The Fundamentals of

PHYSICS

Supplemental Readings

Robert A. Cohen
Physics Department
East Stroudsburg University

October 23, 2023
Version 1.2

Contents

1. Why Physics?

1.1 Introduction

The reason we study physics is because of the way physics is able to identify and use a small set of powerful ideas to make testable predictions. These ideas consist of scientific laws and theories.

What's the difference between a scientific law and a scientific theory?

A scientific law is based upon observations of nature and does not attempt to *explain* why the relationship is the way it is. A scientific **theory** *explains* the relationships that a scientific law describes.[i]

For example, while I was growing up I noticed that the price of a stamp always equaled my age. When I was six, the price of a stamp was six cents. When I was eight, the price of a stamp was eight cents. While I was ten, the price of a stamp increased to ten cents.[ii] I did not know why this was. It just was a "law" that the stamp price seemed to follow. I had no "theory" for why this was.

The most powerful scientific laws and theories are those that have been thoroughly tested and can be applied to a large number of situations. In this way, the scientific definition of a theory differs from how the general population uses the term. Many people outside of science use the word "theory" to refer to a guess. That is *not* the way we use the term in science.[iii] Section 3 provides more information about laws and how they differ from theories.

How many scientific laws are there?

[i] This distinction is not universal but tends to be true in general.

[ii] This continued up to age 32. Alas, the price of a stamp didn't increase to 33 cents until I was 35.

[iii] A scientific **hypothesis** is more similar to a guess, in that a hypothesis is a tentative relationship and explanation. A scientific hypothesis is testable, though, so it is not simply a guess. Further complicating matters is the use of the *statistical* hypothesis, which is pretty much just a prediction, not a tentative relationship or explanation.

There can be many, many laws, but we'll focus mainly on just a couple, of which the law of force and motion is the main one.

1.2 Why study scientific laws and theories?

Scientific laws, like scientific theories, are powerful because we can use them to make a **prediction**.

In other words, rather than use a new relationship for each problem we encounter, we can focus on a small set of relationships, which can then be used for every problem.

WHY IS IT IMPORTANT TO LEARN HOW TO APPLY THESE SMALL SET OF RELATIONSHIPS, RATHER THAN JUST USE A NEW RELATIONSHIP FOR EACH PROBLEM WE ENCOUNTER?

To appreciate the value of our approach, consider the following scenario:

> Suppose you anticipate getting a job that requires you to get around a region where GPS is not available. You decide to take a class that will teach you how to do this.
>
> Two sections of the course are being offered. One section consists of a series of lectures on how to get to various locations in the region. During the first lecture, the instructor describes the route one must take to get to a particular location. All of the steps are written on the board, which the students dutifully copy and try to memorize. The next lecture, the instructor describes the route one must take to get to another particular location. This continues, with each lecture being dedicated to memorizing a route to a particular location.
>
> In the other section, the students are provided with a map of the entire region. A series of activities are utilized to familiarize the students with the map, the conventions used with the map, and how the map is best utilized to determine the appropriate route. Lectures focus on how to use the map in various situations.

Most people would choose the second section because it requires less memorization and is more powerful and universal.

With the first course, students will learn lots of routes. However, they need to memorize them, and it simply isn't possible to memorize every possible route. Even a subset is difficult to remember unless used soon after the course ends. Furthermore, if a route has changed since the course was taken, there is no way for the student to figure out an alternate route.

With the second course, on the other hand, students will learn *how* to determine the best route to any location. While they may not learn any particular route, using a map is a much more powerful technique over the long term, since it can be used for almost every possible route.

So what does the story have to do with scientific laws?

❢ Scientific laws are valuable because of their generality; they can be applied to a wide range of phenomena and circumstances.

Scientific laws, like maps, are powerful in the sense that they can be applied to a variety of situations. The more general they are, the more powerful they are. The sample we'll use here can be applied to practically *all* of the situations we'll examine.

The specific routes described in the story are analogous to problems in physics. Each problem represents a particular destination. Meanwhile, the map is analogous to the law of force and motion.

Rather than memorize how to solve each particular problem, in this course you will learn the underlying idea (represented by the law of force and motion) so you can solve a much wider range of problems.

For example, the same basic principles used to predict the motion of a falling rock can also be used to predict the energy released in a chemical reaction like combustion or even a nuclear reaction like fission.

We must start with baby steps, however, which means that for much of this volume, the examples will be very simple. Since the complications are often the things that make situations interesting, things will get more interesting in the second volume when more complicated situations are examined.

> ✍ Albert Einstein suggested that scientific laws are like tall buildings. Just as a tall building allows us to see the roofs of many other buildings, these laws allow us to solve many problems. In addition, just as a taller building can bring into view buildings that we had not seen before, these laws provide insight into more than just the problem that provided the impetus for identifying the law in the first place.

What if we only had a couple of situations to worry about?

If we were sure that we'd encounter only one specific situation then, yes, it would probably be more efficient to simply memorize how to solve problems involving that one situation. In a similar way, you don't need to learn how to read maps if all you need is to get from home to school and back (and the route never changes).

However, if we can master this set of scientific laws and definitions, we can then solve problems involving lots of situations.

✍ If you are only interested in a few specific things then you may become frustrated with the process we'll be following. Keep in mind that not everyone in the class is only interested in the same things you are interested in. Your patience is appreciated. And, along the way, you may come to appreciate the power and elegance of physics.

What use is this to me if I am not going to be a physicist?

By mastering the ability to apply a small set of ideas to a large array of problems, you'll obtain a very powerful skill set that is useful in all science areas, not just physics.

In this class, our context will be on the law of force and motion basically because it is a relatively simple relationship and can be applied to such a large array of situations. For example, it can be used not only to predict the motion of a ball but it can also be used to predict the motion of an atom or even your blood as it moves through veins and arteries. The law can thus be used to predict whether a bone will break, how fast an animal can move and how much energy can be produced by photosynthesis.

Does physics consist of learning about blood and bones?

Not really. While physics can certainly be applied to situations in the life sciences as well as the earth and space sciences, rigorous examinations of such phenomena require ideas outside what is traditionally considered to be physics (which is why you also take courses in other subjects).

In addition, not everyone taking a physics class has the same needs in terms of the phenomena they are interested in. Instead of covering every possible situation, physics classes tend to focus on a few ideas that are very general and, as such, can be applied to a wide range of phenomena and circumstances.

For example, in volume I you learn how to apply the law of force and motion to a variety of situations, but the goal is not simply to learn about the law of force and motion. Instead, it is to learn how to apply general ideas to specific situations. The law of force and motion is simply the context in which we develop that skill.

✓ *Checkpoint 1.1: Which type of course do you need in physics: one where you memorize how to solve every possible problem you might encounter, or one where you learn how to use general laws and definitions for a wide variety of phenomena? If you don't know why you are taking this course, first talk to your advisor and ask them why.*

Summary

This chapter examined the purpose of this course and the value of learning and applying scientific laws and definitions. The main point of this section is that scientific laws and theories are valuable because of their generality; they can be applied to a wide range of phenomena and circumstances.

Frequently Asked Questions

Why do I need to learn how to do physics?

See section 1.2.

Do I need to be good at math to do physics?

The physics provides a motivation to use math, and the math provides a way to gain insight into the physics. So, it certainly helps to be good at math. However, this book is written in a way that allows you to develop your skills gradually.

The most crucial pre-requisite math skill to have is the ability to understand and utilize ratios and proportions. For example, trigonometry is delayed until near the end of volume I and even trigonometry is just a matter of ratios and proportions. There is a lot you can do with a strong conceptual understanding.[iv]

[iv] I've noticed that students weak at math tend to be good at understanding concepts.

On the other hand, the purpose of this book is to teach you physics, not math. If you are very weak at math, particularly algebra and ratios, that will likely cause problems eventually. You'll likely be better off strengthening your math skills before taking on this course.

Why is it that people complain about physics being difficult?

Physics is difficult only if you focus on the equations (rather than the concepts) and on memorizing answers to problems. If you are used to such memorization, you are not alone. Many courses inadvertently encourage this memorization. To illustrate, consider how you might approach a question in a typical exam. In many subjects, you are expected to recognize each problem or question as soon as you read it and remember the answer to it (perhaps from your notes or from the book). Does that seem familiar?

That is not the case with physics. In physics, you are *not* expected to know the answer right away. Instead, you are expected to first figure out how to answer the question based upon clues provided within the question or problem. Only after you have figured out *how* to answer the question can you then answer the question.

For example, suppose you want to predict whether four inches of bubble-wrap is sufficient to protect you from a fall out of a third story window. Don't panic if you have never seen this question before. First consider what is relevant. Is it relevant to consider what color the bubble-wrap is? What about the air pressure? How about your weight?

These are not simple questions (well, maybe the color one is) and they can be very difficult to answer if you don't have a firm foundation in physics, and in the laws of force and motion in particular. Remember, knowing physics is supposed to make it easier to solve the problems, not harder.

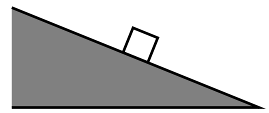
However, these same students mistakenly rely on their shaky understanding of math rather than on their strong conceptual understanding. Suppose you are in a ship that has a lot of holes and in danger of sinking. You need both putty and tape to fix the holes. You have lots of tape on board but not much putty. What would you do? Using only the putty is like relying on your math weakness to solve physics problems. Your boat will sink that way. Use what you have, while working to strengthen what you don't have.

Terminology

Hypothesis Prediction Theory

2. Objects on Inclines

Puzzle #2: In every problem involving surfaces up to now, the surface was horizontal. How does our approach change if the object is on an inclined surface, as shown in the picture?



Introduction

This chapter examines situations that involve objects on inclined surfaces. We already know enough physics and mathematics to solve such problems. Basically, we first use trigonometry on those vectors that aren't aligned with the vertical and horizontal directions. We then solve the problem using the physics, with vertical components treated separately from horizontal components.

However, this process is rather difficult when dealing with inclines.

One reason it is difficult is that for an object on an incline, only the gravitational force will be aligned with one of the component directions (vertical in this case). The surface repulsion force and the friction force will not be. That means that we have to use trigonometry on more than just one force.

A second, and more important, reason is that if there is a force imbalance on the object, it will likely be directed parallel to the inclined surface. And, since the surface is not horizontal, that means both the horizontal expression and the vertical expression will involve a force imbalance of some kind.

In this chapter, we examine two ways of addressing these issues so that the problem is easier to solve.

2.1 Tilted component directions

As mentioned in the introduction, the physics doesn't change when dealing with objects on inclined surfaces. We still have the law of force and motion and the surface repulsion force still has the same properties as before (i.e., just enough to prevent the object from passing through the surface).

The only thing we do differently is use component directions that are parallel and perpendicular to the surface, rather than component directions that are vertical and horizontal.

Can we do that?

Yes. There is nothing special about using vertical and horizontal as our component directions.[i] We should choose whatever component directions make our lives easier, as long as they are perpendicular to each other. And, in this case, we can bypass the two issues identified in the introduction by using component directions that are parallel and perpendicular to the *surface*.

To illustrate what I mean, consider the block on an inclined frictionless surface drawn in Figure 2.1. In part (a), I draw the block and the incline, along with arrows for the two forces acting on the block.

The gravitational force is downward, toward the center of Earth, as always. The surface repulsion force is directed away from, and perpendicular to, the surface as always.

✍ Note that the surface repulsion force is not always directed opposite the gravitational force. The direction of the surface repulsion force is normal (perpendicular) to the surface, as the surface prevents the object from seeping into the surface. The surface repulsion force is directed upwards only if the surface is horizontal and the object is on top of the surface.

In parts (b) and (c) of the figure, I draw the force diagram for the block along with dashed lines to represent the two perpendicular component directions. The only difference is that in part (b) the directions are horizontal and vertical, while in part (c) the directions are parallel and perpendicular to the surface.

[i] In fact, there is nothing special about using east-west and north-south as our component directions, either, except that they are perpendicular.

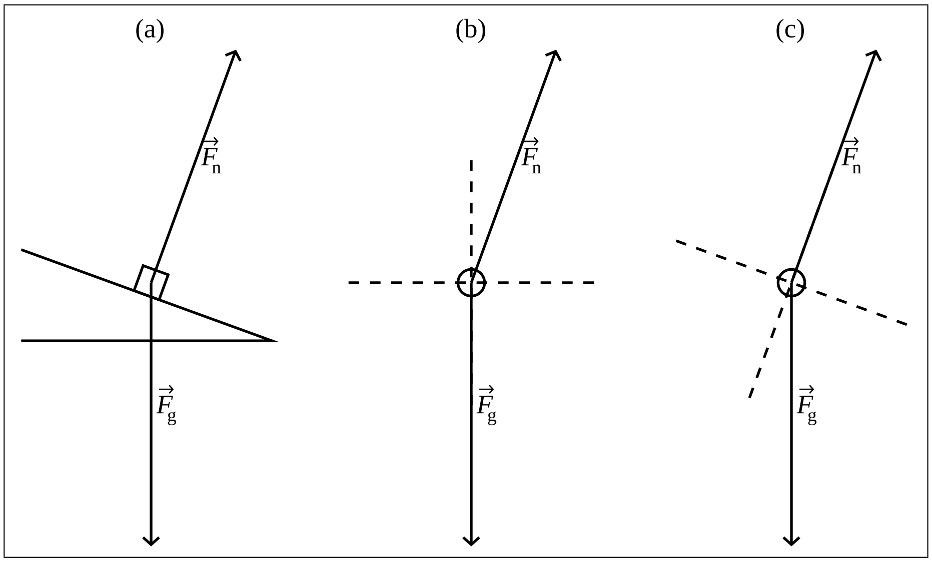


Figure 2.1: (a) A block on an inclined frictionless surface; (b) force diagram with horizontal and vertical component directions; (c) force diagram with surface parallel and surface perpendicular component directions.

Our choice of component directions does not change the physics. All it does is change the math. Indeed, all vector quantities are unchanged – they are just represently differently in terms of perpendicular components.

❢ The orientation of the component directions only changes the math, not the physics.

And, in terms of the math, the component directions in (b) mean that we need to use trigonometry on the surface repulsion force, since that force is not aligned with our component directions. On the other hand, the component directions in (c) mean that we need to use trigonometry on the gravitational force.

✍ Based on just this difference, there is not any advantage to using the component directions shown in (c). However, as mentioned in the previous section, the force imbalance (net force), if any, will be directed along the surface and it is a good idea to choose component directions such that one aligns with the force imbalance direction. The choice shown in (c) fits that criterion.

