$$

Now we need to know the eccentricity for an orbit where the aphelion distance is 4 times the perihelion distance. Writing these distances in terms of the semi-major axis gives:

$$\left.\begin{aligned}r_{min}&=(1-e)\,a\\ r_{max}&=(1+e)\,a\end{aligned}\right\}\quad\Rightarrow\quad 4=\frac{r_{max}}{r_{min}}=\frac{1+e}{1-e}\quad\Rightarrow\quad e=\frac{3}{5}$$

Plugging this in above gives us that the velocity must be increased by a factor of $\sqrt{\frac{5}{2}}$ *at the aphelion of the eccentric orbit to turn it into a circular orbit. This corresponds to a percentage increase of:*

$$\%\ increase=\left(\sqrt{\frac{5}{2}}-1\right)\times 100\%=\boxed{58\%}$$

Suppose an orbiting body is bound by the gravitational attraction of a gravitating body that is a distance R away. If it is bound, it must be that it possesses insufficient kinetic energy such that when it is added to the (negative) gravitational potential the total energy makes it to zero. We define *escape velocity* as the minimum speed that an orbiting body must have in order to (barely) go infinitely far away from a gravitational source. Setting the total energy equal to zero, for the case of the object moving at escape velocity, we have:

$$\frac{1}{2}mv_{escape}^2+U\,(R)=0\quad\Rightarrow\quad v_{escape}=\sqrt{-\frac{2U\,(R)}{m}}=\sqrt{\frac{2GM}{R}}\tag{7.3.12}$$

Notice that whether or not an object can escape a gravitational attraction doesn't depend upon the would-be escaper's *mass*, but only its *velocity*. This makes for an interesting discussion when it comes to light. Light has no mass, so we would think that Newton's law of gravitation would indicate that it is unaffected by gravity. But when it comes to escape velocity, the mass does not factor in at all, so is the conclusion that light is unaffected by gravity incorrect?

It turns out that in fact light *is* affected by gravity (and specifically how this happens requires Einstein's improved theory of gravity), but what is more, we can compute whether light can escape a gravitating body. The speed of light is a well-defined constant, so the question of whether light can escape boils down to how close the origination of the light is to the gravitating body, and of course the mass of that body. The distance from which light will not escape is known as the *Schwarzschild radius*, and is found by plugging the speed of light (typically designated as c) into the escape velocity formula:

$$c=\sqrt{\frac{2GM}{R}}\quad\Rightarrow\quad R_S=\frac{2GM}{c^2}\tag{7.3.13}$$

The Earth's mass is $5.97\times10^{24}kg$, and the speed of light is $3.0\times10^{8}\,\frac{m}{s}$. Its Schwarzschild radius therefore comes out to equal about 8.8 millimeters. Okay, that is quite small, and since the distance is measured from the center of the Earth, it is clear that we will not witness the phenomenon of light being trapped by the Earth's gravity. For light to be trapped by gravity, the gravitating body must fall *inside* the Schwarzschild radius. This requires that the gravitating body be incredibly dense, and since the only force available to pull the matter into that kind of density is the same gravity force, generally a very large amount of mass is required. Assuming this occurs, the object thus created is called a *black hole*. This is a nicely descriptive name, as "black" indicates that light does not escape its gravitational influence. Explaining how appropriate the word "hole" is in this name requires some knowledge of Einstein's theory, which is unfortunately beyond the scope of this work.

CHAPTER OVERVIEW

8: Small Oscillations

8.1: Simple Harmonic Motion

Equation of Motion for an Elastic Force

We have discussed the idea of a restoring force a few times already. If such a force counteracts displacements in both directions of one-dimensional motion, then it can cause the object to move back-and forth across the equilibrium point: An object subject to a restoring force is displaced from its equilibrium point and released. It accelerates toward the equilibrium point thanks to the restoring force. Upon arrival at the equilibrium point, it doesn't stop, because the restoring force is zero there. As it continues past, the restoring force acts to slow down and eventually stop the object, whereupon the object accelerates back toward the equilibrium point and the motion repeats in the opposite direction. This is called *oscillatory motion*, and it results from all two-way one-dimensional restoring forces.

The most common sort of restoring force we study is the elastic force. Indeed, other restoring forces occurring in nature (such as those between particles exhibiting a Lennard-Jones potential energy, as discussed at the end of Section 3.7) are often modeled as masses on springs. The oscillatory motion induced by the elastic restoring force is quite special, as we will see, and is called *simple harmonic motion*. We seek here the equation that relates the position of the mass as a function of time (with the equilibrium point being the origin), usually referred to as the *equation of motion* for this force.

Start (naturally) with Newton's second law, where the net force is simply that of a spring (Hooke's law). As we are working in one dimension, we once again have the luxury of treating our vector directions as simply (+) or (–):

$$F_{net}=-kx=ma \quad \Rightarrow \quad \frac{d^2x}{dt^2}+\frac{k}{m}x=0 \tag{8.1.1}$$

We seek to determine the function $x\left(t\right)$ that satisfies this differential equation. This is actually simpler than it might at first appear, if thought about in the following way: First, let's imagine that the ratio $\frac{k}{m}$ is just the number 1. Can we think of a function that after two derivatives becomes the negative of itself? We don't know a whole lot of special functions, but amazingly, there are actually a couple that do satisfy this: sine and cosine. The derivative of sine is cosine, and then a second derivative brings it back to negative sine.

There are lots of different features we can include with a sine (or cosine) function, so let's write one out in all its glory:

$$x\left(t\right)=A\sin(\omega t+\phi) \tag{8.1.2}$$

Two derivatives of this function gives:

$$\frac{d^2}{dt^2}\left[A\sin(\omega t+\phi)\right]=\frac{d}{dt}\left[\omega A\cos(\omega t+\phi)\right]=-\omega^2A\sin(\omega t+\phi) \tag{8.1.3}$$

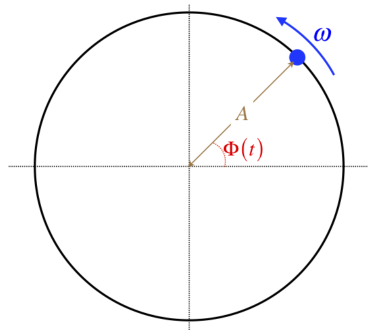
Plugging this into the differential equation gives a solution if we have:

$$\omega=\sqrt{\frac{k}{m}} \tag{8.1.4}$$

Total Phase

It may seem crazy at this point to introduce the greek letter ω as a constant here when we so recently used it as a measure of angular velocity, and there is no rotational motion going on here. But there is a good reason to do this. Consider a bead moving at a constant speed on a circular loop of wire of radius A.

