Notes follow and parts taken from sources in Bibliography

Introduction

Scientific notation & big/small numbers – Any number can be written in scientific notation. It’s a shorthand way for writing very large or very small numbers, though, so you may not have seen it before. The basic idea is this – rewrite every number so that it’s in the form (something \( \geq 1 \) and \(< 10\)) \( \times 10^{\text{Some whole number}} \)

For example, you can rewrite 50 as \( 5.0 \times 10^1 \), 6.35 as \( 6.35 \times 10^0 \), and 1,000,000 as \( 1.0 \times 10^6 \) because \( 10^1 \) is just 10, \( 10^0 = 1 \) (anything to the zero power is 1), and \( 10^6 \) is a million. The reason for this is the size of numbers we’ll use – for example, if you want to know the distance between our galaxy and its nearest large neighbor, we could write it as about 11,800,000,000,000,000,000,000 miles (we will NOT be using miles) or you could just say “about \( 1.18 \times 10^{19} \) miles”. This gives you an idea of why large numbers are called astronomical.

You can also write numbers \(< 1\) in scientific notation. The thing to remember here is that multiplying by a number raised to a negative power = dividing by the number raised to the same positive power. For example, \((\text{anything}) \times 10^{-3}\) is the same as \((\text{anything}) / 10^3\) which is also the same as \((\text{anything}) / 1000\). So, since we could rewrite 0.02 as \( 2 / 100 \), it becomes \( 2 \times 10^{-2} \) in scientific notation.

Metric units & prefixes – units are typically either MKS (meter-kilogram-second) or CGS (centimeter-gram-second). We could measure large distances in meters if we want – we can either write out large numbers or use scientific notation to make it easier. Of course, we could also just use a larger unit of measurement. If you want to give someone directions to the next town, kilometers would be more appropriate than meters (1 kilometer = 1 km = about 5/8 of a mile). In the same way, we could measure someone’s height in kilometers, but it makes more sense to say “1.8 meters” (about 5’ 11”) than “1.8 \( \times 10^{-3} \) km” or “0.0018 km”. To make this easier, I’ve listed the most common prefixes below:

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Giga</td>
<td>( 10^9 )</td>
<td>1,000,000,000</td>
</tr>
<tr>
<td>Mega</td>
<td>( 10^6 )</td>
<td>1,000,000</td>
</tr>
<tr>
<td>Kilo</td>
<td>( 10^3 )</td>
<td>1,000</td>
</tr>
<tr>
<td>Centi</td>
<td>( 10^{-2} )</td>
<td>0.01</td>
</tr>
<tr>
<td>Milli</td>
<td>( 10^{-3} )</td>
<td>0.001</td>
</tr>
<tr>
<td>Micro</td>
<td>( 10^{-6} )</td>
<td>0.000001</td>
</tr>
<tr>
<td>Nano</td>
<td>( 10^{-9} )</td>
<td>0.000000001</td>
</tr>
</tbody>
</table>

“Giga” and “Mega” are not used as often in science as some of the other prefixes – they pop up more in the computer world, where the numbers they represent are not exactly equal to a billion or a million.

Finally, differences between numbers are sometimes classified by “orders of magnitude”. This essentially means “powers of 10”, so a 600 ft. tall building is two
orders of magnitude taller than a person. We can use this idea with weights, times, or whatever. A small car is about one order of magnitude (10 x) heavier than a large person. Generally, this comes up when you are making rough calculations or the things you are comparing are not known to high accuracy. As an example, I could guess that everyone in the class can lift 100 lbs. and be sure that I was within one order of magnitude either way – probably anyone can lift > 10 lbs., & it’s unlikely anyone will be able to lift more than 1,000 lbs.

Why are these systems of units and ways of writing numbers so important? It’s because we have to have a way to describe things objectively in science. We can’t communicate effectively with other scientists if we only say things like “this new star is far away from Earth”. Not everyone will agree on what “far” means. Even the nearest star (besides the Sun) is very far away if we are considering sending a probe to it. However, if you’re talking to someone who studies cosmology (the universe, its beginning, evolution, and end), the next star over is practically in the same place as the Sun! Using a consistent system of units and giving numbers instead of adjectives prevents misunderstandings.

Other units which are common, but not tied to the Metric system are: Astronomical Unit (AU) – this is the average distance between the Earth and the Sun, or about 150,000,000 km (93,000,000 miles). This is useful when you’re talking about distances to things inside the solar system. Light Year (ly) – distance light travels in one year – about $9.5 \times 10^{15}$ m or $9.5 \times 10^{12}$ km (about 6 trillion miles). Very useful for distances inside our galaxy or distances to nearby galaxies. It’s not uncommon, though, to see the size of the universe expressed in light years. Parsec (pc) – About 3.26 ly. We’ll see later where this value comes from.

The Solar System – the major members of the solar system are the Sun & planets (there are other things we’ll discuss later – comets, asteroids, etc.). The planets can be further divided into two groups – small, rocky bodies the size of Earth or smaller, and much larger gaseous ones. There is also an interesting distribution of these in space – the small ones are closer to the Sun, and the large ones are further away (except Pluto, which is small and far away, and doesn’t really fit with its neighbors. Data on it is included below, but it is no longer considered a planet by most astronomers). As the semester progresses, we’ll see that this arrangement is something we could expect to find in any other planetary systems we may discover later.