Let's see how the orientation choice influences our trigonometry. For the purpose of our analysis, I'll suppose the surface is inclined at 20 degrees above the horizontal.

- Vertical and horizontal

  With the vertical and horizontal component directions, as indicated in part (b), we would need to use trigonometry on the surface repulsion force. That force is directed between upward and rightward, giving us two surface repulsion force components: an upward component and a rightward component. Furthermore, since the surface repulsion force is directed closer to vertical than horizontal, we expect the vertical component to be larger than the horizontal component.

  Taking the sine and cosine of 20°, I get 0.342 and 0.940. The cosine, being larger, must correspond to the upward component.

  So, whatever the surface repulsion force magnitude, we'd multiply that by the sine or cosine fractions to get the upward and rightward components of the surface repulsion force. We could then carry out the physics as needed, vertically and horizontally.

- Tilted component directions

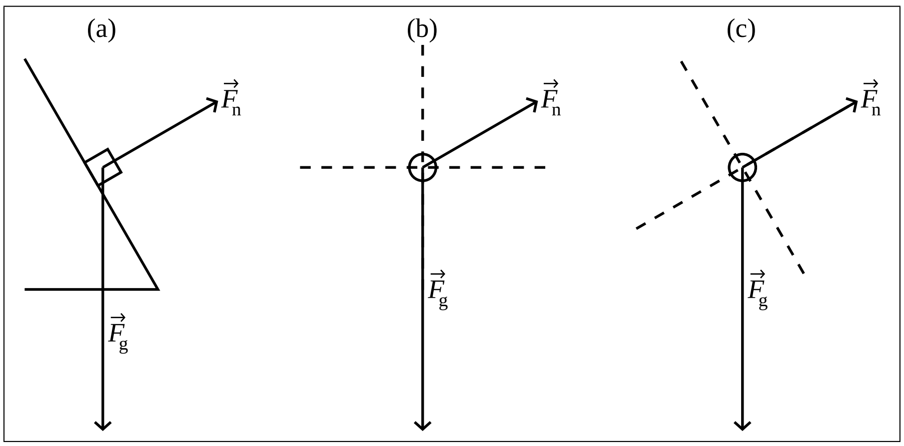
  With the component directions indicated in part (c), parallel and perpendicular to the surface, we would need to use trigonometry on the *gravitational* force. That force is directed between "into the surface" and "down along the surface", giving us two gravitational force components: a component into the surface and a component down the incline. Furthermore, since the gravitational force is directed closer to "into" than "along," we expect the component into the surface to be larger.

  Again, taking the sine and cosine of 20°as before, I get 0.342 and 0.940. The cosine, being larger, must be the component into the surface.

  Whatever the magnitude of the gravitational force, we'd multiply that by the sine or cosine fractions to get the "into" and "along" components of the gravitational force. We could then carry out the physics as needed, along the two tilted component directions as shown in part (c).

✓ *Checkpoint 2.1: Consider the block on an incline that is oriented at an angle of 60 degrees above the surface, as illustrated below.*
(a) Using the component directions in (c), which trigonometric function

would provide the component of the gravitational force that is "into" the surface: the sine of 60 degrees or the cosine of 60 degrees? Why?
(b) Suppose a horizontal force was exerted on the block directly leftward. Using the component directions in (c), which trigonometric function would provide the component of that force that is "into" the surface: the sine of 60 degrees or the cosine of 60 degrees? Why?

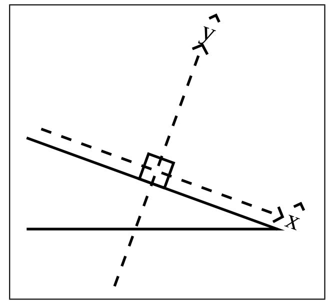


2.2 Using the hat notation

One of the things that makes inclines a bit difficult is that there are multiple directions of interest. Not only do we have vertical forces (like gravity), but we also have forces that are parallel to the inclined surface (like friction) and forces that are perpendicular to the inclined surface. We also specify the angle of the incline relative to the horizontal direction. Consequently, we have four different directions we have to pay attention to.

To distinguish between the various directions, I'll use letters for the various directions and, to make it clearer that the letter refers to a direction rather than a unit or variable, I'll include a little caret on them. For example, I'll use $\hat{x}$ to refer to the x direction or $\hat{N}$ to indicate "northward." In that way, you won't confuse the letter with a variable (as with x) or a unit (as with N).

For example, in the figure I've used $\hat{x}$ and $\hat{y}$ to indicate the two perpendicular directions, one parallel to the inclined surface and another perpendicular to the inclined surface. The caret or "hat" always goes on top of the letter, regardless of what the direction actually is (i.e., the hat on top doesn't necessarily mean the direction is "toward the top of the page").



If $\hat{y}$ is defined as a particular direction, what symbol do we use for the opposite direction?

Just as before, we use the negative sign to indicate an opposite direction.[ii] For example, if $\hat{y}$ is defined as "up" then $-\hat{y}$ means "down."

❢ Just as in one dimension, if two directions are opposite, they have opposite signs. So, the $\hat{x}$ direction is opposite the $-\hat{x}$ direction.

✍ When dealing with a component of a vector value, like the horizontal component of the velocity, I've used the letter as a subscript but without the hat (e.g., v_x).

✓ *Checkpoint 2.2: On or near Earth's surface, the gravitational force on a 2-kg object is 19.6 N downward. Suppose we set $\hat{y}$ to be upward. In terms of $\hat{y}$, what is the direction of the gravitational force on the object?*

2.3 Without friction

In this section, we consider situations where the surface is frictionless. To illustrate, consider the following situation:

> A 10-kg block is on a frictionless surface that is inclined at an angle of 20°(above the horizontal). By how much does the velocity of the box change during 2 seconds of its motion?

This is, perhaps, the classic problem in physics. It is a classic problem in physics because it looks complicated at first glance but, using what we know

[ii] I suppose you could assign a new letter to mean "down" but there is no reason to do that when we already have a way to indicate opposite directions: using negative signs.

about the surface repulsion force and the law of force and motion, we can answer it quite easily (with a little trigonometry thrown in).

We start with the physics.

There are two forces acting on the block: the gravitational force (directed downward) and the surface repulsion force (directed perpendicular to and away from the surface). The force diagram is shown in Figure 2.1 (on page 11).

There will be a force imbalance on the block because the two forces are not directly opposite in direction. In this case, there is an unbalanced portion of the gravitational force pushing the block down along the surface. As described by the law of force and motion, the force imbalance is associated with the moving along the incline (slowing on the way up or speeding up on the way down).

Next, we do the trigonometry on the forces that are not aligned with our component directions. In this case, we will use component directions that are parallel and perpendicular to the surface, as in Figure 2.1(c).

✍ I'm using $\hat{x}$ and $\hat{y}$ to refer to the directions "down the incline" and "away from the surface". As mentioned earlier, I could use any two letters, not just x and y, but I'm lazy so I tend to use the same letters over and over again. While it is quite common to represent "rightward" as the x direction, there is nothing stopping us from using $\hat{x}$ to represent some other direction.

Using the component directions in Figure 2.1(c) means that the gravitational force is at an angle, relative to our tilted component directions. The gravitational force is directed between $-\hat{y}$ (into the surface) and $\hat{x}$ (down along the surface) so we'll replace it with two perpendicular forces: one in the $-\hat{y}$ direction and one in the $\hat{x}$ direction. Furthermore, since the gravitational force direction is closer to the $-\hat{y}$ direction than the $\hat{x}$ direction, we expect the $-\hat{y}$ component to be larger.

Taking the sine and cosine of 20°as before, I get 0.342 and 0.940. The cosine, being larger, must be the $-\hat{y}$ component (into the surface).

In this case, the gravitational force has a magnitude of 98 N (multiply the mass by 9.8 N/kg), so the $-\hat{y}$ component is 92.1 N and the $\hat{x}$ component is 33.5 N.

Now we can carry out the physics.

The force imbalance is due to the $\hat{x}$ component of the gravitational force (component down the surface). That is equal to 33.5 N. From the force and motion equation, the change in velocity is the net force per mass times the elapsed time. Dividing by the mass (10 kg) and multiplying by the time (2 s), we get a change of velocity equal to 6.7 m/s down the incline (i.e., if moving down the incline, it speeds up by 6.7 m/s, and if moving up the incline it slows down by 6.7 m/s).

This solution is rather straightforward only because we knew the direction of the force imbalance (down the surface) and aligned our component directions accordingly.

✍ You'll get the same answer regardless of what mass is used in this case because the net force is proportional to the mass. Whenever that is the case, the change in velocity will be independent of the mass (since we then divide by the mass when using the force and motion equation).

✓ *Checkpoint 2.3: (a) Suppose the mass of the block was not given for the problem given above (and described in Figure 2.1). Would you still be able to determine the change in velocity of the block (given the time of 2 seconds)? If so, what is it? If not, why not?*
(b) The initial velocity of the block was not given. Does the change in velocity depend on whether the block starts at rest or is moving upward or downward?
(b) Suppose the surface in Figure 2.1 was inclined at 30° instead of 20°. Would the surface repulsion force be directly opposite to the gravitational force? Why or why not?

Do we need to know the value of the surface repulsion force?

No. We know that the net force must be parallel to the surface. The surface repulsion force is perpendicular to the surface, so it doesn't contribute to the net force.

❢ If the net force *direction* is known then the net force can be determined without knowing the magnitudes of forces that are perpendicular to that direction.

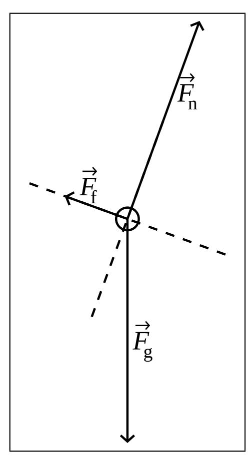
In fact, that is why it is easier to use component directions that are parallel and perpendicular to the surface: the surface repulsion force ensures that the forces are balanced along one of our component directions.

We could determine the surface repulsion force, then, by recognizing that it must balance the other forces acting perpendicular to the surface. In this case, that means that the surface repulsion force must have a magnitude equal to the component of gravitational force acting into the surface, which is 92.1 N (from trigonometry; see above).

As you can see, dealing with inclined surfaces is not much different from dealing with forces that are applied at an angle. In both cases, you'll likely have to determine the components of a force via the sine and cosine functions. The only difference is that the component directions will be tilted, in order to align with the orientation of the surface, which means a force like the gravitational force is now at an angle (unlike how it would be using horizontal and vertical axis, as in volume I).

✓ *Checkpoint 2.4: Suppose the surface in Figure 2.1 is inclined at 30°.*
(a) Would the surface repulsion force be greater or less than when the angle was 20°?
(b) Suppose the mass of the block was not given. Could you determine the surface repulsion force acting on the block? If so, what is it? If not, why not?
(c) Is there any angle for which the magnitude of surface repulsion force will equal *the magnitude of the gravitational force? If so, what angle? If not, why not?*

2.4 With friction



Adding friction adds to the mathematical complexity but the approach remains the same. To illustrate, consider the following scenario:

Suppose we again have our 10-kg block on our surface inclined at an angle of 20° above the horizontal. Let's further suppose that it is sliding down the incline and there is some friction, which is directed up the incline (opposing the sliding; see force diagram to the right).

If it speeds up at a rate of 0.5 m/s every second, what is the coefficient of friction between the box and incline?

The first step is to think about the physics.

The coefficient of friction is the ratio of the friction and the surface repulsion force (friction equation).

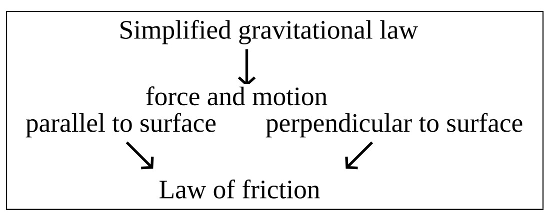
We know there is a friction because that is the reason the block doesn't accelerate as fast as before (0.5 m/s every second, compared to 6.7 m/s every two seconds), as the friction opposes the motion as the block slides down the incline. The friction doesn't totally cancel out the component of the gravitational force directed down the incline because the block still accelerates. The force and motion equation can give us the imbalance and, from that, we can calculate the friction.

As for the surface repulsion force, we know that the surface repulsion force balances the forces perpendicular to the surface. The only other force perpendicular to the surface is a component of the gravitational force.

The second step is to do the trigonometry on the forces not aligned with our component directions. As before, the only such force is the gravitational force and we know from before that the $-\hat{y}$ component is 92.1 N and the $\hat{x}$ component is 33.5 N, where $\hat{x}$ and $\hat{y}$ are as indicated on page 14.

We can now carry out the physics.

The process is illustrated in the diagram below.



We have already used the simplified gravitational law to get the components of gravity, one parallel to the surface (33.5 N; down the incline) and one perpendicular to the surface (92.1 N; into the incline).

We then apply the force and motion equation in the two component directions. Parallel to the surface, we know there has to be a force imbalance in order to for the box's velocity to change. From the change in velocity (0.5 m/s down the incline), the time (1 s) and the mass (10 kg), the force and motion equation says that the imbalance has to be 5 N down the incline.

There are only two forces acting parallel to the incline: the component of gravity parallel to the incline (33.5 N; down the incline) and the friction.

Since they add up to 5 N down the incline, that means the friction force has to be 5 N less than 33.5 N, which is 28.5 N up the incline.

✍ The friction is directed up the incline because the block is sliding down the incline. If the block was sliding up, the friction would be directed down the incline.

Perpendicular to the surface, there is no change in velocity so the law of force and motion states that the net force along that component direction must be zero.

There are only two forces acting perpendicular to the incline: the component of gravity perpendicular to the incline (92.1 N; down the incline) and the surface repulsion force. Since they add up to zero, that means the surface repulsion force must be 92.1 N (enough to balance out the y component of the gravitational force).

According to the friction equation, the coefficient of friction is the magnitude of the friction divided by the magnitude of the surface repulsion force. That gives a coefficient of 0.31.

✓ *Checkpoint 2.5: Suppose the block was sliding up the incline instead of down. Further assume the same coefficient of friction (0.31) and the same angle (20°).*
(a) Would the friction force be the same as calculated above (28.5 N up the incline)? Why or why not?
(b) Would the block still experience a change of 0.5 m/s downward in one second (obtained when the block was sliding down the incline)? Why or why not?

2.5 Vertical surfaces

Given the discussion so far about inclines, one might wonder what happens if the surface is vertical. In that case, parallel and perpendicular to the surface are just vertical and horizontal, respectively, so we can actually solve it the same way we solved problems in volume I. The only difference is that the surface repulsion force is horizontal (instead of vertical) and the friction force is vertical (instead of horizontal).