Figure 8.1.1 – Bead Moving on a Circular Loop of Wire



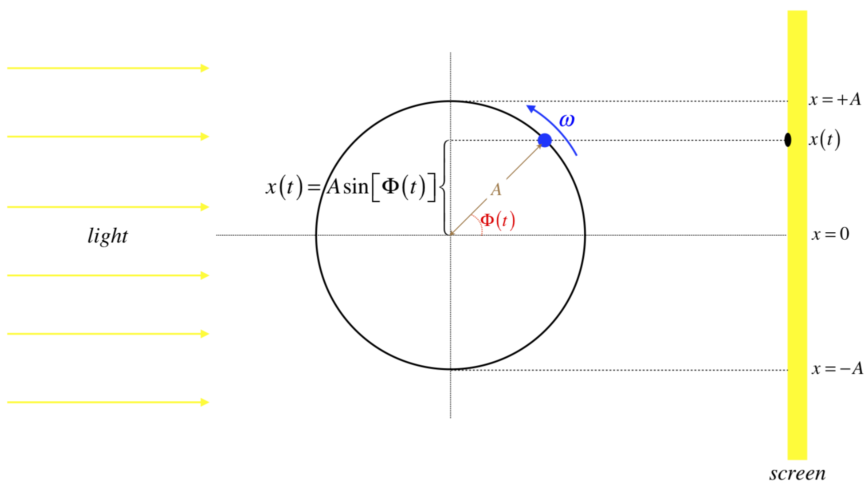
The equation describing its motion is something we are quite familiar with:

$$\Phi(t) = \omega t + \phi \tag{8.1.5}$$

We have changed notation a bit here from what we used previously for circular or rotational motion: $\Phi(t)$ represents the total angle traversed, in place of our previous $\theta(t)$, and ϕ represents the starting angle at $t = 0$, in place of our previous θ_o.

Next let's imagine placing a light source to the left of this loop, and a screen (the plane of which is perpendicular to the plane of the loop) to the right of the loop. The light would project a shadow of the bead onto the screen, and the motion of this shadow can be described mathematically:

Figure 8.1.2 – Motion of a Projection of a Bead moving on a Circular Loop



The motion of the shadow is simply a component of the motion of the bead – if we know the angle the bead makes on the circule, we know the height of the shadow on the screen. In terms of the starting angle and angular velocity of the bead, the motion of the shadow can be written explicitly as:

$$x(t) = A\sin(\omega t + \phi) \tag{8.1.6}$$

This is precisely the same as the equation of motion of a mass on a spring, Equation 8.1.2. That is, if we placed a mass on a spring at the screen with the equilibrium position at $x = 0$, pulled the spring to a maximum stretch (or pushed it to a maximum compression) of $x = A$, and then waited for the shadow to land on the mass before releasing it, the shadow would remain on the mass as it moves, *if* the angular velocity of the bead happens to equal $\sqrt{k/m}$.

We can now discard our bead-on-circular-loop model, but we keep the mathematical structure it leaves behind. The argument of the sine function $\Phi(t)$ is called the *total phase* of the harmonic motion of the mass-on-spring. The maximum expansion/compression of the spring A is called the *amplitude* of the harmonic motion. The constant ω is referred to as the *angular frequency* (not angular velocity – we've left the bead model behind), and still is expressed in units of radians per second. This constant sometimes gives way to a frequency f that is measured in cycles per second (or *hertz*), with the translation between the two being:

$$\omega = 2\pi f \tag{8.1.7}$$

The *period of oscillation* is the time it takes the system to come all the way back to where it started, and as the time per cycle, it is the inverse of the frequency:

$$T=\frac{1}{f}=\frac{2\pi}{\omega}=2\pi\sqrt{\frac{m}{k}} \tag{8.1.8}$$

Alert

The period is the time required for the system to complete a full cycle, which is not the same as the time required for the mass to return to a previous position. The mass must return to the same position and it must be ***moving in the same direction****. In other words, the total phase Φ must change by 2π.*

And finally, the constant ϕ is called the *phase constant*, and it carries the information of where the mass is at time $t=0$.

Kinematics of Harmonic Motion

Once we have a formula for the position of an object following simple harmonic motion, we can use the usual calculus tools to determine the velocity and acceleration at various times as well. The velocity as a function of time is:

$$v\left(t\right)=\frac{d}{dt}x\left(t\right)=\frac{d}{dt}\left[A\sin(\omega t+\phi)\right]=A\omega\cos(\omega t+\phi) \tag{8.1.9}$$

We note a couple of features of this result. First, since the cosine function never exceeds 1, we have the maximum speed of the object:

$$v_{max}=A\omega=A\sqrt{\frac{k}{m}} \tag{8.1.10}$$

And second, this maximum speed is achieved at $x=0$ (the equilibrium point), which makes sense, since the spring was accelerating it toward that point, and immediately after passing it, the spring starts slowing it down.

The acceleration of the mass as a function of time we get from another derivative:

$$a\left(t\right)=\frac{d}{dt}v\left(t\right)=\frac{d}{dt}\left[A\omega\cos(\omega t+\phi)\right]=-A\omega^2\sin(\omega t+\phi) \tag{8.1.11}$$

The fact that this result is only different from the position function by a factor of $\left(-\frac{k}{m}\right)$ brings us back to what started all of this, Equation 8.1.1.

Mechanical Energy

We already know that the elastic force is conservative, so mechanical energy is conserved during simple harmonic motion. At any given time during the motion, the mass will have kinetic and potential energy, with its total energy remaining constant. It's easy to write an expression for the total energy in this system by choosing a convenient point in the motion – when the mass is stationary. This occurs when it reaches its maximum separation from the equilibrium point, i.e. when the displacement equals the amplitude:

$$E_{tot}=\frac{1}{2}kx_{max}^2=\frac{1}{2}kA^2 \tag{8.1.12}$$

We can double-check this result by looking at the moment in time when there is zero potential energy and all of the mechanical energy is kinetic – at the equilibrium point. Using Equation 8.1.9, we get:

$$E_{tot}=\frac{1}{2}mv_{max}^2=\frac{1}{2}m\left(A\sqrt{\frac{k}{m}}\right)^2=\frac{1}{2}kA^2 \tag{8.1.13}$$

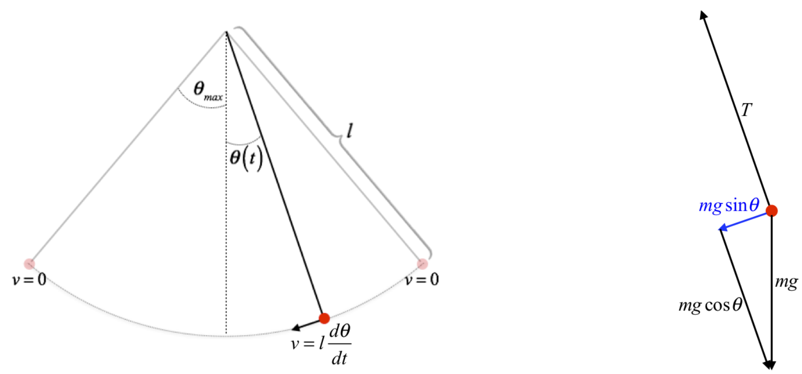
8.2: Other Restoring Forces

Pendulums

A mass on a spring is not the only physical system that exhibits simple harmonic motion. Another example is – at least to a good approximation for small amplitudes – *pendulums*. To say that a pendulum has a restoring force is imprecise – a pendulum is characterized by angular motion, and therefore it is affected by a restoring *torque*. The first type of pendulum we will consider is the *simple pendulum*. This is exactly as it sounds – it consists of a point mass under the influence of gravity at the end of a massless string which is attached to a fixed point.

It's clear that if one defines the motion of the simple pendulum in terms of angular position, the motion is oscillatory – gravity keeps producing a torque that seeks to restore vertical alignment. But is it simple harmonic motion? We need to do the analysis to figure it out. Figure 8.2.1 gives a diagram with lots of labeling, along with a free-body diagram.