As we just saw, saying “larger” and “smaller” is not an objective way to describe things. So, we’ll look at a few numbers in the table below
### Table: Planetary Data

<table>
<thead>
<tr>
<th>Planet</th>
<th>Radius (km)</th>
<th>Distance to Sun (AU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mercury</td>
<td>2439</td>
<td>0.39</td>
</tr>
<tr>
<td>Venus</td>
<td>6052</td>
<td>0.72</td>
</tr>
<tr>
<td>Earth</td>
<td>6378</td>
<td>1</td>
</tr>
<tr>
<td>Mars</td>
<td>3374</td>
<td>1.5</td>
</tr>
<tr>
<td>Jupiter</td>
<td>71400</td>
<td>5.2</td>
</tr>
<tr>
<td>Saturn</td>
<td>60300</td>
<td>9.5</td>
</tr>
<tr>
<td>Uranus</td>
<td>25600</td>
<td>19</td>
</tr>
<tr>
<td>Neptune</td>
<td>24800</td>
<td>30</td>
</tr>
<tr>
<td>Pluto</td>
<td>1150</td>
<td>39.5</td>
</tr>
</tbody>
</table>

You can see that Jupiter, for example, has a diameter about 11x Earth’s. When talking about the relative sizes of any two things, we also have to be clear about what we’re comparing. We can’t just say “Jupiter is 11 times the size of Earth” because it doesn’t tell the whole story. Another measurement we could look at is the volume of Jupiter vs. Earth. The good part about this is, we don’t need another table & we don’t need a calculator to use on it to get the ratio of volumes. That’s because of two simple facts: 1) Earth, Jupiter, and almost every other large (100’s of km or more) single object in the universe (stars, planets, moons, etc.) is spherical to a pretty good approximation – we’ll find out why when we look at the law of gravity. 2) There’s a simple relation between the radius of a sphere and its volume which you learned a few years ago. It is: $\text{Volume} = \frac{4}{3} \pi (\text{radius})^3$. Here’s the part that makes things easier. If we want to know the relative volumes ($\text{Vol. of Jupiter} / \text{Vol. of Earth}$), we can do this

$$\frac{4}{3} \pi (\text{radius of Jupiter})^3 \quad \text{Cancel} \frac{4}{3} \pi \quad \text{to get} \quad \frac{(\text{radius of Jupiter})^3}{(\text{radius of Earth})^3} \quad \text{which is also} \quad \left(\frac{\text{radius of Jupiter}}{\text{radius of Earth}}\right)^3.$$  

Remember that we already know the ratio of the radii (it’s about 11). This means the ratio of volumes is about $11^3$ or 1331 (we’re doing rough calculations here – we didn’t start with EXACTLY 11 so we can’t really quote our final result to 50 decimal places. We’ll call it 1300). So, we see that Jupiter has 1300 x the volume of Earth.

There’s only one thing stopping us from saying that Jupiter has 1300x the mass of Earth, and that is the fact that Jupiter is mostly light gas and Earth is mostly rock. Because of that, Jupiter is really only about 315 times the mass of Earth. For comparison, all the other planets in the Solar System combined have a smaller mass than Jupiter.
Because Jupiter is the largest of the gaseous planets, they’re sometimes called the Jovian planets (this was the Roman adjective for things relating to Jupiter). The others are Saturn, Uranus, and Neptune, in that order from the Sun. These worlds are different from the inner ones in several ways – they all have ring systems (only Saturn’s are obvious, but the others are there), they all have multiple satellites (8 or more), they are all far from the Sun (compared to the Earth), they have very heavy atmospheres, and they’re cold. As we study these planets in more depth, we’ll see that many of these things are related and therefore not too surprising (lots of satellites could reasonably be associated with ring systems, large distances from the Sun coincide with thick, cold atmospheres, etc.). A large part of the course will be separating things that really are unexpected with things that are only odd until we learn how they really work.

The inner planets also have many similarities – they’re mostly rocky, warm, and covered with thin atmospheres, if any (these properties aren’t as uniform as they are for the Jovian planets – Venus actually has a very thick atmosphere compared to the Earth, but it’s still thin on the scale of the planet’s size, unlike the Jovian planets). Also, while the Earth has a large moon, Mars only has two extremely small moons (about the size of cities) and Venus & Mercury have no satellites at all.

When we compare the planets in the Solar System to try & figure out why things are the way they are, how they got that way, and what we could expect in other planetary systems, we’ll use an idea called comparative planetology. We can use this to explain things we observe around us – why do we not expect to see “moonquakes” on the Moon the way we have earthquakes here? Why doesn’t the Moon have an atmosphere? Will the Earth always have one?

In this discussion of the planets, we haven’t mentioned Pluto. The reason is, it doesn’t seem to fit in with the other planets in many ways. It’s in the wrong part of the Solar System to be so small. It also has a very large moon for its size. Where most of the rest of the planetary orbits are so close to being circles that we’d have a hard time noticing they weren’t without careful observation, Pluto’s is noticeably elliptical. Its orbit is also the one that is at the most dramatic angle to the Earth’s orbital plane. Finally, it’s much smaller than even Mercury – about ½ its radius and less than 10% of its mass. The current theory to explain these disagreements is that Pluto didn’t form the way the rest of the planets did and is more likely a captured traveler from the outskirts of the Solar System. If it didn’t form in the same area, it’s not such a big surprise that it doesn’t fit in with the other planets. A few years ago, it was actually “demoted” from a planet to a Kuiper Belt Object.

The Sun is the central object in the Solar System. Its diameter is about 100x Earth’s or 10 x Jupiter’s. From our knowledge of how volume scales with size, that tells us that it’s about 1000x the volume of Jupiter or 1000000 (10^6) x the volume of Earth. If we look at masses, the Sun’s mass actually is about 1000x Jupiter’s mass. That means that the Sun and Jupiter have about the same density (which is reasonable because we’ll see they are both mostly hydrogen). The Sun provides the light & heat
life need to survive on the Earth. It was demonstrated about a hundred years ago that the source of the Sun’s energy can’t be chemical (an ordinary fire) given what we know about the age of the Earth, which should be no older than the Sun, and the size of the Sun. If you take the next semester of astronomy, you will study the Sun more closely and see that its source of energy is the same process that makes the hydrogen bomb such a powerful weapon.