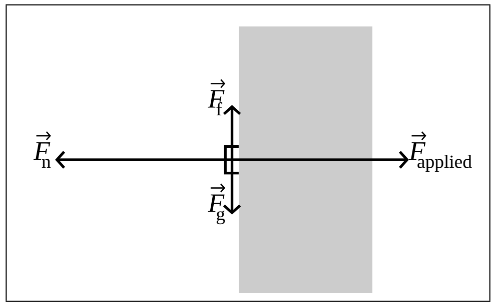


Figure 2.2: A force diagram of a book (represented by the square) held up against a wall. Four forces are exerted upon the book: the surface repulsion force due to the wall ($\vec{F}_n$), the gravitational force ($\vec{F}_g$), the friction preventing the book from sliding down the wall ($\vec{F}_f$) and a force perpendicular to the wall ($\vec{F}_{applied}$) that is due to a person pushing on the book.

For example, consider the following scenario.

> A 2-kg book is held against a rough vertical wall. By exerting a force horizontally into the wall (see Figure 2.2), a friction force results (between the book and the wall) that prevents the book from sliding down the wall. If the coefficient of friction between the book and the wall is 0.3, how much force do I need to exert on the book to keep the book stationary?

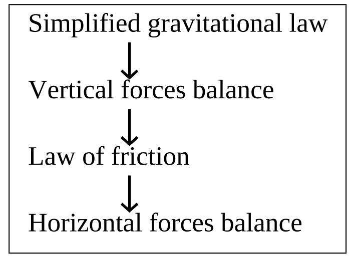
The physics tells us that the forces must be balanced in order for the book to remain at rest.

Given the situation, you might wonder why would I have to apply a force to keep the book stationary.

The answer can be found in the force diagram see Figure 2.2). Notice how the gravitational force is downward (as always) whereas the surface repulsion force is leftward (perpendicular to the surface). They are not parallel. The surface repulsion force balances the applied force in this case, not the gravitational force. The reason why the book doesn't fall is because of friction along the wall. The friction balances the gravitational force.

To produce the friction, though, I need to press the book against the wall. This increases the surface repulsion force. The magnitude of the friction force is proportional to the magnitude of the surface repulsion force.

This reasoning also helps us to identify the steps we must follow to answer the question:



We can get the gravitational force from the *simplified gravity equation*. Then, using the idea of *vertical force balance*, we can get the friction force.

Knowing the friction force, we can use the *friction equation* to get the surface repulsion force and then *horizontal force balance* to get the applied force.

The solution, then, is to carry out these four steps using numbers.

Since the book is near the surface of the Earth, the magnitude of the gravitational force is the mass of the book (2 kg) times 9.8 N/kg. In this case, that would be 19.6 N.

Since the forces must balance, the friction force (acting upward) must balance the gravitational force (acting downward). That means the friction force is also equal to 19.6 N.

Since we know the coefficient of friction, we can get the surface repulsion force from the friction equation ($\mu = |\vec{F}_{\mathrm{f,max}}|/|\vec{F}_{\mathrm{n}}|$). Since the coefficient of friction is 0.3, I get a surface repulsion force of 65.3 N. Do the math yourself, so you can get a sense of how to use algebra to solve for the surface repulsion force.

Since the surface repulsion force balances the applied force, that means the applied force must likewise be 65.3 N (into the wall).

✍ This process is very similar to the process carried out in volume I where the same setup is provided but what may appear at first glance to be a very different problem because of what is being asked.

Can more than 65.3 N be applied?

Yes. If the applied force is more, the surface repulsion force would likewise increase, leading to a greater maximum friction force. That just means the friction can be more than 65.3 N, if needed. However, the friction force only needs to be 65.3 N, so increasing the maximum it can be does not change the actual value it would have.

Can less than this be applied?

Not if we want to keep the book at rest. If the applied force is less, the surface repulsion force would likewise decrease, leading to a smaller maximum friction force. That means the friction would no longer be able to be 65.3 N, the amount need to balance the gravitational force.

✓ *Checkpoint 2.6: For the box on a horizontal surface, the magnitudes of the surface repulsion and gravitational forces were equal. For the book on a vertical surface, the magnitudes of the surface repulsion and applied forces were equal. Why the gravitational force in one situation and not the other?*

Summary

This chapter examined how the orientation of the component directions can be changed in order to simplify the solution. Basically, we orient the component directions such that one of the two aligns with the net force direction. That way, the net force along the other component direction is known (i.e., it is zero).

The main points of this chapter are as follows:

- The orientation of the component directions only changes the math, not the physics.
- Just as in one dimension, if two directions are opposite, they have opposite signs. So, the $\hat{x}$ direction is opposite the $-\hat{x}$ direction.
- If the net force *direction* is known then the net force can be determined without knowing the magnitudes of forces that are perpendicular to that direction.

By now you should be able to use the law of force and motion with any problem that involves an inclined surface.

Frequently Asked Questions

Should the component directions always be parallel and perpendicular to the surface?

No. The problem can be solved regardless of the orientation we choose for the component directions. However, the mathematics is simpler (and the result easier to interpret) if the directions are chosen such that one direction is oriented along the net force direction.

If the gravitational force is downward, isn't it already aligned with the component directions?

Not if the component directions are tilted relative to vertical and horizontal.

Remember, the orientation of the component directions are arbitrary – we can choose whatever orientation we want – as long as the two directions are perpendicular to each other. For inclined surfaces, it may be easier to use component directions that are perpendicular and parallel to the surface, not vertical and horizontal.

Doesn't the surface repulsion force have to balance the gravitational force?

No. The two do not have to balance nor do they have to have the same magnitude. The surface repulsion force only has to be sufficient to prevent the object from crossing through the surface. See volume I for more information.

Is the surface repulsion force always directed upwards?

No. It is directed upwards only if the surface is horizontal.[iii] See page 10.

Is the surface repulsion force always directed vertically (i.e., up or down)?

No. The surface repulsion force is always *perpendicular* to the surface. So, the surface repulsion force is *vertical* only if the surface is *horizontal*. If the surface is *vertical*, like with a wall, the surface repulsion force is *horizontal*.

Problems

Problem 2.1: A 1.5-kg block is sliding up an inclined frictionless surface at an initial speed of 5 m/s. If the surface is inclined at 20° above the horizontal, how far up the surface does the block slide before coming to rest?

[iii]Technically, the object also has to be on *top* of the surface for the surface repulsion force to be upward. If you jump up and hit your head on the ceiling, the ceiling exerts a surface repulsion force on you, but it is directed *downward*, since the surface in that case is *above* the object (your head), not below it.

Problem 2.2: Why is it advantageous to choose component directions that align with one or more of your forces when determining the net force?

Problem 2.3: An object is placed on an inclined surface.
(a) Does the direction of the surface repulsion force depend on whether the surface is frictionless or not? How about whether the object is moving or not?
(b) Does the direction of the gravitational force depend on whether the surface is frictionless or not? How about whether the object is moving or not?
(c) Does the direction of the friction force depend on whether the object is moving upward, moving downward or at rest?

Problem 2.4: Consider the following three frictionless surfaces: (a) a horizontal surface, (b) a surface inclined at 30 degrees, and (c) a vertical surface. Suppose a box is in contact with the surface. In which case is the surface force exerted on the box parallel to the gravitational force exerted on the box? Why?

Problem 2.5: A 2-kg object has the following three forces exerted on it:
$F_1 = 120$ N at $0°$
$F_2 = 200$ N at $126.87°$
$F_3 = 160$ N at $270°$
(a) Using component directions oriented at $0°$ and $90°$, calculate the net force along each component direction then calculate the magnitude and direction of the net force.
(b) Suppose I shift each direction by $60°$ as follows:
$F_1 = 120$ N at $60°$
$F_2 = 200$ N at $186.83°$
$F_3 = 160$ N at $330°$
Again using component directions oriented at $0°$ and $90°$, calculate the net force along each component direction then calculate the magnitude and direction of the net force.
(c) Compare the magnitudes of your two answers. Should they be the same? Why or why not?
(d) Compare the directions of your two answers. Should they be the same? Why or why not?

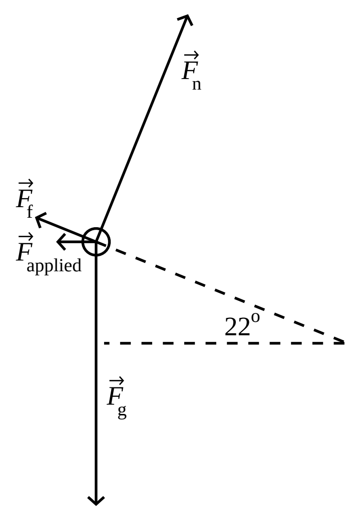
Problem 2.6: A 10-kg block is placed on a surface that is inclined at an angle of $20°$ above the horizontal. Suppose the coefficient of friction between the box and surface is 0.3.

(a) If the block doesn't move, is there friction between the block and the surface? If so, in which direction? If not, why not?
(b) Determine the minimum coefficient of friction between the block and incline that would prevent the block from sliding.
(c) if the block is sliding down the inclined surface, in what direction is the friction force? What is the acceleration of the block?
(d) If the block is sliding up the inclined surface, in what direction is the friction force? What is the acceleration of the block?

Problem 2.7: (a) Suppose we have the same situation as described in the previous problem but the surface was frictionless. Determine the net force acting on the block.
(b) Suppose there was friction present with a coefficient of friction between the box and surface equal to 0.3. Suppose further that the block was sliding *up* the inclined surface, so that the friction was directed *down* the inclined surface. Draw the appropriate free-body diagram and determine the net force acting on the block.

Problem 2.8: A 10-kg block is on a surface that is inclined at an angle of 22° above the horizontal. When I exert a horizontal force F_{applied} on the block with a magnitude of 14 N (see force diagram to right), I am able to keep the block moving down the incline with a constant speed. Determine the coefficient of friction between the box and incline.

Notice that the direction of the friction force is *up* the incline because the block is sliding *down* the incline. If the block was instead sliding *up* the incline, the friction would be *down*.



Problem 2.9: When you lean against a wall, what is the direction of the surface force exerted on you due to the wall? Ignore friction.

Problem 2.10: A book is held against a rough vertical wall. By exerting a horizontal force of 50 N into the wall (perpendicular to the wall), a friction force (between the book and the all) keeps the book at rest and prevents it from sliding down the wall. If the coefficient of friction between the book and the wall is 0.5, what is the maximum mass of the book?

3. Definitions, Laws and Theories

3.1 Definitions

A **definition** specifies how a particular quantity relates to other quantities. For example, the definition of velocity relates the velocity with displacement and time. Definitions are important because we don't really understand what something is until we define it in terms of other things that we already know.

For example, suppose we want someone to know what we mean by the word "table." To do this, we'd have to describe it in terms of things they are familiar with. One definition of a table, then, may be that it is a piece of furniture with a flat level top and one or more legs. That means you can use the word "table" for any piece of furniture with a flat level top and one or more legs. However, this definition doesn't help if one doesn't know what "furniture" means. We'd have to either use a different word or, in turn, provide a definition of the word "furniture".

Definitions, as with any new word, are things that you just need to memorize. Fortunately, as you continue to use a term, it becomes easier and easier to remember what it means.

In physics, you'll probably be familiar with most of the terms already. That can be good and bad. The more familiar you are with a term, the less work you have to do to learn it. On the other hand, you may not know the *exact* definition and your familiarity with the term may lead you to overlook some aspects of the term that are crucial for using it correctly.

For example, you may already be familiar with the term "velocity" as relating to an object's speed but you may not know the *exact* definition. While similar to speed, velocity doesn't exactly mean speed, and you'll likely make errors if you treat them as the same thing.

As I introduce terms and their definitions, there are two things to keep in mind.

The first thing to keep in mind is that there has to be some *value* in defining a new term. We don't just introduce terms because we like it.

Consequently, I'll be careful to introduce terms with precise definitions, when possible, and explain *why* the term is useful. Most of the time, a term is introduced because it either simplifies the description of how things interact or simplifies our explanation of how things interact.

✍ In physics, we use a lot of terms that may appear at first glance to be the same. Keep in mind that a new term would have no value if we already had a term that meant the same thing.

The second thing to keep in mind is that definitions are things that the scientific community has decided upon. Nature doesn't care how we define things. We can't test whether a definition is true or not, or prove that it is or isn't true. Basically, a definition is true because, well, that is the way it has been defined.

❢ Defined relationships are *always* true, by definition.

We can question the *value* of a definition and we can question whether people *use* the term in a way consistent with the definition, but we can't debate the definition itself.

The disadvantage of this is that you need to make sure you understand the *exact* definition and use it in a way consistent with that definition. On the other hand, the advantage of this is that we don't have to worry about whether we can or can't apply a definition. We always can. It just may not be *useful* to do so.

✓ *Checkpoint 3.1: In volume I, an object's momentum is defined as the product of the object's mass and the object's velocity. Is this something that has been proven to be true? If not, could there be instances when the momentum is <u>not</u> equal to the product of the object's mass and its velocity?*

3.2 Scientific laws

Defining "law" is a bit tricky because scientists use the word a little differently than the general population. A scientific **law**, like a definition, describes how two or more things are related but, unlike definitions, the purpose of a law is not to *introduce* a new term. Instead, the purpose of a law to describe how two or more measurable quantities are related, so that we can **predict** one of those quantities.

For example, the law of force and motion relates changes in an object's motion with the forces exerted on that object. Each quantity, the changes in motion and the forces, can be measured separately. With the law, we can predict one by knowing the other. I'll talk more about making predictions in the next section.

ARE THEY CALLED LAWS BECAUSE THEY HAVE BEEN PROVEN CORRECT?

Scientists use the word "law" and "prove" a little differently than how the general population uses the words. In science, we don't *prove* relationships and ideas to be correct. Instead, we *test* whether the relationships and ideas hold. The better they hold, the more useful they are in making reliable predictions.

So, while many scientific laws are treated as being true in all cases, that is not necessarily the case. For example, the law of force and motion is extraordinarily accurate when applied to a wide range of ordinary situations, from bacteria to galaxies and lots of things in between, which is why it is the focus of volume I. However, there are still some situations where we can't apply the law of force and motion.[i]

In addition, whereas a law outside of physics is something a community decides that everyone must follow, a scientific law is just our attempt to describe something that *nature* follows.

IF A LAW ISN'T PROVEN, WHY ARE SOME LAWS ASSOCIATED WITH PARTICULAR PEOPLE?