Figure 8.2.1 – The Simple Pendulum



We wish to describe the motion of the pendulum, which means finding the function $\theta(t)$. We do this using Newton's second law, as we did with the mass-on-spring. We can use either the linear or the rotational form of Newton's second law – naturally both lead to the same result. Let's use the rotational version, as we will need to do later when the pendulum is not "simple." Choosing counterclockwise as the positive direction (so the pendulum to the right of the vertical is in the positive region), we see that the torque for the diagram above is negative – the restoring force has the opposite sign of the displacement, as it must.

$$\left.\begin{array}{l}\tau = -rF_\perp = -l\left(mg\sin\theta\right) \\ \tau = I\alpha = \left(ml^2\right)\alpha = \left(ml^2\right)\dfrac{d^2\theta}{dt^2}\end{array}\right\} \quad \Rightarrow \quad \frac{d^2\theta}{dt^2} + \frac{g}{l}\sin\theta = 0 \tag{8.2.1}$$

This is not the same as the differential equation for the mass-on-spring, given in Equation 8.1.1. However, if we assume the pendulum exhibits a *small amplitude*, then the value of $\sin\theta$ is very close to the value of θ, when measured in radians. For example, for a 30^o angle, the sine is 0.5000, and in radians this angle is $\frac{\pi}{6}$ = 0.5236, for a deviation of less than 5%.

> Alert
>
> *If you are not comfortable with this $\sin\theta \approx \theta$ approximation, start getting used to it – it is used **all the time** in physics!*

Applying this approximation, we get the differential equation:

$$\frac{d^2\theta}{dt^2} + \frac{g}{l}\theta = 0 \tag{8.2.2}$$

This does match the differential equation we found for the mass-on-spring, so the solution is the same (simple harmonic motion):

$$\theta(t) = \theta_{max}\sin(\omega t + \phi), \qquad where \;\; \omega = \sqrt{\frac{g}{l}} \tag{8.2.3}$$

> Alert
>
> *Note that the ω used in the argument of the sine function above is the angular frequency of the motion, and is a constant value. It is **not** the angular velocity of the pendulum $\frac{d\theta}{dt}$, which is constantly changing.*

All of the same things that we followed with for the mass-on-spring follows here, such as the velocity and acceleration functions and their maximum values, as well as the energy stored in the system. This last item bears some examination before we move on. The energy in the system is the gravitational potential energy stored at the maximum angle (measured relative to the bottom of the swing). Doing some geometry, we can get the height of the mass above the bottom of the swing, and from it the total energy:

$$y = l\left(1-\cos\theta_{max}\right) \quad \Rightarrow \quad E_{tot} = U_{max} = mgy = mgl\left(1-\cos\theta_{max}\right) \tag{8.2.4}$$

Okay, now we get to use the approximation of the cosine function for small angles:

$$\cos\theta \approx 1 - \frac{1}{2}\theta^2 \tag{8.2.5}$$

Plugging this in above gives:

$$E_{tot} = \frac{1}{2}mgl\theta^2_{max} \tag{8.2.6}$$

The reader will note that this bears a resemblance to Equation 8.1.12, which also indicates that the total energy in the system is proportional to the square of the amplitude.

The leap to other pendulums (those that are not "simple") is not a difficult one to take, if the restoring torque is also based on gravity (i.e. the pendulum swings). All this requires is replacing the point mass's rotational inertia of ml^2 with whatever the pendulum's rotational inertia around the fixed point happens to be, and computing the restoring torque based on wherever the center of mass happens to be. Calling the rotational inertia I and the distance of the center of mass from the fixed point d, and following the same procedure as above, we get for the differential equation:

$$\frac{d^2\theta}{dt^2} + \frac{mgd}{I}\theta = 0 \tag{8.2.7}$$

The angular frequency is still the square root of the coefficient of θ, so the period of oscillation for this more general pendulum is:

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{I}{mgd}} \tag{8.2.8}$$

Example 8.2.1

A thin, uniform rod with a length of $40cm$ is suspended vertically from one end, from which it is free to rotate without friction. If it is rotated a small angle from the vertical and released from rest, find how long it takes to reach a vertical orientation.

Solution

The rod is displaced a small angle, so we can treat it as a pendulum. When it starts from rest, it starts at its maximum angular displacement, and are asked to find the time it takes to get to its equilibrium point. This constitutes exactly one fourth of a cycle (it takes the same time to swing up to its maximum angular displacement from the equilibrium point), so all we need to compute is one fourth of a period. The rotational inertia is that of a rod about its end, and the center of mass is at the middle of the rod, so:

$$\frac{T}{4} = \frac{2\pi}{4}\sqrt{\frac{\frac{1}{3}ml^2}{mg\left(\frac{1}{2}l\right)}} = \pi\sqrt{\frac{l}{6g}} = \boxed{0.26s}$$

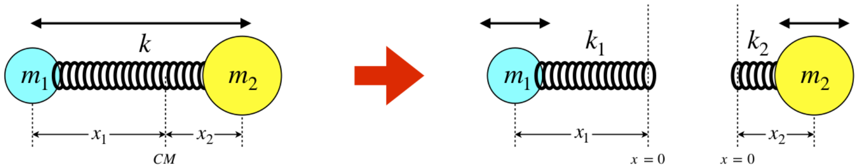
Potential Wells

At the end of Section 3.7, we looked at how we can model chemical bonds as springs. A program was outlined there which provided a way to derive an effective spring constant for any potential with a local minimum. It turns out that the "natural" vibration frequencies of these bonds are quite important when it comes to things like wavelengths of light that the material will absorb or emit, so being able to derive an effective spring constant from the potential function gives us a lot of information about how the material will behave.

There is one complication that arises with these bonds-as-springs models, however. The two molecules attached by the spring are *both* moving. How do we even fit this into our model where a single mass is oscillating through an equilibrium point? It turns out that there is a nice trick we can use for all two-body problems like this by considering the center of mass frame. [*Note: This trick is also used for gravitation when the orbiting body and gravitating body have comparable masses.*]

Consider a spring with masses at both ends. There is no net external force on the system, so as they vibrate, the center of mass remains at rest (we are assuming it started off at rest). We can therefore break this into two separate mass-on-spring systems, with the center of mass being a fixed point for each of them.

Figure 8.2.2 – Two Masses Connected By a Spring



There are a number of things we can say about this model. First, there is a relationship between the variables x_1 and x_2 and the amount each of these changes. They are both measured from the center of mass in the left diagram, and are both positive values in the right diagram, giving:

$$m_1x_1 = m_2x_2 \quad \Rightarrow \quad m_1\Delta x_1 = m_2\Delta x_2 \quad \Rightarrow \quad \Delta x_2 = \frac{m_1}{m_2}\Delta x_1 \tag{8.2.9}$$

The amount of force exerted on each mass by the spring is the same at every moment (Newton's third law), and the magnitude of this force is determined by the stretch (or compression) of the full spring according to Hooke's law. The stretch/compression of the full spring is equal to the sum of the stretches/compressions of the two springs in the separated diagram, so:

$$F = k\Delta x = k\left(\Delta x_1 + \Delta x_2\right) = k\Delta x_1\left(1 + \frac{m_1}{m_2}\right) \tag{8.2.10}$$

But looking at this force from the perspective of just m_1 in the right diagram, the force exerted on it is due to its own spring and its diplacement. Making this comparison gives us k_1 in terms of k:

$$F = k\Delta x_1\left(1 + \frac{m_1}{m_2}\right) = k_1\Delta x_1 \quad \Rightarrow \quad k_1 = \left(\frac{m_1 + m_2}{m_2}\right)k \tag{8.2.11}$$

The angular frequency of oscillation for m_1 is determined by its mass and the spring constant of the elastic force acting on it:

$$\omega = \sqrt{\frac{k_1}{m_1}} = \sqrt{\frac{k}{\mu}}, \qquad where: \quad \mu \equiv \frac{m_1m_2}{m_1 + m_2} \tag{8.2.12}$$

The motion of m_2 mirrors that of m_1, except with a different (smaller) amplitude. We know this because the two masses have to reach their maximum and minimum displacements at the same time to keep the center of mass stationary. So the angular frequency of oscillation for m_2 should come out to be the same as it is for m_1, and sure enough, it does (repeat all of the steps above with the subscripts 1 and 2 reversed).