Now we have some kind of idea about the typical sizes of other objects in the Solar System and their distances from the Sun. Moving to the next level involves finding the distances to the nearby stars (which may have planetary systems of their own). The distances to the nearest stars are found by parallax measurements, which are not too different from the way surveyors used to measure distances on the Earth. These numbers are very accurate and give typical interstellar distances of several light-years. Examining other stars has shown that the Sun is very much an average star – by no means is it one of the largest or brightest even among its neighbors. There is a wide range of stellar sizes and temperatures which represent different initial conditions for each star’s birth, and also different points in the life cycle of the stars. Many stars are part of binary-star systems where two stars orbit each other at distances of several AU’s or so. This kind of star system may not generally allow the formation of planets, and may be very hostile to the development of life on any planets that do form.

Stars can also tend to be found in huge collections of millions or billions of stars known as galaxies. Our galaxy is called the Milky Way and is approximately 100,000 l.y. across. We’re about halfway out from the center, and it takes about 250,000,000 years for the Sun & Solar System to make one orbit around the galaxy. That means that one “galactic year” ago, dinosaurs were roaming the Earth. The nearest large galaxy to the Milky Way is the Andromeda galaxy, about $2.5 \times 10^6$ l.y. away. As you might have expected, these galaxies clump together in groups, and the groups clump together in superclusters. This is about the end of the organizational scale though; the size of the universe (best guess so far – watch the newspapers for updates) is about 10-20 billion l.y. There is an interesting pattern you may have spotted by now – things seem to be separated from other similar things by empty spaces large compared to their own size. For example, the planets are separated by AU’s, which are huge compared to planetary sizes. The stars are separated by ly, huge compared to their sizes. That’s why it is no exaggeration to say that most of the universe is empty space. (This also applies at the small scale, too. Over 99.9% of the mass of an atom is concentrated in the nucleus, which is about $10^{-15}$ the volume of the atom!)

**Science and Pseudo-Science**

Science makes predictions based on observations, interpretation of data, and logical reasoning. Science should be completely unbiased – a scientist will generally try to think of every possible way to knock down his/her own work, for the simple reason that others will certainly do it. If you can think of some reason why what you have suggested is wrong, you also know more about what the right answer might be.
Pseudoscience generally starts from the end and works backward to find facts which seem to match the desired endpoint.

**Difficulties of Astronomy** – The time scale of modern astronomy (~100 years) is virtually nothing on the scale of stellar/galactic/cosmological change. What could you learn about a human’s life cycle by watching a snapshot of thousands of humans? You’d have to guess which of these different images represent steps on a timeline, and which don’t. Baby boys don’t grow into middle-aged women on the way to being old men. Also, size plays a role, but it’s not absolute – babies grow into larger people, but people in their nineties are not generally the largest people around, and NFL linebackers are not the oldest people on Earth.

**Philosophies of Science** – Modern science is based on the testing and formation of hypotheses. These are basically ideas that make testable predictions and explain pre-existing observations. A hypothesis is what an experiment will support or refute. Using a hypothesis to make predictions which are then evaluated by experiments is called deductive reasoning. Starting with experiments and using them to find a hypothesis explaining things is called inductive reasoning.

If you observe specific things and then try to generalize them to a hypothesis, that’s inductive (you drop 10 different kinds of metal into water, and they all sink – inductively, you conclude that metals are denser than water). Now, someone wants to test your hypothesis that metals are denser than water – they drop lithium into water (it would actually explode, but let’s not worry about that). It floats, because it’s less dense than water. Your hypothesis now needs to be modified. SOME metals are denser than water. That testing of your hypothesis was deductive.

If your hypothesis survives many tests over a period of time (which can include some modifications) it will eventually rise to the level of a theory. For some reason, there is a popular belief that anything which is “only” a theory is somehow untested. The word theory in science actually means a great deal. We talk about the theory of relativity or the theory of quantum mechanics in physics – these are actually the most accurate theories the world has ever known. They make predictions (confirmed by experiment) to accuracies of parts-per-billion. As pointed out by Richard Feynman, this is like knowing the distance between New York and Los Angeles to within a few centimeters! Still, we call them theories.

Astronomy operates under restrictions which don’t constrain the other sciences. Chemists can combine elements to make the compounds they want to study, geologists can collect samples of rocks for study in the lab, biologists can put specimens under a microscope, but astronomers basically have to sit on Earth & collect light (some of which is not in the visible range) to answer their questions. The upside is that, with so much space and so many stars around, almost anything that can physically happen is happening somewhere – you just have to find it. Another thing to keep in mind about astronomy is the connection between distance and time. Light travels at about 300,000 km/sec so whenever you look anywhere, you’re really
looking into the past. If you’re looking across the room, it’s only a few nanoseconds into the past, but if you’re looking at distant galaxies, you could be looking billions of years into the past. As we just said, we are in one sense restricted by the fact that a human lifetime is so short on the cosmic scale that changes to the stars are not readily apparent. However, we can use the fact that great distances = many years ago to get some sense of stellar or cosmological evolution through time. We can’t see what our galaxy was like 5 billion years ago, but if we’re looking at light that left other galaxies 5 billion years ago, that could give us some important clues.