It is often the convention to name a law after the person who first identified and presented it. For example, most people refer to the law of force and motion as **Newton's second law**, after Isaac Newton.[ii] However, as noted

[i] This includes quantum mechanics, special relativity, general relativity, and non-inertial frames. We won't examine such situations.

[ii] Isaac Newton was an English scientist who lived from 1643 (or 1642 according to the local calendar in use at the time) to 1727. Apparently, his interest in mathematics didn't really start until he was 20. When he was 22, an epidemic (associated with the bubonic plague) closed the university from which he had just graduated, so he returned to his home in the country. During the next two years, he laid the foundation of calculus

earlier, laws aren't proven to be true. They are simply our descriptions of how various quantities are related.

Indeed, I've chosen to call it the law of force and motion in order to make this clearer. Not only is it more reflective of what the law describes, but I also want to emphasize that it is irrelevant who the scientific community has decided to recognize as being the first to propose the law in a formal way.

✓ *Checkpoint 3.2: The law of gravity is often called Newton's Law of Gravitation. Does this mean that Newton was the first to prove this law?*

3.3 Making predictions

The single most valuable skill that a scientist has is the ability to make predictions. To illustrate what we mean by predictions and how a scientific law is a relationship that can be used to make predictions, consider the following story.

> A student, named Jay, wants to drive to Philadelphia next month. So he asks his friend Kaye how long it will take him to drive to Philadelphia. Kaye says three hours.

How long will it take Jay to drive to Philadelphia?

If you said "three hours" you are using a form of logic called "appeal to authority." The real answer is "we don't know." Just because Kaye says it will take three hours doesn't mean it will take three hours. No one knows for certain how long it will take Jay until he actually takes the trip. By relying solely on Kaye's estimate, we are putting our faith in Kaye's expertise.

(which he called the method of fluxions) and optics. It was also during this two-year period that he developed early versions of several of the laws we'll be using in this book. A formal presentation of his work wasn't published until he was 44 when he published the *Mathematical Principles of Natural Philosophy* (commonly known as the *Principia*, from the original Latin title), one of the greatest scientific works ever published.

> Knowing this, Jay asks Kaye *why* she thinks it will take three hours. Kaye responds that last month she drove to Philadelphia and it took her three hours.

How long will it take Jay to drive to Philadelphia?

If you now say "three hours" you are making a prediction based upon something more than just an appeal to authority. You are basing it upon an **observation**.

The real answer, though, is that we still don't know. Just because it took Kaye three hours doesn't mean it will take Jay three hours. Again, no one will know for certain how long it will take Jay until he actually takes the trip. By relying on this single observation, we are assuming that the time would be the same for everyone.

> Knowing this, Jay asks another friend, Elle, how long it will take him to drive to Philadelphia. Elle says *two* hours. Jay, having read through the previous couple of paragraphs, astutely asks Elle why she thinks it will take two hours. Elle responds that she took the trip yesterday and that is how long it took her.

How long will it take Jay to drive to Philadelphia?

Jay's in a dilemma now. He has two observations and they don't agree. Apparently, the assumption made previously, that the time would be the same for everyone, is incorrect.

> Stumped, he asks another friend, Bea, for some help. Bea has never taken the trip to Philadelphia but Bea is a scientist. And, as a scientist, she is used to making predictions based upon observations.
>
> The key, Bea asserts, is to identify the pattern. Noting that it took Kaye three hours *last* month and it took Elle two hours *this* month, Bea points out that the time went down by one hour from one month to the next. By next month, Bea reasons, the time will go down one hour further. Since Jay will be driving to Philadelphia *next* month, Bea boldly predicts it will take him one hour (i.e., one hour less than it took Elle to drive there *this* month).

How long will it take Jay to drive to Philadelphia?

If you now say "one hour" you are making a prediction based upon an observed *relationship*. Such a relationship is called an **empirical** relationship (from a Greek word meaning "experience").

> ✍ An example of an empirical relationship might be the relationship between saturated fat and heart disease. Suppose someone does a study and finds that people with diets high in saturated fat have a higher incidence of heart disease. That would be an empirical relationship, since the pattern was observed to hold true, but no explanation was provided to explain why such a relationship might be.

In this case, the relationship is between the drive time and the month.

However, just because this relationship holds for the two observations does not mean it will continue to hold for all observations. As before, the real answer is that we still don't know how long it will take.

In general, though, the more observations that are used to identify the relationship, the more likely the relationship will hold for future observations.

> By now, Jay is obsessed with getting a good prediction. Determined to find the pattern between the drive time and the day of the trip, he asks all of his friends who have driven to Philadelphia in the past month. He finds that each day the time went down by two minutes. If two people drove there a week apart (seven days), the time difference was 14 minutes. If two people drove there a month apart (30 days), as in the case of Kaye and Elle, the time difference was 60 minutes. The later in the month one drove, the less time it took. The new observations *support* the empirical relationship identified by Bea, which Jay now confidently calls "Bea's law of driving time."

How long will it take Jay to drive to Philadelphia?

If you now say "one hour" you are making a prediction based upon Bea's law.

Jay refers to the Bea's empirical relationship as "Bea's law" because it is relationship that seems to hold really well. In such cases, it is common to call it a **law**.

An important point about scientific laws, such as Bea's law, is that just because we call it a law does not mean it has been *proven* to always hold true. What makes laws powerful is not that they have been proven to be true but because they are very general.

For example, Bea's law applies regardless of what time we apply it over the past month. We still don't know the answer to Jay's question (how long will it take him to drive to Philadelphia), however. Just because this relationship holds for all of the observations made so far does not mean it will continue to hold for all observations in the future. Calling it a law doesn't make it perfect.

If we can never know, why bother making any predictions at all?

Although we can never be 100% sure our prediction will be right, the fact is that scientific analysis allows us to make the prediction in the first place. *That* is the value of science.

3.4 Laws vs. theories

What is the difference between a law and a theory?

In everyday language, the definitions of the two terms may be ambiguous. In science, however, they mean very specific, and different, things.

To illustrate the difference, and show how generality provides more powerful predictions, let's continue with our story.

> Jay, confident in his prediction that it will take him one hour to reach Philadelphia, arranges to leave one hour before an important meeting. He happens to mention this to another scientist friend of his, Dee. Being a scientist, Dee recognizes the predictive power of an empirical relationship, especially one based on as many observations as this one. However, she suggests that Jay use an even more powerful method.
>
> While empirical relationships are powerful, Dee explains, you have to be careful not to apply them to situations that are different from those in which they were developed. She points out

that there was construction being done on the route to Philadelphia during the month Jay had done his study. When Kaye went, there were many miles of construction. As time went on, parts of the route were finished and the commute time went down. By the time Elle drove to Philadelphia, all of the construction was finished.

Consequently, if we assume that there is no more construction on the route (since all construction was finished by the time Elle made her trip), everyone who drives after Elle will find that the commute takes the same amount of time: 2 hours. Jay's commute time should be the same as Elle's since his driving conditions would be the same as hers.

How long will it take Jay to drive to Philadelphia?

If you now say "two hours" you are making a prediction based upon Dee's explanation.

Once again, the real answer is that we still don't know for sure. While knowing the explanation for the relationship helps us recognize that we cannot simply extend Bea's law indiscriminately, there might still be *other* additional factors that can influence the driving time (e.g., whether Jay drives during rush hour or whether it is raining). The more Jay learns about *why* the driving time would be long or short, the more capable he becomes in making a prediction.

If it turns out that it actually takes two hours for Jay to drive to Philadelphia, Dee's explanation would be supported. In science, an explanation that is well supported by the observations is known as a scientific **theory**.

In general, theories are more powerful than laws because we can use them to determine when the laws apply and when they don't.

On the other hand, we don't need to have a theory behind every law. There are certain situations where we just don't know why a law exists. Still, all observations have been shown to support it and that is why we call it a law.

✍ Even though some laws, which I call *general laws*, can be used to explain why other relationships exist, in my book a theory is a theory only if it is a model that explains *why* a relationship exists. If it is simply a relationship, even if it is one that can be used to "explain" other relationships, I will refer to it as a law.[iii]

So, a scientific theory isn't just a guess that, when proven, becomes a law. First of all, theories don't become laws, as discussed above. Theories explain why a relationship exists while laws describe what the relationship is. Second, it is not possible to *prove* laws or theories to be correct. Rather, we only *support* them via additional testing.

3.5 Hypotheses and testability

What may confuse the issue for non-scientists is the similarity between the word "hypothesis" and the word "theory."

To non-scientists, a **hypothesis** is just an educated guess, much like they would use the word theory or prediction. We've already discussed what a scientific theory is. What about a scientific hypothesis?

To scientists, a hypothesis is more than an *educated* guess. It is a *testable* guess. An educated guess isn't useful. To be useful, the guess must be *testable*.

To test an idea, we need to be able to utilize the idea to make a prediction that can then be compared with observations.

Does that mean an observation must be repeatable?

If by "repeatable" you mean that you will observe the same things only when the conditions are "reset" to exactly how they were, the answer is no. That is one way but not the only way. In fact, most times it is impossible to set up things exactly how they were. Fortunately, all that is needed is that the prediction needs to be specified in terms of an observation that has yet to be made.

For example, one prediction I can make with Bea's law of driving time is that if Jay drives to Philadelphia next month his trip will take one hour. Alternatively, a prediction I can make with Dee's theory is that if there is construction on the route to Philadelphia then Jay's trip will take longer than two hours.

[iii] Not surprisingly, the line between theories and laws is not universally agreed upon. Don't be surprised if you run across relationships in science that don't adhere to the strict definition of law and theory provided here.

If the prediction turns out to be correct, does that mean I've proven the hypothesis?

No. As mentioned in the story of Jay's trip to Philadelphia, we can never say for *certain* that a law or theory is true. On the other hand, the evidence in support of a theory or law can be so overwhelming that its validity is "beyond a reasonable doubt." Just keep in mind that what counts as "reasonable" is subjective.[iv]

Is this textbook about testing laws of force and motion?

Not really. Although we frequently check our predictions to make sure they are reasonable, the bulk of the text focuses on the skills and techniques needed to apply relationships to specific situations.[v] Still, you should keep in mind that the predictions we make should be testable. If they don't seem supported by your observations, check with the instructor.

So how long will it take for Jay to travel to Philadelphia?

We won't know how long it will take until he actually makes the trip. All we know is that Bea's law predicts that it will take one hour and Dee's theory predicts that it will take two hours.

If it does indeed take two hours then Dee's theory is supported and Bea's law is not. On the other hand, if it takes one hour than Bea's law is supported and Dee's theory is not.

Either way, a theory doesn't become a law or visa-versa. If Dee's theory is supported, that won't make it a law and if Bea's law is not supported, that won't make it a theory.

> ✍ A real-life example that illustrates the difference between theory and law is an examination of how the pressure of a gas depends upon its volume (for a given temperature and mass). We call the observed relationship between the two Boyle's law whereas we use the Kinetic-Molecular Theory to explain the relationship in terms of the molecular model.

[iv] For example, while scientists tend to be convinced in the validity of the second law of thermodynamics, many non-scientists assume it is false in their pursuit of perpetual motion machines or machines that make energy from nothing.

[v] The course also includes a lab component where you learn techniques needed for testing laws and theories. The activities serve to illustrate what it means to find agreement (or disagreement) between a prediction and an observation.

Why be so picky about the difference between a law and a theory?

Because conceptually they are a little different and so how you use them is a little different. In addition, the subject of volume I of this book is on a small set of general laws. As laws, not theories, it is important to recognize that they *describe* certain relationships, not *explain* why those relationships occur. On the other hand, as general laws, they can be used in lots of different situations. How we apply them to the various situations is the focus of the bulk of this text.

I've taken a lot of science courses in my life and I've never needed to know the difference between law and theory. Why is it so important in this course?

Chances are your other courses focused on the *results* of science rather than the *process* of science. It turns out, however, that the real power of science is not in the answers but in making predictions and using a small subset of general ideas to make specific predictions (like using a map to determine specific travel routes). To do this, you need to have a firm grasp of how laws and theories are identified and used to make the predictions.

Science, it turns out, is *not* simply a list of facts about nature. Rather, it is a *process* by which laws and theories are identified and tested.

In this volume, we explore the meaning of some laws of force and motion and introduce techniques for applying them to make specific predictions. If you aren't comfortable with the meaning and limits of the laws, you end up missing the point that the laws form the framework for everything we'll do. Instead, you'll see everything as a million disconnected equations.

Terminology

Definition	Law	Observation
Empirical	Newton (Isaac)	Predict
Hypothesis	Newton's second law	Theory

Frequently Asked Questions

Are all relationships considered to be laws?

Usually we reserve the word "law" for those relationships that hold really well. After all, just because a relationship holds for a bunch of observations does not mean it will continue to hold for all observations. In general, the more observations that agree with the relationship, the more likely the relationship will hold for future observations.

Which is more powerful: a law or a theory?

It depends on the law and theory. Both a general law and a general theory can be used to consolidate what at first glance appears to be disconnected phenomena. So, just as a few general laws (e.g., the law of force and motion) can be applied to a large number of very specific situations (rather than coming up with a new empirical relationship for every situation we encounter), a few general theories (some of which will be discussed in this volume) can be applied to a large number of very specific situations (rather than coming up with a new explanation for every situation we encounter). So, like the general laws we examined in volume I, a good general theory can be very powerful.

However, since a law simply *describes* a relationship that was observed, it cannot be used to predict when the relationship will hold and when it will not hold. On the other hand, an *explanation* for the relationship, because it explains *why* the relationship holds, can also be used to explain why the relationship would *not* hold (i.e., when the assumptions inherent in the model are not valid). In that sense, a theory can be more powerful.

Is a law simply a theory that has been proven correct?

No. This is a common misconception. Laws, like the ideal gas law or Hooke's law, are called laws because they describe relationships between variables. Theories are called theories because they are models that *explain* observations. For example, the ideal gas law *describes* how properties of gases (like pressure and temperature) are related whereas the kinetic-molecular theory is based upon using a model of individual molecules in motion to *explain* why gases behave the way they do.

Does "theory" mean "guess"?

To non-scientists, maybe, but not to scientists. To scientists, theories are models that are well-supported by observations. Some theories are better supported and have been more rigorously tested than some laws.[vi] To scientists, “guesses”, if you can call them that, are called **hypotheses**. Hypotheses differ from guesses in that hypotheses are testable.[vii]

WHAT IS THE DIFFERENCE BETWEEN “TESTABLE” AND “REPEATABLE”?

To be testable, one must be able to make a prediction as to what one might observe in a given circumstance. A phenomenon need not be repeatable to be testable. For more information, see page 35 in section 3.