The quantity μ has units of mass, and is commonly referred to as the *reduced mass* of the system. Its use is a common shortcut for reducing two-body problems to one-body problems. The angular frequency here takes on the usual form for a one-dimensional simple harmonic oscillator, and all that needs to be done is to calculate the reduced mass from the two masses involved and use the full spring constant (possibly computed from a potential function with a local minimum).

8.3: Damping and Resonance

Damping

If an oscillating system experiences a non-conservative force, then naturally some of its mechanical energy is converted to thermal energy. Since the energy in an oscillating system is proportional to the square of the amplitude, this loss of mechanical energy will manifest itself as a decaying amplitude. A common damping force to account for is one for which the force is proportional to the velocity of the oscillating mass, and in the opposite direction of its motion (naturally – it has to do negative work to take out mechanical energy). Air resistance (and other fluid drag) behaves like this when an object moves through the medium at low relative speeds. Let's therefore add the following force into the mix

$$\overrightarrow{F}_{damping} = -\beta\overrightarrow{v} \tag{8.3.1}$$

Treating this is one-dimension (so we can drop the vector signs) and writing the velocity as the first derivative of position, Newton's second law gives (once again measuring position from the equilibrium point of the spring):

$$F_{net} = ma \quad \Rightarrow \quad -kx - \beta\frac{dx}{dt} = m\frac{d^2x}{dt^2} \tag{8.3.2}$$

Rearranging terms changes our differential equation from Equation 8.1.1 to:

$$\frac{d^2x}{dt^2} + \frac{\beta}{m}\frac{dx}{dt} + \frac{k}{m}x = 0 \tag{8.3.3}$$

It is beyond the scope of this work to discuss how such differential equations are solved, but the solution will be given, and the reader is encouraged to plug the solution back into the differential equation to confirm that it works (actually, guessing-and-confirming is pretty much how such differential equations are solved!):

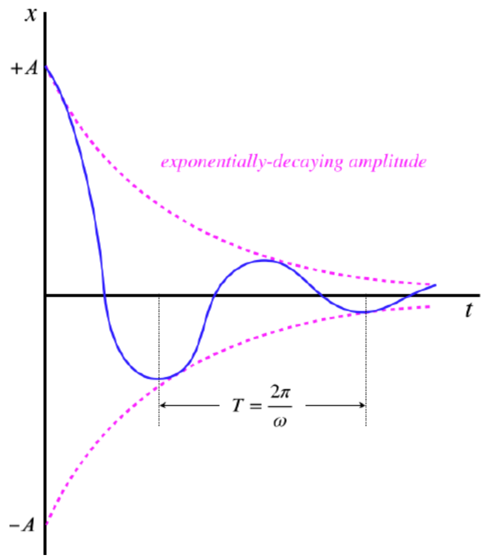
$$x\left(t\right) = Ae^{-\frac{\beta}{2m}t}\sin(\omega t + \phi), \qquad where: \;\; \omega \equiv \sqrt{\frac{k}{m} - \frac{\beta^2}{4m^2}} \tag{8.3.4}$$

We see that the introduction of the damping force affects the angular frequency ω – it is different from the solution for the undamped case, Equation 8.1.4. The fact that we can independently change the quantities that appear in the square-root provides three interesting possibilities for ω, depending upon whether the quantity in the radical is positive, zero or negative. Let's look at each in turn...

Case 1: $\boldsymbol{\beta < 2\sqrt{km}}$

Small values of β correspond to small drag effects, and don't affect the motion of the system enough to keep it from being oscillatory. In such cases, the motion is called *underdamped*. The effect of the drag in this case is twofold: It reduces the frequency of oscillation, and (as evidenced by the decaying exponential factor that includes a β in the exponent) it causes the amplitude to grow smaller with every oscillation. A graph of position vs. time looks something like this:

Figure 8.3.1 – Underdamped Motion



[*Note: This graph starts at $t = 0$ with $x = +A$ in order to simplify the picture of the exponential envelope. This means we have started the motion at rest at its maximum separation, which corresponds to a phase angle of $\phi = \frac{\pi}{2}$.*]

The effect on the energy of the system is obvious – the non-conservative drag force converts mechanical energy in the system into thermal energy, which is manifested as ever-decreasing amplitude (recall the simple relationship total energy has to amplitude, shown in Equation 8.1.12). Notice that if we gradually start increasing the value of β, the period T gets longer, stretching out the spacing of the bumps on the graph (and moving the point where the graph first crosses the t-axis to the right), while at the same time making the exponential decay (the pink curve) steeper. It should not be surprising therefore what we find in the next case...

Case 2: $\beta = 2\sqrt{km}$

In this case, the sinusoidal behavior goes away. One way to see this is to note that this condition causes ω to vanish, making the period infinite, thereby making it so that the system never completes an oscillation. This kind of motion is called *critically-damped*. The easiest way to get a handle on this is to simply plug the condition into the solution, Equation 8.3.4:

$$x(t) = [A \sin\phi]\, e^{-\sqrt{\frac{k}{m}}t} \tag{8.3.5}$$

We see that the oscillatory motion is gone (the sine function just includes the phase constant, so there is no time dependence in the sine function. In fact, if we start the motion at rest at the maximum spring stretch ($\phi = \frac{\pi}{2}$, the value of $x(t)$ never even goes negative – it just exponentially decays, and is simply one of the dashed pink graphs shown in Figure 8.3.1.

Case 3: $\beta > 2\sqrt{km}$

This is strange case is called *overdamped*. The value under the radical is now negative, which makes the angular frequency *imaginary*. Dealing with the complex numbers is a bit cumbersome, but fortunately we don't have to do this. We can simply go back to the original differential equation armed with the knowledge about β and obtain a new solution from scratch. Again, we are not in a position to delve into the details of solving differential equations, but this case also results in no oscillations, and exponential decay of the displacement toward the equilibrium point without ever crossing it. As one might expect, with a stronger force opposing the motion than the critically-damped case, the system heads toward the equilibrium point more slowly than that case.

Resonance

If we can take energy out of the system with a damping force that acts in opposition to the motion, it makes sense that we can also add energy into the system by introducing a force in the direction of motion. While a damping force by something like drag through a fluid works "automatically," adding energy to the system requires some coordination in the application of the force. For example, if the force can only act in the $+x$ direction, the force can only be applied periodically – exerting it all the time would have it acting in the direction opposite to its motion half the time.