**Growth of Science in the 20th century** - Science has seen tremendous growth in the last 100 years. Newton’s ideas, which most of science rests on, have been shown to be incomplete (NOT the same thing as being wrong – they are still good enough for many purposes. Only when we talk about speeds which are close to the speed of light or distances which are smaller than an atom or incredible concentrations of mass do we start to notice significant deviations from the laws we’ve used for 350 years.) There are many different reasons for the increase in knowledge lately. First, society can afford the “luxury” of having a larger percentage of the population involved in science. We don’t need to have 50% of the country farming to feed themselves and the other 50%. Second, governments and societies have seen great benefits come from scientific studies, and that encourages investment in science. Technology, which can be thought of as the movement of an idea or device from the laboratory to mass production, has made similar leaps since about 1900. At that time, cars were rare, planes didn’t exist, and electricity was somewhat new. Since then, science has helped refine cars & planes, designed rockets, created lasers, transistors, computers, etc. There has been a beneficial feedback from technology into science – the presence of computers allows scientists to gather far more data, analyze it more exhaustively, and test theories faster than would ever be possible without them.

**Separating Science and Pseudoscience** – there are lots of tools you can use to do this. First, skepticism is important. If you hear an extraordinary claim (this magnet will relieve pain anywhere in your body caused by anything), that should set off alarm bells for you. Examine things like

1) Is there some way this could possibly work? This requires knowing some of the basic principles of science that we’ll hopefully learn in here. For example, getting more energy out of something than you put in (perpetual motion) seems to crop up now and then. Most scientists won’t even bother trying to deflate these claims anymore – they’re too ridiculous to waste time on. The Patent Office will reject any application like this without reading further.

2) Where did you hear about this magical device/theory? Has it been published in a reputable scientific journal? Or was it published at the writer’s expense? The reason science operates by publishing in journals is that articles submitted are sent to two or more reviewers – other scientists who are experts in the subject of the article. Their job is to carefully analyze the experiment or theory discussed and look for any holes
the author might have missed. If a mistake makes it past the reviewers and editors, it will be published for hundreds or thousands (depending on the journal) of other experts in that general area of study to see. This process has a tendency to quickly kill bad ideas. As a general rule, scientists who announce discoveries via news conferences rather than journal articles are looked at suspiciously by others scientists.

3) How does the author present the work? Frequently, pseudoscientists will claim that “organized science” is somehow out to get them. There is some conspiracy between the government, big businesses, and/or scientists to keep news of their discovery quiet.

4) Have other people been able to confirm or reproduce the results? If not, that should be a clue that something is off. As mentioned above, if you can quickly poke a hole in the experiment, analysis, or conclusions in a published paper, you have a quick paper of your own. The other scientist’s reputation will take a small (or large, depending) hit if he/she has published nonsense.

5) Are testable predictions made by the theory? The theory should be more than a “fit” to selected pieces of data (or even all of the available data) – something new should come out of the theory. You can use Excel or other computer programs to fit a curve perfectly to an arbitrary amount of data; if there is no science behind the fit, though, you will probably find that the next data point measured will have no relation to the complicated curve that fit the others.

6) Existing data shouldn’t be ignored – in some cases, it will be mentioned and then brushed off for some less-than-valid reason. Going back to an earlier example, Einstein’s theory of relativity did NOT change the validity of Newton’s laws in most cases (almost all cases confined to the Earth). In fact, if you look at certain limits (small masses, low velocities), you’ll find Einstein’s laws produce Newton’s laws.

These points will help you decide the value of a new theory or new product. Even with the number of governmental agencies regulating food, medicines, vitamins, etc., there’s a huge amount of garbage which dishonest people will try to sell to anyone who doesn’t know better. This brings up the question of motivations for pseudoscience, and there are many of these: First, there are people who are genuinely convinced that they have made a fundamental breakthrough and it is being ignored or suppressed for some reason (generally referred to as crackpots). Frequently, the reason given for this obstruction is that “these scientists don’t want to admit that everything they know is wrong and that they have to go back to square one to understand my theory”. Possible? Sure. There will be people in any field who resist change just because it is a change. What history has shown, though, is that when someone presents a “crazy” idea which actually holds up to scrutiny, the scientists who learn it and accept it first are generally the ones who are able to publish lots of important papers on the further implications of the new idea. It has happened with both quantum theory and relativity in the 20th century.
Another motivation is profit – the person presenting the new theory is selling something based on it, and they have a financial interest in convincing others that they’re right. Sales of wearable magnets, copper bracelets, etc., are an example of this. The government is as susceptible to pseudoscience as any person. NASA has, in the past, funded truly ridiculous ideas. This is sometimes defended by using arguments like “Well, you always need to question everything.” The problem with this idea is that, once an idea has been questioned over and over and over again, and never failed even once, you reach a point where you’re wasting your time and money by investigating further. How many glasses do you need to drop before you accept the fact that gravity makes things fall toward the Earth?

Why do people fall for pseudoscience scams? Some of the same reasons.

1) One popular one is that they are promised something that they should recognize is probably too good to be true. If you have chronic pain, or some incurable disease, the promise of relief or a cure is going to be very attractive. If a person is desperate, he/she may try almost anything.

2) The belief that a company couldn’t make claims unless they were true. In fact, there are many commercials or products for sale from companies large and small that are obviously worthless. The process of the FDA, FTC, etc. getting enough reports to investigate and then successfully prosecuting fraudulent sellers can take years (literally). By the time the people behind a dishonest commercial are finally forced to stop airing it, they will have long since moved on to some other scam.

3) A strong economic incentive to believe in the pseudoscience. People who had heavily invested in tobacco stocks decades ago were probably among those who wanted to deny that smoking causing cancer. Similarly, large coal and oil companies want to stir up as much dust around the issue of global warming as possible.