Problems

Problem 3.1: (a) Identify a relationship or model in a field of science (it need not be physics) that is called a *theory* (i.e., the kinetic molecular theory, the atomic theory, the theory of biological evolution, the theory of plate tectonics, etc.). Write down the name (if it has one) and also describe the relationship or model.
(b) Does the theory identified in (a): (i) describe a observed pattern (e.g., describe how two things are related) or (ii) explain why the pattern occurs?
(c) Based on your answer, how well does the relationship correspond to the definitions provided here (i.e., theory vs. law)?

Problem 3.2: Suppose there is a theory that states that there is a parallel universe that exists but we cannot exchange any information with it and that is why there is no evidence of it. Is this theory scientific according to the definition of testable presented in section 3.5? Why or why not?

Problem 3.3: According to the American Institute of Biological Sciences, biological evolution consists of change in the hereditary characteristics of groups of organisms over the course of generations. There is a theory that states that this change is related to natural genetic variations within a species (i.e., that characteristics that are beneficial are more likely to be found in succeeding generations). Is such a theory scientific? If not, why not? If so, is there a way to test it without having to observe succeeding generations?

[vi] For example, Ohm’s Law, which is discussed in Volume II, has been shown to fail in certain situations yet it is still called a law.

[vii] To see what we mean by “testable,” see volume I.

Problem 3.4: A law describes a relationship. For example, Bea's law of driving time describes the relationship between the driving time and the day of the trip. However, a law can also describe the *lack* of a relationship. An example of such a law in science is *Mendel's law of independent assortment*, which states that the characteristics inherited from one parent is independent of the characteristics inherited from the other parent.
(a) In Bea's law, the driving time is being related to the day of the trip. In his law of independent assortment, what two items is Mendel stating are *not* related?
(b) Provide a prediction that one might make based upon Mendel's law of independent assortment. Your prediction should involve the variables you identified in (a). For example, in Bea's law, a prediction might be that Jay's driving time next month would be one hour.

Problem 3.5: The Big Bang Theory is a popular theory for explaining why distant galaxies appear to be traveling away from us at great speeds. One of its predictions is the existence of cosmic background radiation (which was eventually observed several years later). A friend argues that the Big Bang Theory isn't scientific because it describes an event (the start of the universe via a big bang) that cannot be repeated. How do you respond?

Problem 3.6: Why is "Bea's law of driving time" called a law instead of a theory? Hint: Does it explain why the relationship between the drive time and the day of the trip is the way it is or does it simply provide the relationship between the drive time and the day of the trip?

Problem 3.7: The American Physical Society provides the following definition of science.

> Science is the systematic enterprise of gathering knowledge about the universe and organizing and condensing that knowledge into testable laws and theories.

In comparison, Merriam-Webster Online defines science as follows:

> Knowledge or a system of knowledge covering general truths or the operation of general laws (especially as obtained and tested through scientific method)

One important part of the APS definition of science is that science is a **process**. Does the Merriam-Webster definition emphasize the procedural nature of science? How could it be improved?

4. Variable Abbreviations

Letter abbreviations are used to represent variable values in equations. For example, we could use s to represent distance and t to represent time. The abbreviations s and t are known as **variable abbreviations**. The two variables, **distance** and **time**, are represented by the abbreviations s and t.

✍ Mathematical abbreviations are used for both units and variables in mathematical equations. Variable abbreviations will be italicized whereas unit abbreviations will not (see section 6 for the list of unit abbreviations).

Be careful. An upper-case letter typically means something different than a lower-case letter (compare, for example, a and A). In addition, we sometimes use Greek letters (e.g., λ for wavelength). Subscripts are used to distinguish between closely-related variables (or variables corresponding to different times).[i]

Also, be aware that some letters are used for two or more variables (see, for example, P being used for both power and pressure). To avoid confusion, the duplication will involve variables that will rarely, if ever, be in the same relationship. Still, don't be surprised when you see the same letter used for a different variable in a different context.

To indicate quantities that are vectors (i.e., have a direction), I use an arrow above the letter (e.g., $\vec{F}$). The magnitude of a vector is indicated by vertical bars (e.g., $|\vec{F}|$. A two-dimensional vector can be written as the sum of two perpendicular vectors, called projections. Subscripts are used to indicate a vector projection. For example, $\vec{F}_x$ indicates the projection of the force vector along the x direction. Such projections can also be written as the product of a component value and a direction. For example, $\vec{F}_x$ could also be written as $F_x\hat{x}$, where $\hat{x}$ is the x direction.

[i] As you get more familiar with certain relationships, you might find the subscripts are unnecessary and, at that point, you are free to drop them. After all, knowing the history (and being familiar with the unit abbreviations) is the best way to avoid getting too confused. However, until you reach that point, I recommend that you keep the subscripts.

The list of variable abbreviations is like a vocabulary list.[ii] And, just like a language, knowing what the letters represent becomes easier the more you use them. Fortunately, our "language" is mathematics so you don't have to also learn the structure and syntax of a different "language."

To learn the "language," you should probably write down each abbreviation you encounter in a list somewhere as a reference. Section 4 is a good reference but you won't want to constantly be referring back and forth to the list. By writing your own list, you will be able to recall them quickly without checking the list every time.

Δ	finite change or difference
$\Delta\vec{s}$	displacement (change in position)
$\Delta\theta$	rotational (angular) displacement
$\Delta\vec{s}_x$	displacement in $\hat{x}$ direction
$\hat{x}$	the x direction (i.e., unit vector in x direction)
α	rotational (angular) acceleration
κ	dielectric constant
λ	wavelength
μ	coefficient of friction
μ	magnetic permeability
Φ	magnetic flux
π	ratio of circle circumference to diameter
ρ	density
θ	angle or rotational (angular) position
θ_i	incident angle
θ_r	reflected angle
θ_t	transmitted angle
ρ	resistivity
ρ	density
τ	torque (rotational force)
ω	rotational (angular) velocity (angular frequency)

[ii] Scientists in different fields may use different conventions for their variable abbreviations. For example, whereas I am using s for position, someone else might use x. Throughout this text I will use the convention used in physics unless I see a pedagogical advantage for changing it. I can't change it too much because then no one outside of this class would understand it.

a acceleration
a_{cent} acceleration toward the center of the circle
a_{circ} acceleration around the circle
A cross-sectional area
A Avogadro's number

B magnetic field

C capacitance
c speed of light in a vacuum

d slit separation distance
d_i image distance
d_o object distance
D diameter of a circle

E energy
E_{e} electric energy
E_{g} gravitational energy
E_{k} kinetic energy
$\vec{E}$ electric field
$\mathcal{E}$ open circuit voltage (emf)

f frequency
f focal length
F magnitude of force
$\vec{F}$ (vector) force
F_{circ} force directed around the circle
F_{e} electric force
$\vec{F}_{\text{f}}$ frictional force
$\vec{F}_{\text{g}}$ gravitational force
F_{m} magnetic force
$\vec{F}_{\text{n}}$ normal (surface) force
$\vec{F}_{\text{T}}$ force due to a string, rope or cable

g gravitational field strength
G gravitational force constant

h height
h_{i} image size
h_{o} object size

I rotational mass (moment of inertia)
I current

J Impulse

k electric force constant

L angular (rotational) momentum
L inductance (or self-inductance)
ℓ length (e.g., of wire)

m mass
m magnification
M mass (large)
m_e mass of electron
m_p mass of proton

n index of refraction
n whole number (as in number of elements)
N whole number (as in number of loops)

P power
P pressure
p momentum

q charge
q_e charge on an electron
q_p charge on a proton
Q charge stored on a capacitor

r distance between two objects
r length of radial arm or distance from object to axis
r internal resistance
R radius of an object or circle
R resistance

s (magnitude) position
$\vec{s}$ (vector) position
s_x position in $\hat{x}$ direction

t time
T period
T temperature

v	speed (magnitude of velocity)
$\|\vec{v}\|$	speed (magnitude of velocity)
$\vec{v}$	(vector) velocity
v_{circ}	velocity (or speed) around the circle

w	width (of slit)
W	work

X_C	capacitive reactance
X_L	inductive reactance

Z	impedance
Z_{cap}	impedance of the capacitor
Z_{ind}	impedance of the inductor
Z_R	impedance of the resistor

5. Constants

Astronomical constants:

mass of the sun	1.9891×10^{30} kg
mean Earth-Sun distance (center to center)	1.4959789×10^{11} m
mass of the moon	7.349×10^{22} kg
mean radius of the moon	1.7371×10^{6} m
mean Earth-Moon distance (center to center)	3.844×10^{8} m
mass of Earth	5.9723×10^{24} kg
mean radius of Earth	6.371×10^{6} m

Particle constants:

mass of electron	9.11×10^{-31} kg
mass of proton	1.673×10^{-27} kg
mass of neutron	1.675×10^{-27} kg
charge of electron	-1.60218×10^{-19}C
charge of proton	$+1.60218 \times 10^{-19}$C

Physical constants:

gravitational force constant (G)	6.67430×10^{-11} N $\cdot$ m^2/kg^2
gravitational force constant (G)	6.67408×10^{-11} N $\cdot$ m^2/kg^2
electric force constant (k)	8.98755×10^{9} N m^2/C^2
magnetic permeability (μ)	$4\pi \times 10^{-7}$ N/A^2
speed of light in vacuum (c)	2.998×10^{8} m/s
Avogadro's number (A)	$6.02214199 \times 10^{23}$ molecules/mole

6. Units

Letter abbreviations are used for the units of measurement. For example, "m" stands for "meters" and "s" stands for "seconds." Unit abbreviations are written in non-italic font to distinguish them from variable abbreviations, which are written in italic font (see section 4) for the list of variable abbreviations).

6.1 International System of Units (SI)

The scientific community[i] uses the International System of Units, sometimes referred to as SI (for Système International) or the metric system.[ii]

Five basic physical quantities and their SI units:

In the SI system, there are only a couple of basic physical quantities and each one is assigned a unique unit. The basic physical quantities are length, mass, time, temperature and current. In SI, the associated units are meters, kilograms, seconds, kelvin and ampere. By convention, the unit names are lowercase, even if named after a person.

Abbreviation	Name	Quantity
kg	kilograms	mass
m	meters	length
s	seconds	time
K	kelvin	temperature
A	ampere	current

Derived units:

[i]For more information, see the NIST publication *Guide for the Use of the Internal System of Units (SI)* by Barry N. Taylor [available via `https://physics.nist.gov/cuu/pdf/sp811.pdf`.

[ii]Technically, SI is just a particular type of metric system, but most people treat the two as meaning the same thing.

Units for all other quantities are derived in terms of these basic units (for this reason, this is sometimes called the MKS system – M for meters, K for kilogram and S for seconds). Examples of such derived units are as follows:

Abbreviation	**Name**	**Quantity**
Ω	ohm (V/A)	resistance or impedance
C	coulomb (A·s)	charge
°C	degrees Celsius[iii] (K - 273.15)	temperature
eV	electron-volt (1.6×10^{-19} J)	energy
F	farad (C/V or $\Omega^{-1}\cdot s^{-1}$)	capacitance
H	henry ($\Omega\cdot$s)	inductance
Hz	hertz (cycles/s)	frequency
J	joules ($kg\cdot m^2\cdot s^{-2}$)	energy
K	kelvin	temperature
N	newton ($kg\cdot m\cdot s^{-2}$)	force
Pa	pascal (N/m^2)	pressure
rad	radians (m/m)	angle
T	tesla ($N\cdot A^{-1}\cdot m^{-1}$)	magnetic field
V	volt (J/C)	electric potential
W	watt (J/s)	power
Wb	weber ($T\cdot m^2$)	magnetic flux

6.2 Other units

Abbreviation	**Name**	**Quantity**	**Value (in SI units)**
°	degree	angle	$\pi/180$ rad
bar	bar	pressure	10^5 Pa
ft	feet	length	12 in (0.3048 m)
h	hour	time	60 min (3600 s)
in	inches	length	2.54 cm
L	liter	volume	1000 cm^3
lb	pounds	mass	0.45359237 kg
mi	miles	length	5,280 ft (1609.344 m)
min	minutes	time	60 s
ua	astronomical unit	length	$\sim 1.496 \times 10^{11}$ m

6.2.1 Metric prefixes

Prefixes[iv] are then used to represent factors of ten for each unit. For example, the "kilo" represents 10^3. Consequently, a kilometer is equivalent to a thousand meters and a kilogram is equivalent to a thousand grams.

Abbreviation	**Name**	**Quantity**	**Abbreviation**	**Name**	**Quantity**
Y	Yotta-	10^{24}	y	yocto-	10^{-24}
Z	Zetta-	10^{21}	z	zepto-	10^{-21}
E	Exa-	10^{18}	a	atto-	10^{-18}
P	Peta-	10^{15}	f	femto-	10^{-15}
T	Tera-	10^{12}	p	pico-	10^{-12}
G	Giga-	10^{9}	n	nano-	10^{-9}
M	Mega-	10^{6}	μ	micro-	10^{-6}
k	kilo-	10^{3}	m	milli-	10^{-3}
h	hecto-	10^{2}	c	centi-	10^{-2}
da	deka-	10^{1}	d	deci-	10^{-1}

Some notes:

- You don't need to memorize them (after all, there is always the table) but you should become familiar with the most commonly used prefixes, which are Mega-, kilo-, centi-, milli- and micro-.
- The abbreviation for "micro" is the only one listed that is a Greek letter. It is the Greek letter "mu" (some people mistakenly think it is the Roman letter *u*).
- All of the abbreviations are one letter except for deca-.
- Many people get centi and milli confused. If it helps, consider that the "centi-" prefix comes from the same root as "century" (100 years) and the "milli-" prefix comes from the same root as "mile" (1000 paces) and "millennium" (1000 years).

6.3 Unit conversion

After all of the work involved in getting an answer to problems, one's work is not necessarily finished. This is because the answer needs to be interpreted

[iv] http://physics.nist.gov/cuu/Units/prefixes.html

and, in some cases, revised so that the audience can interpret the answer appropriately.

For example, suppose you multiply (5 ft/s) by 10 h. The result is 5 ft·h/s. This is a distance but it isn't in units that anyone is familiar with. Consequently, you cannot leave the answer as "5 ft·h/s." You must convert it to a unit that is familiar.

The problem is that the answer contains two different units of time (i.e. hours and seconds). To simplify, convert one of the time units so that the two units are the same.

This can be done by simply replacing one of the units with its equivalent. For example, since 1 h = 3600 s, replace "h" with "3600 s" to get

$$5 \text{ ft} \frac{3600 \text{ s}}{\text{s}}$$

which simplifies to 18,000 ft.