Generally when we talk about one of these energy-adding forces, we assume they are periodic (we call this a *periodic driving force*). The simplest such force is sinusoidal:

$$F(t) = F_o \sin \omega_d t \tag{8.3.6}$$

The quantity F_o is the maximum force applied, and ω_d is the *driving frequency*. If we add this to the equation for Newton's second law (including damping), we get:

$$F_{net} = ma \quad \Rightarrow \quad -kx - \beta \frac{dx}{dt} + F_o \sin \omega_d t = m \frac{d^2x}{dt^2} \tag{8.3.7}$$

These differential equations just keep getting uglier! We won't go into the solution for this one, but a look at the physical effects from a perspective of energy is enlightening.

The first thing we note is that this force, over time, adds energy to the system, which means that while the damping force takes energy away, the total energy doesn't decay all the way to zero. This means that an amplitude of oscillation doesn't dissipate away. Second, it seems clear that the more time that the force spends pushing in the correct direction, the more energy it can add to the system. It will maximize its pushing in the correct direction when it is completely synchronized with the *natural or resonant frequency* $\omega_o \equiv \sqrt{\frac{k}{m}}$ of the system, which is the frequency at which it would oscillate without damping. To see this, we'll take a peek at the result of solving the differential equation. This solution gives the following expression for the amplitude resulting from forced, damped oscillatory motion:

$$A = \frac{F_o}{\sqrt{m^2 \left(\omega_d^2 - \omega_o^2\right)^2 + \beta^2 \omega_d^2}} \tag{8.3.8}$$

We can see from this expression that the closer the driving frequency ω_d is to the natural frequency ω_o, the larger the amplitude is. Given that the amplitude is a proxy for the energy in the system, this means that more energy is added to the system by a driving force whose frequency is well-tuned to the natural frequency of the system. This phenomenon is called *resonance*.

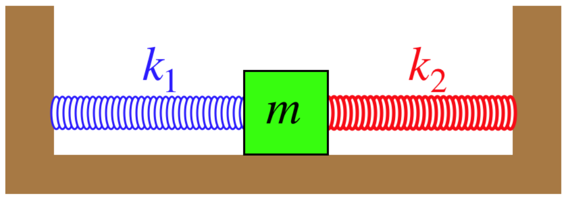
8.4: Coupled Oscillators and Normal Modes

Multiple Springs

As the next step in our examination of small oscillation theory, we move from a single oscillator into combinations of oscillators. A very basic first step is to consider the effect of multiple springs on a single mass. We will keep the discussion basic by continuing to insist that the mass only moves in one dimension.

As a first case, consider the simple case of a mass attached to two different springs. We will call this case *parallel* springs, because each spring acts on its own on the mass without regard to the other spring.

Figure 8.4.1 – Two Springs in "Parallel"



We will assume for simplicity that the mass is attached between the two springs when both are at equilibrium. This is not a necessary assumption, because if they are not at equilibrium, a new equilibrium point (where both springs are slightly stretched or slightly compressed) exists, not changing any of the analysis. Note that if we had placed both springs on the same side of the mass, then the physics would not be different from what we have here – instead of two forces in the same direction because one spring is compressed and the other expanded, the two forces would be in the same direction because both are stretched or expanded, and the magnitudes of the Hooke's law forces are the same in both directions.

This system behaves exactly like a single-spring harmonic oscillator, but with what frequency? To answer this, we basically need to find the single spring constant that is equivalent to these two springs. We can do this by displacing the mass a distance Δx and seeing what restoring force is the result for each case. As stated above, the forces by both springs are in the same direction, so:

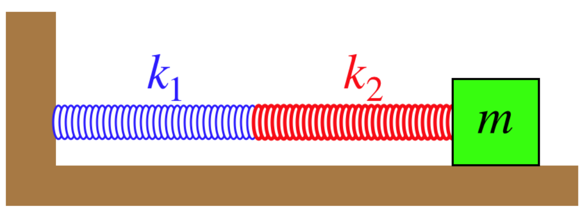
$$\begin{array}{ll}\text{two springs:} & |F| = k_1\Delta x + k_2\Delta x \\ \text{one "equivalent" spring:} & |F| = k_{eq}\Delta x\end{array} \tag{8.4.1}$$

In order to be equivalent, these restoring forces must be equal, so we get a way of writing these two springs as a single equivalent spring:

$$k_{eq} = k_1 + k_2 \tag{8.4.2}$$

Simple enough, but now let's look at what happens if we connect the springs to each other. This configuration we call *series*, because the springs are in direct contact and therefore effect the mass by successive stretching. Note that in the parallel cases, the stretching was coincident rather than successive, and the resulting physical effect is different in the two cases.

Figure 8.4.2 – Two Springs in "Series"



In this case, the springs are connected, which means that according to Newton's third law, whatever force is stretching or compressing one spring is also the force that is stretching or compressing the other, and this is also the force exerted on the mass. They have different spring constants, so the equal forces will stretch or compress them different amounts. If we are to combine them into a single equivalent spring, then the sum of the stretches or compressions will be the total displacement of the mass. We therefore have:

$$\begin{array}{ll}\text{two springs:} & |F| = k_1\Delta x_1 = k_2\Delta x_2 \\ \text{one "equivalent" spring:} & |F| = k_{eq}\left(\Delta x_1 + \Delta x_2\right)\end{array} \tag{8.4.3}$$

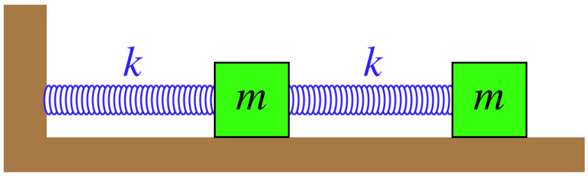
Solving for Δx_1 and Δx_2 in terms of the force in the "two springs" equations and plugging them into the equivalent spring equation gives:

$$\frac{1}{k_{eq}}=\frac{1}{k_1}+\frac{1}{k_2} \tag{8.4.4}$$

Coupled Oscillators

Our next step is to increase the number of masses. We already considered the case of two masses connected by a single spring in Section 8.2, but found that case to just be equivalent to one "reduced mass" on a single spring. Here we will introduce a second spring as well, which removes this simplification, and creates what is called *coupled oscillators*. Let's start with the simplest conceivable case – two identical masses connected with two identical springs to a single fixed point:

Figure 8.4.3 – Simple Coupled Oscillators



What makes this so much more complicated than the previous cases is that the "origin" for the second mass is moving, which complicates the measurement of the stretch or compression of the second spring. We begin by defining position variables for both blocks, and then use these variables to relate the forces on each. We actually have a great deal of freedom in defining how we measure the positions of the two blocks, and there is no reason to define a common origin for them, as long as we incorporate these variables into the dynamical equations correctly.

For this problem, we will define the origin of the left mass as its position when the left spring is at its equilibrium length, and the origin of the right mass as its position when both springs are at their equilibrium lengths. So if both blocks are at their origins, then there is no force on either block.

Figure 8.4.4 – Coordinates for Blocks

Now we need to make an accounting of the forces on each block according to the coordinates defined. The force on the right block is easy – it is determined only by the stretch or compression of the spring between the blocks. That stretch/compression is easily expressed in terms of the coordinates defined:

$$F_2=-k\left(x_2-x_1\right) \tag{8.4.5}$$

Note that if both blocks move the same distance in the same direction from their origins, then $x_1=x_2$, and the force is zero because the spring between them isn't stretched. A bit of thought should also convince the reader that the sign on the force works out correctly for various positions for the blocks.