4) A need to feel “special” in some important way. Developing some “revolutionary” new idea, or being one of the (few) first people to believe in it, means you were right before everyone else. That suggests you are smarter than everyone else.

**Angles & Errors**

**Measurement of Angles** - The use of an angular measurement is the only way we can say how far apart two objects in the sky appear to be from each other. The big things to remember here are 1) this information alone is not enough to tell us how far apart these two things actually are from each other or from us 2) the angular measurement depends on the true distances between the three objects – for example, from the Earth, the angular separation between the Sun & the full Moon is 180°. If you were on another planet, that separation might be anywhere from 0° to 180° (see below).
Small angle approximation - the formula for the angle subtended by something which is far away from you (far meaning a large distance compared to its true diameter) is:

\[ \theta = \frac{True \text{ object diameter}}{True \text{ dist to object}} \times 57.3^\circ \]

The degree is actually a rather large unit of angular measure for most objects in the sky (just like you could measure your height in miles, but it’s not very convenient). The degree can be divided into 60 pieces, called minutes of arc or arc-minutes, and abbreviated with an apostrophe (35' = 35 minutes of arc). The Moon is about 30' in angular size as seen from Earth. The arc-minute can then be divided into 60 arc-seconds, abbreviated with quotation marks. This means that we can write: 1' = 60' = 3,600". Make sure you don’t confuse minutes/seconds of time with minutes/seconds of arc, because they measure completely different things (time vs. angle). You’ve actually encountered this before when you’ve talked about ounces – an ordinary soft drink can holds 12 ounces (volume measurement) and a pound is 16 ounces (weight measurement). The number of arc-seconds in 57.3° is 206,265.

One of the most basic and most accurate ways to measure distances to things we can’t reach involves angular measurements. Imagine you need to measure the distance to a tree across a river. You don’t want to get wet, so what can you do? You can construct a triangle with the tree at one corner. As shown below, the other two corners can be on your side of the river, and if you measure the distance between them and the angles at those two corners, you will have all the information you need to use trigonometry to find the distance to the tree.
In practical terms, you can imagine mounting two small telescopes or spotting scopes on a protractor. You point one at the tree, and one at your partner on your side of the river. Your partner has the same equipment and makes the same measurements. You use a tape measure to get the distance between you and your partner, and the rest is math.

This is very similar to the way we measure distances to nearby stars. The distance between you and your friend (known as the **baseline**) and your ability to measure angles will determine your accuracy. We want the largest baseline possible because stars are ridiculously far away. To do this, we use the Earth’s orbit. We measure a star’s position relative to the background of very distant stars, and then wait 6 months so that we are at the opposite side of our orbit and do it again. The change in that angle is related to (actually twice the size of) what we define as **parallax**. The parallax of a star is measured in arc-seconds since they are all so far away compared to the size of our baseline (2 AUs). An object with a parallax of one second of arc (if we had a star other than the Sun that close) would be 206,265 AUs from Earth. That works out to 3.26 light years, and that's where the **parsec** comes from. It is a contraction of **parallax** and **arc-second**.

To help imagine what this process would look like, see the image below. If the star were very close, the situation (shown on the left side of the picture) would look like the one above with the tree and the river. In reality, the stars are all so far away that the triangle is ridiculously long and narrow, as on the right. The two angles are each very, very close to 90°, and the process of determining parallax (and therefore distance) boils down to the ability to measure the tiny difference between the angles below and a true 90° angle.
Measurement Errors – Precision & accuracy mean two different things when discussing measurements. If a series of measurements is precise, that means the random error in the measurements is small. The measuring device will repeatedly give the same answers (or nearly the same) when measuring the same things over and over. This does not necessarily mean the answers are correct. A measurement can be very precise and completely wrong. If a series of measurements is accurate, that means that, on average, the measurements are correct. This means that the systematic error is small. Systematic errors are always in one direction – random errors can be high or low. An example of random error would be stepping on a scale 10 times and getting 4 or 5 different answers. An example of systematic error would be if a friend is standing behind you & keeps putting a weight on the scale when you get on.

One example of the problems caused by systematic errors in astronomy is found in the 1922 estimate of the size of the universe made by astronomer Jacobus Kapteyn. He didn’t use the assumption that all stars have the same intrinsic brightness and that apparent variations in brightness are only due to different distances to the stars. His estimate of the universe’s diameter was 10,000 l.y. This was later shown to be too small by a factor of $\sim 10^5$. The problem was that absorption of light by interstellar dust wasn’t considered to be too severe from results of preliminary surveys. The star counts were all interpreted incorrectly in a systematic way rather than a random way.

Basic Sky Observations

Early observations – People have studied the motions of the Sun, planets, and stars through the sky for thousands of years. In all parts of the world, astronomers realized that the patterns are predictable. The Mayans, Egyptians, Chinese, and Arabs all made important contributions centuries before western science began. Early astronomers predicted eclipses and observed “guest stars” (exploding stars in our galaxy), developed calendars, understood solstices/equinoxes, and noticed that the planets were definitely different from the stars. By observing the shadow of the Earth on the Moon during lunar eclipses, the Greeks knew the Earth must be a sphere at least as early as about 500 BC. They also correctly determined the sizes of the Earth and Moon.
Motion of the stars – the stars move constantly on different cycles & timescales. The daily rotation of the Earth makes stars (and Sun, Moon, and planets) appear to move from East to West across the sky. The projections of the Earth’s geographic North & South poles onto the sky are known as the North & South Celestial Poles (NCP & SCP). These are the points in the sky that the stars seem to rotate around. We only see the NCP – the SCP is below the horizon for us, and can only be seen from the equator or below, just like the NCP can only be seen from the equator or above. The projection of Earth’s equator onto the sky is known as the celestial equator. There happens to be a bright star very close (currently) to the NCP – it’s known as the North Star because it always lies in the North (unless it’s directly overhead, which means you’re already at the North Pole). As you move from the North Pole down to the Equator, Polaris moves from overhead down to the horizon. Its altitude (number of degrees above the horizon) is equal to your latitude.