Sometimes the result is still hard to interpret because the number is too large or too small. For example, is the distance from class to where you live larger than 12,350,000 inches? Most people cannot tell without converting the number to a more reasonable unit.

In this case, since

$$1 \text{ mi} = 5,280 \text{ ft}$$

we could convert 18,000 ft to miles by replacing "ft" with "(1/5280) mi". This gives

$$\frac{18,000}{5,280} \text{ mi}$$

or 3.41 mi.

Even this may not be satisfactory, since it is not in SI. To convert it to SI, replace "mi" with "1609.344 m" to get

$$3.41(1609.344 \text{ m})$$

or 5490 m.

This should then be converted to a more reasonable number by using the metric prefixes, i.e., 5.49 km.

7. Scientific notation

A popular way of getting rid of big or small numbers is to use scientific notation. In scientific notation, the factor of ten is written explicitly rather than in a metric prefix. For example, instead of writing 5490 m as 5.49 m, we could instead write it as 5.49×10^3 m.

Such notation also makes it much easier to multiply or divide very big and very small numbers. For example, what is 12,350,000,000,000 s times 0.0000000000350 s?

Not only does it take a long time to write down all of the zeros associated with very big or very small numbers, but it is hard to readily see how many zeros are present without carefully counting them. These two problems are addressed via a technique called **scientific notation**.

Scientific notation essentially takes all of the zeros and wraps them up in the form of 10^n, where n is some integer. For example, since ten million (7 zeros) is the same as $10 \times 10 \times 10 \times 10 \times 10 \times 10 \times 10$ (ten multiplied by itself 7 times), we can write ten million as 1×10^7. Since a number like 55,000,000 is just 5.5 times ten million, we can write 55,000,000 as 5.5×10^7.

Likewise, a number can have lots of zeros if it is really small. For example, since one-ten-millionth (0.0000001) is the same as 1/10/10/10/10/10/10/10 (one divided by ten 7 times), we can write one-ten-millionth as 1×10^{-7}. Similarly, since a number like 0.00000055 is just 5.5 times one-ten-millionth, we can write 0.00000055 as 5.5×10^{-7}.

It is easier to multiply and divide very large or small numbers when they written in scientific notation because we can quickly identify how many zeros are present. For example, suppose you had to multiply 5.5×10^7 by 5×10^2. We can simply perform this multiplication in two steps. First, we multiply 5.5 by 5 to get 27.5. Then, we multiply 10^7 by 10^2 to get 10^9. The final result is 27.5×10^9 or 2.75×10^{10}.

Note: By convention, we write a number like 12,300 as 1.23×10^4 rather than 123×10^2. For example, we'd write 2.75×10^{10} instead of 27.5×10^9.

8. Significant digits

When we identify a number, we want to be as precise as possible. For example, if we are pretty sure a distance is 2.3 m, we wouldn't want to round it off and say it is 2 m. On the other hand, we don't want to create the impression that we know the distance to more precision than we actually do. That is, if we are pretty sure the distance is 2.3 m but aren't sure whether it is 2.32 m or 2.28 m or 2.30 m, then we wouldn't want to say the distance was definitely 2.325617 m!

So, we only write down the digits that we are pretty sure of (some people add one additional digit that has been estimated). We call these digits the *significant digits*. All others we have no confidence in and thus should not be written.

Note: If you are pretty sure a number is something like 5.50 and not 5.51 or 5.52, then you should go ahead and include the zero.

8.1 A method of predicting significant digits

To determine the number of digits in an answer that are significant, the best way is to vary numbers as in the following example.

Example 8.1: What is $2.5 \times 3.8765/87.20$?

Answer 8.1: If we plug these numbers into a calculator, we get 0.111138 as the answer. However, not all of these digits are significant. To find which ones are significant, we will go through the process outlined above.

First we assume that all the digits given in the numbers are significant. Consequently, the number "2.5" could be anywhere between 2.45 and 2.55. The number "3.8765" could be anywhere between 3.87645 and 3.87655, The number "87.20" could be anywhere between 87.195 and 87.205.

There are eight combinations we can try.

$$\begin{aligned}
\frac{2.55\times3.87655}{87.205} &= 0.113356\\
\frac{2.55\times3.87655}{87.195} &= 0.113369\\
\frac{2.55\times3.87645}{87.205} &= 0.113353\\
\frac{2.55\times3.87645}{87.195} &= 0.113366\\
\frac{2.45\times3.87655}{87.205} &= 0.108911\\
\frac{2.45\times3.87655}{87.195} &= 0.108923\\
\frac{2.45\times3.87645}{87.205} &= 0.108908\\
\frac{2.45\times3.87645}{87.195} &= 0.108920
\end{aligned}$$

As you can see, the answer varies from 0.108908 to 0.113369 (or 1.08908×10^{-1} to 1.13369×10^{-1}). All of these numbers equal 0.11 when rounded. We would say the answer is 0.11 or 1.1×10^{-1}. Only two digits are significant.

This can get quite tedious, especially if we have many numbers involved. If we were making measurements in a lab, the method used above might be needed. For our purposes, however, we only need a "quick-and-dirty" method that gives the approximate number of significant digits.

This "quick-and-dirty" method has a different rule depending on whether you are adding/subtracting or multiplying/dividing:

1. If you are multiplying or dividing, the number of significant digits in the product or quotient is approximately the number of significant digits in the *least* precise number used in the calculation. For example, in example 8.1, the least precise number was 2.5, which had only two significant digits.
2. When adding or subtracting, each place is significant only if that place was significant in *every* number used in the calculation. For example, if we add $2.5 + 3.8265 + 97.2$, only the tenths place is significant. If we instead write "2.5" as "2.5????" to indicate our uncertainly over the unwritten digits, we can see this clearly when the sum is written as follows.

$$\begin{array}{rr}
 & 2.5????\\
+ & 3.8265?\\
+ & 97.2????\\
\hline
 & 103.5????
\end{array}$$

Example 8.2: What is $1.234567 + 130.0 + 3.41$?

Answer 8.2: If we just put the numbers in a calculator, we get 134.644567. However, the hundredths place is not significant in "130.0"; only the tenths place is significant. Consequently, only the tenths place is significant in the answer. We would write 134.6 (or 1.346×10^2 as the answer).

Example 8.3: What is $2.5 \times 3.8765/13.40$?

Answer 8.3: If we plug these numbers into a calculator, we get 0.723228 as the answer. However, the number "2.5" only has two significant digits. Consequently, the answer should only have two significant digits. We would write 0.72 (or 7.2×10^{-1}).

It is important to keep in mind that this method only gives the approximate number of significant digits in the answer. For example, if we used the tedious method in example 8.3, we would find that the answer could vary from 0.70849 to 0.737977. The answer doesn't really have two significant digits. Instead of writing the answer as 0.72 (as predicted by the short-cut method), it would be better to write the answer as 0.723 ± 0.015. The short-cut method is okay as long as we realize that it only predicts an approximate precision of the answer. If we really need to know the precision of the answer (as when we are doing measurements in the lab), a better method is needed.

Example 8.4: (a) What is 2.513×3.8765?
(b) What is 4.522×1.2486?
(c) What is the product of (a) and (b)?

Answer 8.4: (a) Plugging into a calculator, one gets 9.7416445. Since the least precise number used in the calculation has 4 significant digits, the answer should be written as 9.742.
(b) Plugging into a calculator, one gets 5.6461692. Since the least precise number used in the calculation has 4 significant digits, the answer should be written as 5.646.
(c) The product of 9.742 and 5.646 is 55.003332. With four significant digits, the answer is 55.00 (or 5.500×10^1).

A common mistake is to round the intermediate steps too much. For example, if we instead round the answer in part (a) to 9.7 and the answer in part (b) to 5.6, our answer in part (c) is 54.32. This is not correct. Even if we round to two significant digits, we get 54, which is still not correct.

As a rule, do not round at all until the end.

Example 8.5: (a) What is the product of 4.352 and 2.3?
(b) What is the product of 5 and 2.0123?

Answer 8.5: (a) Plugging into a calculator, one gets 10.0096. Since the least precise number is "2.3", which has two significant digits, the answer must be 10.
(b) Plugging into a calculator, one gets 10.0615. Since the least precise number is "5", which has one significant digit, the answer must be 10.

Notice that the answer is 10 for both part (a) and part (b) of the example. However, in part (a), the number is supposed to have two significant digits while in part (b) the number is only supposed to have one. How can we write the number 10 so that our audience knows how many digits are significant?

The answer is to write it in scientific notation. The answer to part (a) would be 1.0×10^1 whereas the answer to part (b) would be 1×10^1.

9. Mathematics

9.1 Average vs. midrange

To get a sense of what an average means, consider the following example.

> Suppose you buy five apples at five different stores. The prices are \$1.00, \$0.95, \$0.75, \$0.80, and \$0.80. To find the average price per apple, add up the total amount you spent, and divide by the number of apples. You should get \$0.86 per apple as the answer.

✍ There are many different schemes for computing an average. This particular type of average is called the **mean**. This is the type of average that is relevant to the equations of motion and so it is the one we will use.

This works for more than just apples. You are probably familiar with figuring out your average grade, so consider the following example:

> Suppose you take five exams during a semester and receive the following scores: 100, 95, 75, 80 and 80. To find the average exam score, add up all five values and divide the total by the number of exams. You should get an average score of 86.

Notice that in both cases the average is somewhere between the minimum and maximum values (e.g., 86 is between 75 and 100) but is not equal to the **midrange** value, which is the value exactly midway between the minimum and maximum values. In this case, the minimum is 75 and the maximum is 100 so the midrange value would be 87.5.

Is the average ever equal to the midrange value?

Yes, it can be.

Consider, for example, the following series of test scores: 60, 70, 80, 90, 100. In this case, the average score would be 80, and that also happens to be the midrange value.

This particular example happens to be an **arithmetic progression**, which is a sequence of numbers where each number differs from the preceding number by a fixed difference. In this case, each test score differs from the previous test score by 10.

Another example of an arithmetic progression is 5, 3, 1, −1, −3. In this case, the progression is toward negative numbers, so this is called a negative arithmetic progression (as opposed to a positive arithmetic progression). However, either way, the maximum and minimum values are provided by the *first* and *last* values (or *last* and *first* values).

Example 9.1: Suppose you had the following series of numbers: 20, 19, 18, 17, 10. Is the average value also equal to the midrange value? Why or why not?

Answer 9.1: No. The numbers do not uniformly decrease. The midrange value would be 15 (since the minimum is 10 and the maximum is 20). The average is not 15 because there are more numbers greater than 15 than less than 15, which pushes the average to be higher than 15.

Now that you've gotten a sense of what an average is, let's apply this to velocity (see volume I).

Suppose an object starts with a velocity of 5 m/s in the positive direction and steadily speeds up, gaining 2 m/s in the positive direction every second for five seconds. So it starts with a velocity of +5 m/s and at each succeeding second is moving at +7 m/s then +9 m/s then +11 m/s then +13 m/s and finally +15 m/s five seconds later.

Since this is an arithmetic progression, the average value is equal to the midrange value. In this case, the minimum value is +5 m/s and the maximum value is +15 m/s, so the midrange (and thus the average) is +10 m/s.

For the average velocity to be the midrange value, the acceleration must be constant. In the case just considered, for example, the object's acceleration is 2 m/s^2 in the positive direction for the entire five seconds.

Example 9.2: Suppose an object is traveling at a constant velocity of +10 m/s for four seconds and then suddenly slows down to a stop, taking one second to do so. Is the average velocity during the five seconds equal to the midrange value (+5 m/s)? Why or why not?

Answer 9.2: No. The velocity values do not uniformly decrease. The midrange value would be +5 m/s (since the minimum is 0 and the maximum is +10 m/s). The average is not +5 m/s because the object spent more time with the speeds greater than 5 m/s than speeds less than 5 m/s.

As you can see in the example, the midrange is not always equal to the average. However, for situations where the acceleration is constant, the average velocity is indeed equal to the midrange value. This is nice, because most of the situations we'll examine for the time being are those where the forces are constant, and according to the law of force and motion the acceleration is constant when the net force is constant.

9.2 Slope

The slope of a line on a graph is defined as the rise over run, where the "rise" is the change in the vertical coordinate and the "run" is the in the horizontal coordinate.

Typically in math we use y for the vertical coordinate and x for the horizontal coordinate. Consequently, the general expression for the slope is as follows:[i]

$$m = \frac{\Delta y}{\Delta x} \tag{9.1}$$

For a straight line, there is just one slope value. This means the $\Delta y/\Delta x$ ratio is the same regardless of what Δx interval we choose. A larger Δx value just means that Δy will be larger as well.

It is for this reason that the equation of a straight line can be written with a single slope value m. Typically, it is written as follows:

$$y = mx + b$$

[i]According to the "Earliest Uses of Symbols from Geometry" web page at http://jeff560.tripod.com/geometry.html, the earliest known use of the letter m to indicate slope is from 1757, as part of a general equation for a line (written as $y = mx + n$).

This equation can be derived by assuming a constant slope and then rearranging the definition of the slope.

$$m = \frac{\Delta y}{\Delta x}$$

Since the slope is the same regardless of which two points we choose on the line, we'll make things easy and choose one of the points to be where $x = 0$. We'll use y_0 to indicate the y value when x is zero. The other point can be at any x and y value pair, so we'll use x and y to indicate those values. That means that Δy can be written as $y - y_0$, and Δx can be written as $x - 0$ or just x. We thus get:

$$m = \frac{y - y_0}{x}$$

Multiply both sides by x and then add y_0 to both sides to get:

$$mx + y_0 = y$$

Flipping sides, we get:

$$y = mx + y_0$$

By convention, we use b instead of y_0, giving us the traditional format for the equation of a straight line (constant slope):[ii]

$$y = mx + b$$

What if the line is curved and not straight?

If the line is curved and not straight then there is not a single value of the slope and so we cannot use this equation. However, we can still figure out the average slope of various portions of the curve by apply the definition of slope.

For example, consider the graph shown in Figure 9.1.

What is the slope of the curve?

The slope depends on what portion you examine.

If you consider the entire portion shown, from the left end at (0,0) to the right end at (3,15), the total rise is 15 and the total run is 3. From the

[ii]The letter b is used because it seems the general equation for a line may have, at one time, been written as $(x/a) + (y/b) = 1$, which can be re-arranged to get $y = (-b/a)x + b$.