The force on the left block is only a little bit trickier, as there are two springs acting on it. The force on it by the spring between the blocks is the same as what we just found, but in the opposite direction, and the force on the left block by the left spring just follows the simple Hooke's law. The net result is therefore:

$$F_1=-kx_1+k\left(x_2-x_1\right)=-k\left(2x_1-x_2\right) \tag{8.4.6}$$

These two forces result in two differential equations when Newton's second law is applied in each case. Specifically, we have:

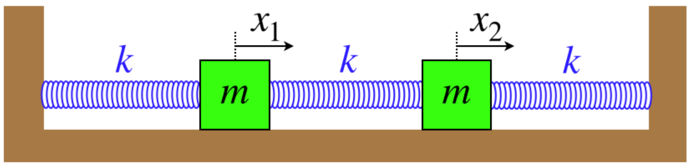
$$\begin{array}{llll} \text{left block:} & F_1=m\dfrac{d^2x_1}{dt^2}=-2kx_1+kx_2 & \Rightarrow & \dfrac{d^2x_1}{dt^2}+2\dfrac{k}{m}x_1-\dfrac{k}{m}x_2=0 \\ \text{right block:} & F_2=m\dfrac{d^2x_2}{dt^2}=kx_1-kx_2 & \Rightarrow & \dfrac{d^2x_2}{dt^2}+\dfrac{k}{m}x_2-\dfrac{k}{m}x_1=0 \end{array} \tag{8.4.7}$$

These coupled differential equations appear quite daunting to solve, but there is a trick for treating these systems, which we will examine next.

Changing Variables

It turns out that the system above for which we obtained coupled differential equations is a bit "ugly" inasmuch as the solution is not especially illuminating, so here we will examine a system with a bit more symmetry by adding another spring. To save space, we will include the two coordinate systems with the diagram.

Figure 8.4.5 – Coupled Oscillators With Three Springs



As before, we will define the origins of the two coordinate systems to be the equilibrium points, and as before, we need to express the net forces on each mass according to the positions. Clearly the equation for the force on the left mass is identical to the example above (Equation 8.4.6). We can work out the force for the second mass as we did before, but because of the symmetry of the situation, the equation for the net force on the right mass should look similar to that of the left mass:

$$\begin{array}{llll} \text{left block:} & F_1=-kx_1+k\left(x_2-x_1\right)=-k\left(2x_1-x_2\right) & \Rightarrow & \dfrac{d^2x_1}{dt^2}+2\dfrac{k}{m}x_1-\dfrac{k}{m}x_2=0 \\ \text{right block:} & F_2=-kx_2+k\left(x_1-x_2\right)=-k\left(2x_2-x_1\right) & \Rightarrow & \dfrac{d^2x_2}{dt^2}+2\dfrac{k}{m}x_2-\dfrac{k}{m}x_1=0 \end{array} \tag{8.4.8}$$

The symmetry in this physical system is reflected in these equations of motion, but we are no closer to a solution than before. Here is where our "trick" comes in. Add the two differential equations together to obtain a new equation:

$$\frac{d^2x_1}{dt^2}+\frac{d^2x_2}{dt^2}+2\frac{k}{m}x_1+2\frac{k}{m}x_2-\frac{k}{m}x_1-\frac{k}{m}x_2=0 \quad\Rightarrow\quad \frac{d^2\left(x_1+x_2\right)}{dt^2}+\frac{k}{m}\left(x_1+x_2\right)=0 \tag{8.4.9}$$

If we define a new variable $q_1\left(t\right)$ (the reason for the subscript will soon become clear), a remarkable thing happens:

$$q_1\left(t\right)\equiv x_1\left(t\right)+x_2\left(t\right) \quad\Rightarrow\quad \frac{d^2q_1}{dt^2}+\frac{k}{m}q_1=0 \tag{8.4.10}$$

This new variable satisfies a differential equation that exactly describes harmonic motion with angular frequency equal to $\sqrt{\frac{k}{m}}$! It's not clear how to physically interpret this at the moment, so we will come back to it shortly, but at least we can write a solution for $q_1\left(t\right)$:

$$q_1\left(t\right)=A_1\sin\left(\omega_1t+\phi_1\right)\,, \qquad \omega_1=\sqrt{\frac{k}{m}} \tag{8.4.11}$$

This is not the only function we can find that has this property. Let's now *subtract* the two differential equations that we previously added:

$$\frac{d^2x_1}{dt^2}-\frac{d^2x_2}{dt^2}+2\frac{k}{m}x_1-2\frac{k}{m}x_2-\frac{k}{m}x_1+\frac{k}{m}x_2=0 \quad\Rightarrow\quad \frac{d^2\left(x_1-x_2\right)}{dt^2}+3\frac{k}{m}\left(x_1-x_2\right)=0 \tag{8.4.12}$$

Again we define a new function $q_2\left(t\right)$ and get a result only slightly different from the previous one – the frequency of oscillation comes out to be greater by a factor of $\sqrt{3}$:

$$\begin{array}{ll} q_2\left(t\right)\equiv x_1\left(t\right)-x_2\left(t\right) \quad\Rightarrow & \dfrac{d^2q_2}{dt^2}+3\dfrac{k}{m}q_2=0 \\ q_2\left(t\right)=A_2\sin\left(\omega_2t+\phi_2\right)\,, & \omega_2=\sqrt{3\dfrac{k}{m}} \end{array} \tag{8.4.13}$$

With these two solutions, we can now reconstruct the original functions, giving us the answer we were looking for:

$$x_1(t) = \frac{1}{2}[q_1(t) + q_2(t)]$$
$$x_2(t) = \frac{1}{2}[q_1(t) - q_2(t)] \tag{8.4.14}$$

It should be noted here that the amplitudes of the two normal modes are free parameters, so the motions of the two blocks can essentially be any mix of the two q-function sinusoids. If one of the amplitudes is zero, then the blocks follow the simple harmonic motion of the remaining q-functions. While this looks like nothing more than a math trick that only works for a highly-symmetric problem, it actually displays a very fundamental principle used throughout physics.

Normal Modes

It turns out that the oscillatory motion described by the functions q_1 and q_2 can be observed in the physical system, given the proper circumstances. Suppose we displace both blocks in the same direction by the same amount, and release them from rest. In this configuration, the spring between them is neither stretched nor compressed, and the force each block feels will be caused only by the other spring attached to them. Since they have the same mass and were displaced equal amounts with identical springs, their motions will mirror each other, which means the center spring will never expand or compress. In this case, they will both oscillate like a single mass on a single spring – with a frequency equal to $\sqrt{\frac{k}{m}}$. This is exactly the frequency we found for q_1. By displacing both blocks the same amount, we have set up the circumstances where the difference in their positions (measured by q_2) never changes. By forcing q_2 to remain constant, only the motion described by q_1 is seen.

If we only want to witness the motion described by q_2, we just need to set up an initial condition where the sum of the positions is unchanged. This corresponds to moving the blocks equal distances from their equilibrium points in *opposite* directions and releasing them from rest.