Motion of Sun – The stars seem to rise a few minutes earlier every day. This is because the Sun seems to move to the East relative to the stars. To observe the relative Sun/stars motion, we need to define two kinds of day. The solar day is the standard one we use every day – it’s the time between two successive observations of the Sun going through its highest point (or through the local meridian, which is the line overhead connecting the NCP and the SCP). This is 24 hours.

The other kind of day is the sidereal day. This is the time between successive observations of the same star going through the local meridian. This is also the time it takes the Earth to rotate 360°. This is 23 h 56 m. The reason for the difference in times is that, over a day, the Earth moves around the Sun by about 1° (1° because it takes 365 days to move the full 360° around the Sun ~ 1°/day). The Earth now has to rotate through 361° rather than 360° to put the Sun back in the same place. How long does this extra degree take to rotate through? It is easily calculated as: 360° in 23 h 56 m (or 1436 minutes). 1436 minutes/360° = 3.99 minutes (~ 4 minutes) per degree. Therefore, it takes 23 h 56 m + 4 m = 24 hours to get the solar day. This also means that the length of the year in days depends on which kind of day you’re using – it’s 365.25 solar days, but 366.25 sidereal days. The other way to understand this is to realize that the trip around the Sun gives one more rotation as seen by the stars. That’s where all the other numbers really come from. Also, this has nothing to do with leap years – the leap year exists because the Earth doesn’t make a whole number of rotations during one trip around the Sun. There’s no reason it should – the two rotational motions are not strongly connected.

Seasons – The Earth’s rotational axis is not perpendicular to its orbital plane. There is an angle between the NCP (Earth’s axis) and the North Ecliptic Pole (North end of a line perpendicular to the Earth’s orbit & passing through Earth’s center) of 23.5°. Because we see everything from the same point of view all year (if we stay in one place), we see the Sun seem to move not just from East to West daily, but also from N to S (Summer & Fall) and then S to N (Winter & Spring). Two times a year, the Earth’s axis of rotation does not point towards the Sun at all – it’s perpendicular to a
line joining the Earth and the Sun. These times are called the Equinoxes (from Latin for equal night – the night & day have equal lengths – 12 h) and they happen on the 1st day of spring (about March 21st) and the 1st day of fall (about September 23rd). On these days, the sun rises exactly in the East & sets exactly in the West.

When the northern end of the Earth’s axis starts to have a component pointing away from the Sun, we are entering the season of days shorter than 12 hours (Fall). We’re in the northern hemisphere, so we see shorter days and less direct sunlight – this is because the Sun is lower in the sky all day long. It also rises more to the south of east and sets more to the south of west. Since the Sun is lower in the sky, less of its light gets spread over more of the Earth, making the rays less intense. A quick example is to shine a flashlight directly down onto a table so that you see a circle of light. Then, tilt the flashlight so the circle gets longer and skinnier. You’re not producing any more light, and it’s being spread over a larger area, so it gets less intense. These shorter days of less intense sunlight lead naturally to cooler temperatures in our hemisphere. Six months later, the northern tip of Earth’s axis starts to point more towards the Sun. The Sun is higher in the sky (though never overhead outside of the tropics) and it rises north of east and sets north of west. Longer days with more direct sunlight heat up our hemisphere & cause summer. There’s actually a lag between the longest, most sunlight-intense day (June 21, known as the summer solstice – the Sun is at its northernmost point & turns around & heads south) and the hottest part of the year (usually July/August) which is caused by the oceans serving as a giant heat-sink. They’re hard to heat up, & they take a while to cool down (which is why the shortest day with the Sun at the lowest altitude (Dec. 21 – winter solstice – Sun’s southernmost point, when it is lowest in the sky & turns around heading north) is a couple of months before the coldest days (usually January/February). Notice that we didn’t mention distance to the Sun anywhere in here. We’re actually closest to the Sun around the first few days of the year. This acts to make the Northern seasons a little less intense than the Southern ones (but local climate effects and the distribution of land/sea on the Earth probably make this impossible to detect).

**Stellar Coordinate Systems** – Because the Earth rotates relative to the distant stars, they are not in the same place every night or even at different times on the same night. There are 2 obvious ways to set your coordinates up – attach them to the Earth, or attach them to the stars. The most convenient way to locate a particular star is probably the altitude-azimuth or horizon system. This specifies a star’s location as a number of degrees away from North (measured to the East) which is the azimuth, and a number of degrees above the horizon (measured towards the zenith) which is the altitude (stars below the horizon have negative altitudes). This system is attached to the Earth because if you go outside and find the altitude and azimuth of the top of a tall building, that will not change from day to day or year to year. This is an easy system in one sense, because everyone can find North & make estimates of where “30° altitude” and “105° East of North” would be. Unfortunately, these numbers depend on your location on the Earth (different for NY vs. LA) and the time of year AND time of day. In other words, somebody had to do a significant number of
calculations to provide you with the altitude & azimuth of a star in advance from another city.