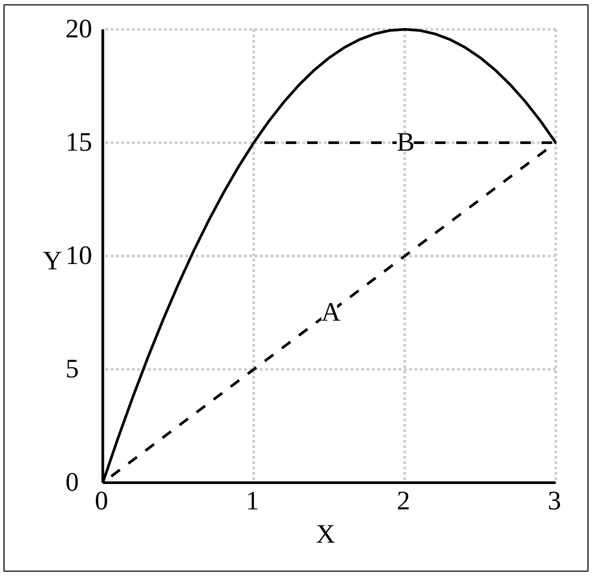


Figure 9.1: A sample graph. Dashed lines are explained in the text.

definition of slope (equation 9.1), we get a slope equal to 15 divided by 3, which equals 5.

Since there is no single slope value for this curve, you might wonder what the value of 5 corresponds to. It turns out that this is the slope of dashed line A in Figure 9.1, which is the straight line drawn from the left end at (0,0) to the right end at (3,15).

The actual curve has portions where the slope is greater than 5 and other portions where the slope is less than 5.[iii] So, using the end points actually gives you the *average* slope of the curve between those two points.

For a curved line, the definition of slope (equation 9.1) gives the *average* slope over the portion selected.

As an example, suppose we want to know the average slope between the points (1,15) and (3,15). We might expect that the slope is close to zero, since the curve rises between $x = 1$ and $x = 2$ (meaning the slope there is greater than zero) and falls between $x = 2$ and $x = 3$ (meaning the slope

[iii] In this case, it turns out that for points to the left of around $x = 1.5$, the slope of the curve is actually greater than the slope of that straight line. Conversely, for points to the right of around $x = 1.5$, the slope of the curve is less than the slope of that straight line.

there is less than zero). To find out, we take the total rise (zero, since then y value on the right is the same as the y value on the left) divided by the total run (2) to get a slope equal to zero. This is consistent with what we expected.

9.3 Scaling

The law of gravity is given in volumes I and II:

$$|\vec{F}_{\mathrm{g}}| = G\frac{m_1 m_2}{r^2}$$

When dealing with an equation that seems as complicated as this, it helps to use the idea of proportions, which is discussed in volume I. Proportions can give us insight into a relationship without needing to do a lot of math.

For example, we know that the gravitational force on two interacting objects is less if the two objects are farther apart. Let's use our knowledge of proportions to determine *how much* less.

For our example, let's suppose the distance is increased by a factor of n so that the distance increases from a distance r to a distance equal to nr. What happens to the force?

To answer this, we first notice that the r is in the denominator. That means the force must decrease. Next we notice that the r is squared. When we replace r by nr, the denominator goes from r^2 to n^2r^2. Since the denominator increases by a factor of n^2, that means the force decreases by a factor of n^2.

For example, if r is doubled then $|\vec{F}_{\mathrm{g}}|$ is quartered (a factor of one-fourth).

Notice how we can get this result without doing a lot of arithmetic. For example, we didn't need to use the numerical value for G or know the mass values or even the actual value of r. All we really needed to know was the ratio between the new and old values of r (e.g., the value is doubled).

✍ One could instead plug in values of G, r and the masses and then see what happens when the value of r is doubled. However, not only would that be more work but it would only tell us what the value does for that particular case.

The process we used above is known as **scaling**, which is a really valuable tool, not just for physics, and shows how understanding proportions can make your life easier.

As an additional example, let's consider the problem described in volume I where it states that the magnitude of the gravitational force on an object (due to Earth) is only 1% smaller at an altitude of 3.19×10^4 m above Earth than what it is when the object is on Earth's surface.

To see how I got this number, we first need to recognize that if r increases by a factor of n then $|\vec{F}_\text{g}|$ must decrease by a factor of n^2. Going backwards, if $|\vec{F}_\text{g}|$ decreases by a factor of n then r must increase by a factor of $\sqrt{n}$.

So, if we want the magnitude of the gravitational force to decrease by 1% (i.e., decrease to 0.99 of its original value) then r must increase by the square root of that factor (i.e., the square root of 0.99).

The square root of 0.99 is 0.995. Consequently, r must increase to 1/0.995 of its original value. Since this ratio equals 1.005, that means it must increase by 0.5%. Half of one percent of 6.371×10^6 m is 3.19×10^4 m, as asserted before.

We can also use scaling to show that the gravitational force is significant only if at least one of the objects is very massive.[iv] Recall that it is because of this that we can ignore the gravitational force between ordinary objects (like you and a ball) because the mass of these objects are simply too small.

❢ Unless one (or both) of the objects is very massive, the gravitational force will likely be insignificant.

Ordinary objects are much closer to each other than the size of Earth. Wouldn't that make the magnitude of the force between ordinary objects even greater?

This is where scaling can help us.

We know that the magnitude of the gravitational force will be larger when the distance between objects is smaller. An object on Earth's surface is 6.37×10^6 m from Earth's center (an amount equal to Earth's radius). Two ordinary objects, like a person and a ball, are much closer to each other than that.

While we could answer this question by calculating the gravitational force between two ordinary objects, using typical masses and a typical separation

[iv]Like the moon, Earth or sun.

distance, and then comparing that to the gravitational force on one of the objects by Earth, we'll instead using scaling.[v]

For example, let's suppose we scale Earth's radius down by a factor of X (so that X is some small fraction). From the law of gravity, we know that decreasing the radius by a factor of X means that the denominator will decrease by a factor of X^2.

However, that assumes the mass will remain the same. If Earth's radius decreases, so will its mass, and that will affect the gravitational force.

Since the volume is proportional to the cube of the radius, decreasing the radius by a factor of X means that the volume decreases by a factor of X^3. Since the mass is proportional to the volume (for the same density), that means the mass must likewise decrease by a factor of X^3.

From the law of gravity, we can see that decreasing the mass by a factor of X^3 means that the numerator will decrease by a factor of X^3.

The X^2 in the denominator only partially cancels the X^3 in the numerator. The end result is a factor of X^3/X^2, which is equal to X.

So overall, when the object gets smaller (and closer) by a factor of X, the gravitational force gets *smaller* by a factor of X (assuming the density stays the same). Consider that a rock 6.37 cm in diameter is one hundred million times smaller than Earth. That means the gravitational force due to the rock will have a magnitude one hundred million times smaller than the gravitational force due to Earth, even though the rock is one hundred million times closer.

This is why we tend to ignore the gravitational force except in those situations where one or both of the objects are very massive (like the sun or Earth). In other words, we will assume that everyday objects (like you, me and the ball) don't exert a measurable force on each unless they are actually in contact, in which case the force is due to contact rather than gravity.

[v] To do it with numbers, we could use 70 kg for the person and 1 kg for the ball with a separation distance of 1 m. Plugging those numbers into the universal law of gravitation, one gets a gravitational force of magnitude 4.7×10^{-9} N. A ball would have to have a mass of about 200 million kg in order for the force to even be one newton, still a small fraction of the gravitational force on it due to Earth.

✍ One could also go backward and infer the density of a planet or moon by using our knowledge of its gravity (which can be obtained by putting something in orbit around it; see volume I). Indeed, this is how NASA can report on the density of the moon (and rule out otherwise-plausible hypotheses about how the moon and Earth formed) even though no one has seen the interior of the moon.

9.4 Vectors vs. Scalars

We use the word **vector** to describe variables that have both a magnitude and a direction. Examples of vectors include displacement and velocity.

A variable that doesn't have a direction, like time, is called a **scalar**. A scalar is represented by only a single number that does not depend on whether we are in one or two dimensions. For example, temperature is always represented by a single number (and unit).

What is the difference between the magnitude and a scalar?

Since the magnitude of a vector does not include the direction, technically the magnitude of a vector is a scalar quantity.

Are magnitudes always positive? What about scalars?

The magnitude is always positive.[vi] So, if we have two velocities, one that is 20 m/s toward the east and another that is 20 m/s toward the west, the magnitude of both is 20 m/s, even though they point in different directions.

Indeed, if we rewrite the second one to be −20 m/s toward the east (since that is the same as +20 m/s toward the west), its magnitude would still be 20 m/s.

✍ Some scalar quantities can have negative values. For example, if a temperature of zero corresponds to when water freezes then we can have negative temperatures.

Are component values vectors or scalars?

[vi] This does not include situations where the word magnitude is used for some other purpose. For example, stellar magnitude is defined in such a way that it can be a negative number.

That depends on what you mean by "component values." As we already know, a vector in two dimensions can be written as two values, one for each component direction. For example, for a displacement three blocks toward the east and two blocks toward the north, the eastward component value is three blocks. That is a scalar. The eastward component direction is, well, eastward.

✍ It is traditional to refer to the component direction as a vector with value of one (no units). In that case, it is called a **unit vector**. An example of a unit vector would be $\hat{x}$.

9.5 Finding Vector Direction

This text examines two-dimensional motion in terms of two perpendicular components. However, sometimes we want to know the magnitude and direction of a result, not just the component values.

In volume I, the Pythagorean Theorem

$$A^2 = A_x^2 + A_y^2$$

was introduced as a way of finding the magnitude (A) of a vector from the two component values (A_x and A_y). However, I never mentioned how to find the direction of the vector.

It turns out that we need to use the **inverse** trigonometric functions to find the direction of a vector from its components.

In a way, the method almost seems like cheating. For, just as the trigonometric functions "take in" an angle (direction) and "spit out" the relative fractions of the components, there are "inverse" trigonometric functions that do the reverse: they "take in" the relative fractions of the components and "spit out" an angle (direction).

What are these inverse trigonometric functions called?

They are called the **inverse cosine**, the **inverse sine** and the **inverse tangent**.[vii]

Which inverse trigonometric function do I use?

[vii] What can I say? Mathematicians aren't very imaginative.

Most people use the inverse tangent function, so that is the one I will show you.

To use the inverse tangent function, you first have to recognize that the tangent function, like the sine and cosine functions, represents a ratio. For the tangent function, the ratio is A_y divided by A_x:

$$\tan\theta = \frac{A_y}{A_x}$$

where:

θ represents the angle between the vector and the $\hat{x}$ direction
A_x represents the component value in the $\hat{x}$ direction, and
A_y represents the component value in the $\hat{y}$ direction.

✍ I'll use θ to indicate an angle. It is the Greek letter "theta." As discussed in volume I, we use Greek letters for angular quantities.

Like the sine and cosine functions, the value of the tangent function depends upon the angle (indicated by θ).[viii] For a particular angle, there is a particular ratio of A_y to A_x. In other words, given the angle θ, the tangent function provides the fraction A_y/A_x.

In our case, we know the values of A_y and A_x (and thus the ratio as well) but we don't know the angle θ. We need to go "backwards" (from A_y/A_x to θ, rather than from θ to A_y/A_x).

The backward process, which we call the *inverse* tangent, is indicated as $\tan^{-1}$ in mathematical expressions.[ix] So, mathematically, we have

$$\theta = \tan^{-1}\left(\frac{A_y}{A_x}\right) \tag{9.2}$$

where θ is the angle between the vector and the $\hat{x}$ direction.

The notation is a little confusing, so I'd like to make a comment about that.

[viii] In the expression, it looks like I am multiplying tan by the angle θ but I'm not. The two together ($\tan\theta$) means to find the tangent value that corresponds to the angle θ. On a calculator, one might type in the angle and then hit the tangent button.

[ix] Likewise, the inverse sine is indicated as "$\sin^{-1}$" and the inverse cosine is indicated as "$\cos^{-1}$."

The convention[x] of indicating the inverse trigonometric functions via a raised "−1" can lead to confusion if you aren't careful. Up until now, you've probably only encountered the raised "−1" to mean "one divided by the number." For example, x^{-1} means $1/x$. For the inverse trigonometric functions, like $\cos\theta$, the raised "−1" does not mean we divide 1 by the function (e.g., $1/\cos\theta$). Rather, it means a separate function that does the reverse process.[xi]

Let's try it out on a calculator now.

The fraction A_y/A_x is unitless, so it doesn't matter what unit you use for A_y and A_x, as long as they are the same when you calculate the ratio. That way, the ratio is unitless (i.e., the units will cancel).

For example, if A_y and A_x are the same, then the ratio will be 1 (with no units).

So, take the inverse tangent of 1 on your calculator.

✍ On some calculators, the inverse function is obtained by pressing a "2nd" or "INV" button before pressing the trigonometric buttons.

The answer you get is an angle (indicated by θ in equations).

There are several units for angle (e.g., degrees and radians). Which unit your calculator is using depends on how your calculator is set up. If your calculator is set up to provide the answer in degrees, you should get 45 degrees. If your calculator is set up to provide the answer in radians, you should get 0.785 radians. Try it now and make sure you can properly interpret the number your calculator is giving you.

WHY IS 45 DEGREES EQUAL TO THE INVERSE TANGENT OF ONE?

If the ratio is one, that means the two components are equal. The two components are equal only if the direction is exactly midway between the two component directions. That means the angle is 45 degrees.

---

[x] This is the mathematical convention. You might find a different convention in other contexts. For example, Excel uses "acos", "asin" and "atan" to represent the three inverse functions.

[xi] The word "inverse" is used in the same way when we say that $1/x$ is the inverse of x. Consider, for example, if you want to convert inches into centimeters. To do so, we multiply the value (in inches) by 2.54 cm/in. If we let x equal 2.54 cm/in, this is saying that we can convert a value from inches into centimeters by multiplying by x. To do the reverse and convert a value from centimeters into inches, we can multiply by $1/x$. In other words, multiplying by $1/x$ does the inverse function that x did.

To show you how to use the inverse tangent, let's look at displacement of 10 m in a direction 30 degrees east of north. Using the sine and cosine of 30 degrees then multiplying by the 10 m, we get that the two components are 5 m and 8.66 m.

Let's suppose we knew neither the magnitude nor the direction — only the values of the components (5 m and 8.66 m). Using equation 9.2, we can get the angle with the inverse tangent function:

$$\begin{aligned} \theta &= \tan^{-1}(\Delta s_y/\Delta s_x) \\ &= \tan^{-1}\Big((8.66 \text{ m})/(5 \text{ m})\Big) \\ &= \tan^{-1}(1.732) \end{aligned}$$

which gives a value of 60 degrees.