There are only two of these "special" modes of oscillation for this system, and these are called the system's *normal modes*. It turns out that these modes exist for every system, including those that are not symmetric (different masses, different spring constants, etc.). It also happens that there is a separate normal mode for every *degree of freedom* that the system possesses. By "degree of freedom" we mean essentially the number of variables we need to introduce in order to completely describe the evolution of the system. For example, if we added another block and another spring to the system above, then we would need another variable x_3 to describe the motion of the third block, giving the system three degrees of freedom, and therefore three normal modes.

One useful way to think about normal modes is like unit vectors we use to describe the "vectors" that are the total motion of the system. Each mode is "orthogonal" to all the others, and when combined in proper proportions, they can build any of the complicated types of motion. Each mode has its own individual frequency (frequencies of different normal modes are sometimes equal to each other, typically due to a symmetry of the system).

Return to the "Simple" Coupled Oscillators

Armed with this idea of normal modes, let's take another shot at the system of coupled oscillators shown in Figure 8.4.3. We have our two differential equations that include x_1 and x_2 in Equation 8.4.7. Now we are looking for a pair of new coordinates, q_1 and q_2, that express the motions of the normal modes. In the highly-symmetric case that we just solved, we found that each of these normal mode functions was a simple combination of the individual object coordinates – q_1 was the sum of x_1 and x_2, while (q_2\) was the difference. In this less-symmetric case, we need to assume that the combination is more complicated, so we'll write the following for a generic q (we'll see that the cases for the two separate modes come out automatically):

$$q(t) = \alpha \cdot x_1(t) + \beta \cdot x_2(t)\ \ , \tag{8.4.15}$$

where α and β are constants.

If we multiply the first differential equation in Equation 8.4.7 by α, the second by β, and add them together, we get:

$$\frac{d^2}{dt^2}[\alpha x_1 + \beta x_2] + (2\alpha - \beta)\frac{k}{m}x_1 + (\beta - \alpha)\frac{k}{m}x_2 = 0 \tag{8.4.16}$$

The quantity in square brackets is just q. We are looking for a normal mode equation, so ultimately this needs to look like the equation for a simple harmonic oscillator with the angular frequency associated with this mode:

$$\frac{d^2q}{dt^2} + \omega^2 q = 0 \tag{8.4.17}$$

If we compare these last three equations, we see that the second derivative is already where we want it to be. To get the rest of the differential equation where it needs to be, we need:

$$(2\alpha-\beta)\frac{k}{m}x_1+(\beta-\alpha)\frac{k}{m}x_2=\omega^2 q=\omega^2\left(\alpha x_1+\beta x_2\right) \tag{8.4.18}$$

Collecting terms gives:

$$\left[(2\alpha-\beta)\frac{k}{m}-\omega^2\alpha\right]x_1+\left[(\beta-\alpha)\frac{k}{m}-\omega^2\beta\right]x_2=0 \tag{8.4.19}$$

The two masses do not have to always be at their origins at the same time, which means that the coefficients of x_1 and x_2 must each equal zero, giving:

$$\begin{array}{lcl} (2\alpha-\beta)\dfrac{k}{m}=\omega^2\alpha & \Rightarrow & \dfrac{\beta}{\alpha}=2-\dfrac{m}{k}\omega^2 \\ (\beta-\alpha)\dfrac{k}{m}=\omega^2\beta & \Rightarrow & \dfrac{\alpha}{\beta}=1-\dfrac{m}{k}\omega^2 \end{array} \tag{8.4.20}$$

Now we can multiply these last two equations together to eliminate α and β, leaving us with a quadratic equation to solve for ω:

$$1=\left(2-\frac{m}{k}\omega^2\right)\left(1-\frac{m}{k}\omega^2\right) \quad\Rightarrow\quad \omega=\sqrt{\frac{3\pm\sqrt{5}}{2}}\sqrt{\frac{k}{m}} \;\approx\; 1.62\sqrt{\frac{k}{m}} \;\; or \;\; 0.62\sqrt{\frac{k}{m}} \tag{8.4.21}$$

We see that the quadratic equation gives us two different possible frequencies – one for each normal mode. If we plug each of these frequency solutions back into Equation 8.4.20, we can determine the ratio of α and β for each normal mode, which gives us q_1 and q_2. [*Note that we can never determine exact values for α and β – we can only find their ratios. For example, even in the symmetric case we could have chosen $q_1=0.3x_1+0.3x_2$ and $q_2=0.3x_1-0.3x_2$ and everything would have worked out the same.*]

$$\frac{\beta}{\alpha}=2-\frac{3\pm\sqrt{5}}{2}=\frac{1\mp\sqrt{5}}{2} \;\approx\; -0.62 \;\; or \;\; +1.62 \tag{8.4.22}$$

If we just choose α equal to +1, then we can put this all together to summarize the two normal modes:

$$\begin{array}{lll} \text{mode 1:} & q_1=(+1)\,x_1+(-0.62)\,x_2 & \omega_1=1.62\sqrt{\dfrac{k}{m}} \\ \text{mode 2:} & q_2=(+1)\,x_1+(+1.62)\,x_2 & \omega_2=0.62\sqrt{\dfrac{k}{m}} \end{array} \tag{8.4.23}$$

How do we interpret these modes physically? Mode 1 involves the two masses being displaced in opposite directions from their equilibrium positions at their turnaround points (i.e. they are always moving in opposite directions). Their maximum displacements are in the ratios given, so if the mass between the two springs is $1cm$ from its equilibrium when it stops, then the other mass is $0.62cm$ from its equilibrium (in the opposite direction) when it stops at the same moment. Mode 2 is interpreted similarly, except that the masses are always moving in the same direction, and a $1cm$ amplitude for the mass between the two springs results in a $1.62cm$ amplitude for the other mass. The system's frequency in mode 1 is greater than it is for mode 2, which makes sense if one visualizes the motions of the two modes – the first has the two masses going in opposite directions, not displacing that far, while the second mode has the whole system stretching far away from the wall, then compressing toward the wall.

There is an obvious symmetry between the two modes, where the ratios of the amplitudes of the masses match the ratios of the normal mode frequencies. This comes from the fact that the masses and spring constants are equal – the reader should not be fooled into thinking that this is a general feature of normal modes of coupled oscillators.

Finally, it needs to be emphasized again that these two modes are not the only ways that the system can evolve – they are just the only ways that the system will repeat itself periodically if the initial conditions are set up just right. If the masses are started in some non-normal mode condition (e.g. they are separated from their equilibrium points by amounts that do not satisfy either of the allowed proportions), then the motion will be a mix of the two modes, and will look very complicated, and will not repeat with the simplicity that the normal modes repeat.

Coming Attractions

When one obtains more mathematical tools (namely the matrix methods learned in linear algebra), the method for solving for normal modes is not so clunky as it is shown to be here. Also, a deeper understanding of these modes becomes possible. As a quick preview of things to come, try the following:

- In the equations for the q's, treat the coefficients of the x's as though they are components of a vector.
- Take the dot product of the "vectors" q_1 and q_2. You will find that it vanishes for both of the examples above, and in fact *this is a general feature of normal modes of all systems of oscillators*.

This works no matter how many degrees of freedom are present – the 5 normal modes for a system of 5 blocks connected by springs form mutually-orthogonal 5-dimensional vectors. Indeed this mutual orthogonality is the origin of the use of the word "normal" in "normal modes."

Preface

Physics and Models

The whole idea of the study of physics is to understand how the universe operates. We cannot actually ever know for sure how this works, but we play a sort of game: We develop a model that explains why things happen the way they do, and then we test the effectiveness of that model when it comes to predicting how other things will unfold. If the model predicts accurately, it is a “good” model, and if it doesn’t, it is discarded.