The other system is known as the Equatorial system. This uses coordinates centered on the sky. The two coordinates in this system are known as Right Ascension and Declination. Declination is the star’s position N or S of the celestial equator and is measured in degrees. Right ascension is measured from the N-S line which passes through the point where the ecliptic and the celestial equator cross (Sun’s position on the vernal equinox = 1st day of Spring). RA is therefore like longitude and declination is like latitude. Also, RA is measured in hours (rather than degrees) East from the crossing point mentioned above. The hours represent the fact that the Earth rotates once in 24 hours. One rotation = one complete circle = 360°, so 24 h = 360° which means 1 hour of RA = 15°. The good part about these coordinates is that, once you know the RA & dec. for a star, they won’t change over your lifetime. The bad part is, you still need to be able to find these stars at night, and RA & dec. are not easily determined by just looking at the night sky.

**Precession** – Something that will change a star’s RA & dec. (VERY SLOWLY) is the phenomena of precession. This is the gradual change of the direction in which the North pole is pointing. The NCP therefore moves around the sky over time. It actually traces out a circle every 26,000 years. This is what we meant by saying Polaris was the **current** North Star, not a permanent one. In about 12,000 years, Vega will be the star closest to the NCP. In 26,000 years, it’ll be Polaris again. This precession is just like the way a spinning top moves. It not only spins very quickly on its axis, but its axis moves around in a smaller circle so that the top seems to trace out a cone. This is because the axis is spinning in the gravitational field of the Earth, and the Earth is pulling down on the “top” of the top. The net result is precession. The same thing happens when the Earth spins rapidly (on a scale of 26,000 years, a day is very quick!) in the gravitational fields of the Sun and Moon.

**Eclipses & the Motion of the Moon**

**Moon** – The Moon moves around the Earth the same direction in which the Earth moves around the Sun. Since the Earth is moving around the Sun while the Moon tries to circle the Earth, we have two distinct “months”, just like we have two kinds of days – solar and sidereal. The Moon’s sidereal period is 27.3 days. This is the amount of time necessary for the Moon to make one trip around the Earth relative to the stars. The time for the Moon to go from one phase through its full cycle and back to that phase is longer, just as the time between two passages of the Sun through the local meridian is longer than that for two passages of a star through the local meridian. This longer period is called the Moon’s synodic period because it is relative to the Sun. The length of the synodic period is easily found from the sidereal period. 360° are covered in 27.3 days, which means the Moon moves 13.2° around the Earth each day. However, after 27.3 days, the Earth has moved (360°/365.25)*27.3 = 26.9° from where it was when the Moon started its orbit. 26.9°/(13.2° per day) = ~ 2 days. 27.3 + about 2 = 29.5 days for the synodic period. (it takes about one more step to
see that because of the constant chase between the Earth and Moon. While the Earth moved almost 27° during the Moon’s sidereal period and the Moon has to move for 2 more days to get near it, the Earth is still running from the Moon for those 2 days and the Moon will have to move a few hours to finally catch it. When these are all added to the 27.3, you get about 29.5)

**Phases of the Moon** – The phase of the Moon (as seen by people on Earth – there’s no such thing as an “absolute” phase) depends on the positions of the Moon, Sun, and Earth. Half of the Moon is always lit by the Sun (except in the rare case of a lunar eclipse when part or all of the Moon enters Earth’s shadow). During a full Moon, the lit half points directly at the Earth. During a New Moon, it points directly away from the Earth (we see only the dark side, so we can’t see anything). Incidentally, there is no dark side of the Moon – ½ of it is always dark and ½ of it is always lit (again, except for lunar eclipses) and the halves change slowly and constantly over the Moon’s orbit. An astronaut landing anywhere on the Moon would see the Sun crawl across the sky, taking about 2 weeks to move from horizon to horizon, and then would have 2 weeks of night waiting for the next sunrise. However, since the Moon always keeps the same face pointed at the Earth, wherever the Earth was in the sky when the astronaut landed, it would always stay there unless he/she moved.

The basic phases of the Moon are New (Moon between Earth & Sun), 1st quarter (Moon directly “behind” the Earth relative to the direction Earth is going in its orbit), Full Moon (Earth between Moon & Sun), and 3rd quarter (Moon “in the way” of Earth’s movement through its orbit). Additionally, there are crescent phases (between New and 1st or 3rd quarter – only a sliver of Moon visible) and gibbous phases (between Full and 1st or 3rd quarter – all but a little of the Moon visible), each of which can be waxing (heading towards Full) or waning (heading towards new).

**Moon’s orbit and eclipses** – The Moon’s orbit is not exactly in line with the Earth’s orbit (=apparent path of the Sun through the sky = ecliptic). If the two orbits were in the same plane, we’d have a solar eclipse and a lunar eclipse every month. As it is, we have them much less often – 2 to 5 times per year, each. The line marking the intersection between Earth’s orbit and the Moon’s orbit does not point in a constant direction. If it did, there would be two times each year (the same time every year) when eclipses might occur.

(There are multiple animations of this available on the internet that will make the process easier to visualize. You should look at the ones at [http://session.masteringastronomy.com/problemAsset/1016295/8/Lesson1.swf](http://session.masteringastronomy.com/problemAsset/1016295/8/Lesson1.swf) and [http://highered.mheducation.com/olcweb/cgi/pluginpop.cgi?it=swf::640::480::/sites/dl/free/007299181x/220730/eclipse_interactive.swf::Eclipse%20Interactive](http://highered.mheducation.com/olcweb/cgi/pluginpop.cgi?it=swf::640::480::/sites/dl/free/007299181x/220730/eclipse_interactive.swf::Eclipse%20Interactive) )
This would be when the line of intersection (called the **line of nodes**) pointed at the Sun. When the Moon passes through the ecliptic plane, it has to be either the New Moon or the Full Moon before an eclipse can happen. Things could only line up this way when the line of nodes points to the Sun. Adding to the complexity of predicting eclipses, the Moon’s orbital plane and therefore its line of nodes does not stay in the same place. It precesses (backwards relative to Earth’s motion) at a rate of one revolution every 18.6 years. This is caused by the fact that the Earth’s rotational axis is not only precessing, but wobbling slightly (changing the 23.5° angle by a few arc seconds) over 18.6 years.