ISN'T THE ANGLE SUPPOSED TO BE 30 DEGREES?

To understand why the inverse tangent is 60 degrees in this example, we have to recognize that we can form the ratio two ways: $\Delta s_y/\Delta s_x$ or $\Delta s_x/\Delta s_y$. One of those ways will give an angle of 60 degrees and the other way will give an angle of 30 degrees.

The angle that is provided is the angle from the component direction used in the *denominator*. Since the x component was used in the denominator, the angle in this case is 60 degrees (i.e., 60 degrees from the $\hat{x}$ direction).

If we had instead used the fraction $\Delta s_x/\Delta s_y$ (i.e., 5/8.66), the inverse tangent would give a value of 30 degrees.

In general, always interpret the angle that the calculator gives you. If you forget that the angle is given relative to the component direction used in the denominator, you can always tell which angle is which simply by comparing the two component values. Which one is bigger? The direction will be closer to that component direction.

✍ Another reason for drawing a picture is that it should be pretty clear from the values of the two components whether the calculator has given you the angle in degrees or not.

Problems

Problem 9.1: For which of the following series of numbers is the average equal

to the midrange value? If so, what is the average? If not, why not?
(a) 1, 2, 3, 4, 5
(b) 5, 4, 3, 2, 1
(c) 5, 5, 5, 2, 1
(d) 1, 4, 7, 10, 13
(e) 1, 1, 1, 1, 5

Problem 9.2: For which of the following situations is the average velocity equal to the midrange value? If equal, identify the average velocity. If not equal, explain why it is not equal.
(a) Initial velocity of +1 m/s and accelerating uniformly at +1 m/s^2 for 4 s
(b) Initial velocity of +5 m/s and accelerating uniformly at −1 m/s^2 for 4 s
(c) Initial velocity of +5 m/s and remaining at that constant velocity for 2 s before accelerating uniformly at −2 m/s^2 for an additional 2 s
(d) Initial velocity of +1 m/s and accelerating uniformly at +3 m/s^2 for 4 s
(e) Constant velocity of +1 m/s for 4 s

Problem 9.3: In math class, the slope or steepness of a line on a graph is equal to the vertical change in the line divided by the horizontal change in the line. Is this a law, a definition or a derived relationship?

Problem 9.4: (a) In Figure 9.1, what is the average slope between $x = 2$ and $x = 3$?
(b) In Figure 9.1, what is the slope right at $x = 2$: positive, negative or zero?

10. Using electronic meters

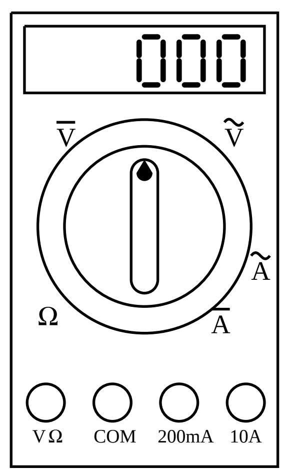
Most likely, you will be using a meter that can work as both an ammeter and a voltmeter. Such a meter is called, naturally enough, a **multimeter**.

10.1 Ammeters

To measure current with a multimeter, you need to "select" the **ammeter**.

There are lots of different types of multimeters, so there is not one set of instructions that will work on all of them. Thus, you need to look at whatever multimeter you encounter and determine how to select the ammeter.

However, there are some features that are common for all multimeters and since you will likely be using them a lot in this course, I'd like to go through some of the general features of a multimeter. In particular, I'd like to point out the features you need to be aware of in order to use it as an ammeter.



A simple schematic of a typical multimeter is shown to the right.

One uses the knob in the center to choose the type of reading you'd like to make. To measure steady current, rotate the knob so that it points to $\overline{\text{A}}$. To measure the RMS value (for AC current) select the "A" with the wavy line on top, $\tilde{\text{A}}$.

At the bottom of the meter are four sockets (indicated by small circles in the figure). The sockets are designed to accept a type of test lead called a "banana lead," which have ends that look a little like a banana.

The sockets are usually color-coded so that we know where the current goes in (red) and where the current comes out (black).

Because the meter can be used to measure lots of different things, there are several red input sockets, depending on what you want to measure.

On the other hand, there is just one black output socket, indicated as "−" or "COM", because that socket is common to all measurements.

> ✍ If the connections are reversed (between the input and the "COM" ports), the ammeter will read a negative current. This does not harm the ammeter. It just means that the current is flowing the reverse way.

To read current, you need to use either the "200 mA" or "10 A" input sockets.

Which input port you chose depends on *how much* current will be passing through the meter.

- The "10 A" scale measures current up to 10 A. You should always start with the 10 A scale, as it will work for all currents (up to 10 A). The disadvantage is that it does a poor job of measuring currents less than 200 mA.
- The "200 mA" scale measures current up to 200 mA. It does a better job of measuring small currents but to do so it requires a sensitivity that can be destroyed with too much current.
  For that reason, you should not use the 200 mA scale until you've first measured the current with the 10A scale and have verified that the current is not more than 200 mA.[i]

> ✍ Because the 200 mA scale is so sensitive, most multimeters use a **fuse** with the 200 mA scale that melts if more than 200 mA of current flows through the meter. This way, if too much current flows through the meter, the fuse melts, breaking the circuit and stopping the flow of current. Of course, that also prevents the meter from working until you replace the fuse.

If the meter starts to beep or just show or flash a "1" that means that the current is higher than the value indicated by the scale.[ii] You should disconnect the meter from the circuit, as it is possible that you are passing too much current through the meter.

---

[i] It can be easy to make a mistake and connect the ammeter in its own separate path rather than along the same path as the other element or elements through which the current is already passing. This will send more current through the meter than you expect.

[ii] Just because you use a particular input, like 200mA, does not mean that the meter display is set to show currents up to that value.

✍ Just turning off the meter is usually insufficient as current may continue to flow through the meter. Similarly, whenever you change the type of measurement the multimeter is making (e.g., from voltage to current), make sure that the multimeter is disconnected from the circuit (otherwise you are in danger of blowing the fuse if the connection and/or scale is not appropriate).

10.2 Voltmeters

One can use a **multimeter** to measure voltage as well as current. An example of a multimeter is shown on page 75. To measure steady voltage, you need to rotate the knob so that it points to $\overline{\mathrm{V}}$ (the "V" with a bar over it). To measure the RMS value (for AC voltage) select the "V" with the wavy line on top, $\widetilde{\mathrm{V}}$.

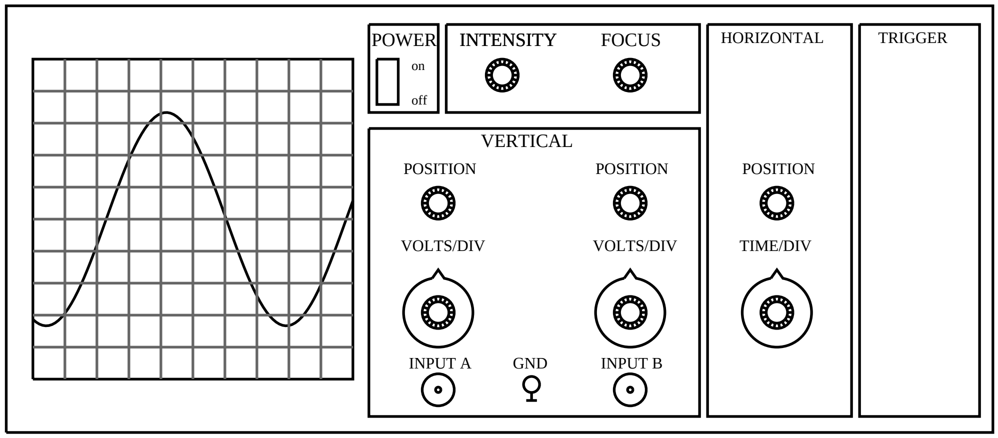
A voltmeter uses two sockets. The "+" socket will likely be labeled "V" (as in the figure on page 75). The "−" socket is usually labeled "COM" (since it is the common socket used for all of the various types of measurements). Flipping the connections doesn't hurt the meter. It just switches the reading from positive to negative (or visa-versa).

10.3 Oscilloscopes

❢ An oscilloscope measures voltage.

As mentioned above, one can use a multimeter (via the "AC" scale) to measure the AC voltage (via the RMS value). Another way is measure the AC voltage to use an instrument called an **oscilloscope**.

An oscilloscope is just a fancy voltmeter. An example of an oscilloscope is shown below.



At first glance, an oscilloscope looks very complicated because there are many controls on the face of the oscilloscope. Fortunately, the controls break down nicely into six areas. Of the six areas, only three are crucial to know in order to use the oscilloscope as a voltmeter:

1. The display area is the gridded area on the left. This is where the voltage signal is displayed. Unlike a multimeter, which displays just a number, an oscilloscope graphs the voltage as it varies in time. In this case, the figure shows a voltage that is oscillating sinusoidally (i.e., like a sine function).
2. The area entitled "VERTICAL" controls the vertical axis of the graph. The vertical axis represents the **voltage** of the source. This area is separated into two parts (A and B) in case the user wants to measure the voltage across two different objects (in which case there would be two plots in the graph area).
3. The area entitled "HORIZONTAL" controls the horizontal axis of the graph. The horizontal axis represents **time**.

The way you make a voltage measurement with the oscilloscope is similar to the way you make a voltage measurement with a multimeter. Remember that to measure the voltage across a battery with a *multimeter*, one connects the "COM" (or "−") port of the meter to one end of the battery and the "V" (or "+") port of the meter to the other end.

The *oscilloscope*, like the multimeter, also has two connections to measure the voltage but it differs from the multimeter in two ways.

One way it differs is that the oscilloscope can measure <u>two</u> voltage values at the same time. As mentioned before, the oscilloscope's "VERTICAL" section has two parts. Each part has a duplicate set of controls. Each set corresponds to a separate input (or "channel").

Why do you need two inputs?

The two inputs are provided in case you want to compare the voltage across one part of the circuit with the voltage across another part of the circuit.

Where are the "+" and "−" ports of the oscilloscope?

That is the other difference between the oscilloscope and a multimeter.

Rather than having two "banana" slots (one for "+" and one for "−"), there is a separate "+" port for each input and a common "ground" port.[iii] In the figure, the ground port is indicated as "GND" and the two input ports are indicated as "INPUT A" and "INPUT B".

So, to measure the voltage across a battery, we connect the input port to one end of the battery and the ground port to the other end of the battery.

How do we use the oscilloscope to measure the voltage across an element in a circuit?

The method is similar to how you would measure the voltage across a battery. Connect the input port to one end of the element and the ground port to the other side.

Can we use the oscilloscope to measure the current through the circuit also?

No. The oscilloscope measures voltage, not current.

However, we can still use it to measure current indirectly. In other words, we could measure the voltage across a resistor. Then, assuming we know the resistance of the resistor, we could then use $V = IR$ to calculate the current through the resistor.

How do you know what the voltage is if the oscilloscope doesn't give you a number?

To determine the voltage, you need to convert the graphical display to a numerical value. Voltage is measured vertically on the graphical display.

[iii] Actually, each input is "coaxial", which means there is an outer conductor wrapped around an inner conductor. This can be converted to the typical dual-input plugs via a "banana to coaxial" converter (or BNC clip). The BNC clip usually has a red port (for "+") and a black port (for "−").

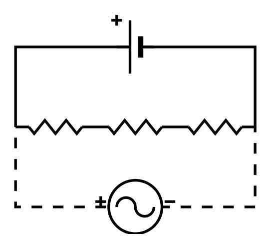
Each horizontal line on the display represents a certain amount of voltage. The voltage that corresponds to each horizontal line is given by the "VOLTS/DIV" scale (see figure). Changing the "VOLTS/DIV" scale does not change the voltage, it only changes how much each horizontal line represents.

For example, if in the figure the "VOLTS/DIV" scale was 1 V/division, then the signal being measured in the figure has an amplitude of about 3.3 V (i.e., 6.6 V peak-to-peak).

What benefit does graphing the voltage have?

The benefit of the graph (on an oscilloscope) as opposed to just a number is that it allows us to show how the voltage varies with time (which is the horizontal axis; the time scale can also be chosen on the oscilloscope).

Because of the graphing ability, we indicate the oscilloscope in a circuit schematic via a circle with a wave in it. This is illustrated to the right, where a battery is connected to three resistors in series. The voltage across the circuit is measured by an oscilloscope (the dashed lines indicate that no current flows through the oscilloscope).



How do we measure two voltages at once?

For the most part, you can treat the two inputs to the oscilloscope as two separate "voltmeters", each with their own controls for scale. However, on some oscilloscopes, when you do this you need to worry about where the "−" port of each input is placed in the circuit.

Why?

For many oscilloscopes, the "−" port is actually connected to parts of the building that are grounded (see volume II). Thus, the two "−" ports are actually connected. If you place those two ports to different points in the circuit, you are essentially shorting out that portion of the circuit.

Furthermore, many signal generators use the same building ground as the oscilloscopes, so we need to make sure that the part of the circuit connected to the ground of the signal generator is also the same part of the circuit that is connected to the ground of the oscilloscope. Otherwise, you'll short out part of the circuit.

✍ One way to avoid this mistake is to just use one of the “−” ports, and leave the other “−” port alone (it will still be physically connected to the circuit via the other “−” port).

11. Lasers

Light in a laser is created via a process called **l**ight **a**mplification by **s**timulated **e**mission of **r**adiation (or LASER). **S**timulated **e**mission refers to the way the laser gets the atoms to emit light. Essentially, an atom can exist at various energy "states". When the atom transfers from a higher energy state to a lower energy state, it emits light (**r**adiation). A laser contains atoms that are stimulated in such a way as to emit the light. Exactly how this works is beyond the scope of this course.

One advantage of using a laser is that only one wavelength is amplified. That means that the light coming out of the laser has only one wavelength (i.e., one color).

Why is only one wavelength produced?

In a laser, the mechanism for producing only one frequency is similar to that for the violin string. Where for a string the traveling waves reflect off the ends, in a laser the light reflects off of mirrors. The reflection causes the light to interfere with itself and form standing waves. The laser cavity is designed so that a standing wave is set up that has the desired frequency. Thus, the dimensions of the cavity will determine the frequency that is amplified.

Since only one frequency is amplified, the light that is produced is all of the same frequency, which is why the laser light is monochromatic[i]. The standard classroom laser uses a mixture of helium and neon and produces light that has a wavelength of 632.8 nm (in vacuum).

[i] This is an idealization. Real lasers are not perfectly monochromatic (due to the Doppler effect; see volume II). However, we can consider them to be monochromatic when compared to random emissions.

Index