Inherent to this description is the idea of “accuracy.” No model we have ever designed has ever predicted a result with 100% accuracy. Mainly this is because 100% accuracy requires 100% accurate measurements of both the starting conditions and of the results, and this simply isn’t possible. What we settle for instead is a sense of what sorts of problems our model is intended to solve. Some models are precise to an incredibly small dimension (like models that predict atomic behavior), but these are not useful for making predictions in the macroscopic world where trillions of trillions of atoms are involved. Conversely, we also make macroscopic models that breakdown when our measurements become too fine.

So all models come with them an understanding that they work “up to a point.” When I discuss a problem involving a “frictionless surface,” one can certainly argue that no such thing exists, but true as that statement is, it is not relevant. The model of the frictionless surface allows us to answer questions about situations where the amount of friction is small, and our coarse measurements can’t distinguish the effects of that small amount of friction. Further, this model can be used as a starting point, to which we can later append a friction effect to make a more inclusive model.

You will sometimes hear me (or future physics instructors) say that such-and-such is true if a certain quantity is “small.” This simply means that if the quantity is small enough, the coarseness of our measurements provide too much noise for us to really notice the effect of that small quantity.

Measurement and Units

While we can make some general predictions about the behavior of our universe, these are not usually particularly satisfying. The statement, “If I drop something, it will fall to Earth” can be considered a “theory of gravity,” but big deal. How long does it take the dropped object to fall some specified distance? How fast is it going when it lands? How do the motions of two different dropped objects differ from each other? All of these are questions we would like to answer as well, and they all require measurement. But if I measure the time for an object to fall and call it “3,” while you measure the same process and call it “17,” we are not going to get anywhere. We need a standardized system of units that we can agree upon so that we can compare results.

Many hundreds of years before Galileo, Aristotle sought to explain everything, but he did so descriptively. Galileo was among the first set out to do so mathematically. Galileo was studying the effects of gravity on motion (Aristotle simply said that things that are heavy fall, and things that are light rise), and did experiments where he rolled balls down ramps and timed their journeys. He started zeroing-in on a precise mathematical description, but every time he got close to accepting his results, the experiment would go haywire and his new results would disagree badly with the early ones. It turns out that the problem was that he was using his own heartbeat to time the motion of the ball, and when his predictions started coming true, he got excited, his heart beat faster, and the predictions began to fail.

There are many systems of units available to us. We could for example measure speed (which is a rate of distance covered over time) in units of furlongs per fortnight. But there is one system that we use in physics as the default, from which we only rarely stray. It is called the *Système Internationale d’Unites*, or *SI units* for short. The three most fundamental measurements we have in this system are *meters* (distance), *kilograms* (mass), and *seconds* (time). For this reason, this system of units is also often referred to as *mks units*.

First-time physics students often pay little attention to units when they are solving problems, thinking of them as more of a nuisance than a help. You should fight this tendency. If you are solving a problem to find a speed and you end up with an answer that (because you carefully carried units through the math) came out to be kilograms per meter, then that is an indication that you made a mistake somewhere. Many students plug in the numbers and then throw the proper units in at the end, and this provides them with no opportunity for catching mistakes.

Index

Detailed Licensing

Overview

Title: UCD: Physics 9HA – Classical Mechanics

Webpages: 60

All licenses found:

- CC BY-SA 4.0: 78.3% (47 pages)
- Undeclared: 21.7% (13 pages)

By Page

- UCD: Physics 9HA – Classical Mechanics - *CC BY-SA 4.0*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Motion - *CC BY-SA 4.0*
    - 1.1: Vectors - *CC BY-SA 4.0*
    - 1.2: Vector Multiplication - *CC BY-SA 4.0*
    - 1.3: Straight-Line Motion - *CC BY-SA 4.0*
    - 1.4: Kinematics - *CC BY-SA 4.0*
    - 1.5: Graphing - *CC BY-SA 4.0*
    - 1.6: Motion in Multiple Dimensions - *CC BY-SA 4.0*
    - 1.7: Examples of 2-Dimensional Motion - *CC BY-SA 4.0*
    - 1.8: Relative Motion - *CC BY-SA 4.0*
  - 2: Force - *CC BY-SA 4.0*
    - 2.1: Properties of Force - *CC BY-SA 4.0*
    - 2.2: Effects of Force - *CC BY-SA 4.0*
    - 2.3: Types of Forces - *CC BY-SA 4.0*
    - 2.4: Problem Solving - *CC BY-SA 4.0*
  - 3: Work and Energy - *CC BY-SA 4.0*
    - 3.1: The Work - Energy Theorem - *CC BY-SA 4.0*
    - 3.2: Conservative and Non-Conservative Forces - *CC BY-SA 4.0*
    - 3.3: Mechanical Advantage and Power - *CC BY-SA 4.0*
    - 3.4: Energy Conservation Models - *CC BY-SA 4.0*
    - 3.5: Thermal Energy - *CC BY-SA 4.0*
    - 3.6: Force and Potential Energy - *CC BY-SA 4.0*
    - 3.7: Energy Diagrams - *CC BY-SA 4.0*
  - 4: Linear Momentum - *CC BY-SA 4.0*
    - 4.1: Repackaging Newton's Second Law - *CC BY-SA 4.0*
    - 4.2: Center of Mass - *CC BY-SA 4.0*
    - 4.3: Momenta of Systems - *CC BY-SA 4.0*
    - 4.4: Momentum and Energy - *CC BY-SA 4.0*
    - 4.5: Collisions - *CC BY-SA 4.0*
    - 4.6: Problem Solving - *Undeclared*
  - 5: Rotations and Rigid Bodies - *CC BY-SA 4.0*
    - 5.1: Rotational Kinematics - *CC BY-SA 4.0*
    - 5.2: Non-Inertial Frames - *CC BY-SA 4.0*
    - 5.3: Rotational Inertia - *CC BY-SA 4.0*
    - 5.4: Dynamics of Rotating Objects - *CC BY-SA 4.0*
    - 5.5: Torque - *CC BY-SA 4.0*
    - 5.6: Static Equilibrium - *CC BY-SA 4.0*
  - 6: Angular Momentum - *CC BY-SA 4.0*
    - 6.1: Linking Linear and Angular Momentum - *CC BY-SA 4.0*
    - 6.2: Effects of Torque - *CC BY-SA 4.0*
    - 6.3: Applications of Angular Momentum Conservation - *CC BY-SA 4.0*
  - 7: Gravitation - *CC BY-SA 4.0*
    - 7.1: Universal Gravitation - *CC BY-SA 4.0*
    - 7.2: Kepler's Laws - *CC BY-SA 4.0*
    - 7.3: Energy in Gravitational Systems - *CC BY-SA 4.0*
  - 8: Small Oscillations - *CC BY-SA 4.0*
    - 8.1: Simple Harmonic Motion - *CC BY-SA 4.0*
    - 8.2: Other Restoring Forces - *Undeclared*
    - 8.3: Damping and Resonance - *Undeclared*
    - 8.4: Coupled Oscillators and Normal Modes - *Undeclared*
  - Preface - *CC BY-SA 4.0*
  - Back Matter - *Undeclared*
    - Index - *Undeclared*
    - Glossary - *Undeclared*
    - Detailed Licensing - *Undeclared*