An eclipse is possible when the Full or New Moon passes near a node (i.e., the line of nodes is pointing near the Sun). For about 15-19 days on either side of perfect alignment (line of nodes pointing at the Sun), an eclipse can happen if the Moon is in the right place. This time is called an eclipse season. Because of the Moon’s regression of nodes mentioned before, this happens a little more often than 2 times a year. It’s 2 times in about 346 days. Since a year is longer than this, by a little, we could conceivably have 3 eclipse seasons in one year. This happens about every 9 years or so. We'll always have at least 2 eclipse seasons in a year, though, so that’s why we say there is a minimum number of eclipses (of all kinds).

The material highlighted in the box below is not essential, but may interest you.

---

Because the eclipse season is a little longer than the cycle of phases (29.5 days) we could also have 3 eclipses within about 30 days. In other words, we could have a solar eclipse at New Moon, a lunar eclipse at the following Full Moon, and another solar eclipse at the next New Moon. This happens about once every 7 years or so. Very rarely, we can have 2 3-eclipse seasons and one 2-eclipse season (1935 was an example) giving us as many as 5 solar eclipses in a single year (total of 7 of all kinds). Eclipses occur in predictable cycles. To go from one eclipse to the exact same “kind” of eclipse, we have to wait for 3 different cycles to coincide. The first cycle is the Moon’s cycle of phases, or synodic month (29.53 days). We know this is important because Solar eclipses (for example) only happen when the Moon is new. The next cycle is the time between passages of the Moon through a node (actually, the same node. The nodes are different – one is called the **ascending node** because the Moon is coming from below the Ecliptic plane to above it as it crosses the node. The other is the **descending node**). This period is known as the **draconic month** and is 27.21 days. We know this is important because, if the Moon is not at or near a node when it’s Full or New, there will be no eclipse because the shadow of the Moon (or Earth) will not fall on the Earth (or Moon).

Finally, will the Moon be close to the Earth (total solar eclipse) or far from the Earth (annular eclipse)? It seems like this would just depend on the draconic month above, but the Moon’s orbit is also sliding around while it precesses, kind of like a bent hula-hoop. The period of this last effect is called the **anomalistic month**, and is 27.55 days long.
These three periods are all nearly the same length, which means it will take a long time for them to all get back to the start of their cycles again. As an example, imagine you’re watching a race where one runner is slightly faster than a second, and the second is slightly faster than a third. They’re all together at the start, but they move away from each other as the race continues. If it goes on for a very long time, the fastest runner will pass the slowest runner, and then later pass the second fastest runner. As you can imagine, it will take many laps before they all end up at the same place again (by which time the fastest runner will have made several more trips around the track than the slowest).

In our case, 223 Synodic months = 6585.32 days, 242 Draconic months = 6585.36 days, and 239 anomalistic months = 6585.54 days. That’s close enough. This time period is known as a Saros cycle or just a Saros. It’s about 18 years, 10 1/3 days (or 11 1/3, depending on leap years in the 18 years). When you observe one eclipse, you can know that if you wait for one Saros cycle, you should see another one. However, since the number of days is not a whole number, the eclipse will not be at the same place on Earth. It’s about 1/3 of a day more than a whole number, so the second eclipse will happen about 1/3 of the way around from the first one. The eclipse of Aug.11 1999 (totality over Europe) and the one of Aug. 21 2017 (totality here) are the same in terms of the Saros. 3 Saros cycles will return an eclipse to approximately where it was on the Earth. That means the eclipse of July 20, 1963 took a path near the one which will be taken by the 2017 eclipse over the US.

It’s important to mention that this does not mean that you have to wait 54 years between eclipses at the same place. It just tells you when you’re essentially guaranteed another one based on your last one. At any given time, there are many different Saros cycles in progress, so you can potentially see multiple eclipses in just one year. The “essentially guaranteed” above means that this doesn’t go on forever. After hundreds of years, the fact that our 3 cycles above are all slightly different (by fractions of a day) and all are different from a whole number of days finally adds up, and that particular Saros is over.

There are a few kinds of solar eclipses – the most dramatic is the total solar eclipse where the full face of the Sun is covered & the corona (outer atmosphere of the Sun) is visible. North or South of the path of totality, observers will see a partial solar eclipse which can be anything from just less than totality (near the path) to just a small segment of the Sun disappearing briefly (far away from the path of totality). Also, since the Moon’s distance to the Earth is not constant, its angular size is not constant – if the eclipse happens when the Moon is farther away from the Earth, its size will appear smaller and a ring of sunlight will remain. This is called an annular eclipse. Places experiencing a total solar eclipse will have the Moon’s umbra (darkest central part of the shadow) pass over them. Partial solar eclipses are seen when the penumbra (outer partial shadow) passes over a place. An annular eclipse is seen if the umbra is too short to reach the Earth’s surface.
Annular Eclipse

Total/Partial Eclipse

Time – Risings and Settings – As the Earth turns, everything in the sky (except some man-made satellites) will rise in the East and set in the West because the Earth is turning from West to East. Sunrise occurs when the part of the globe 90° to your east is pointed directly at the Sun, and sunset occurs when the part of Earth 90° to your west is pointed at the Sun, as shown below.
The Moon, stars, and planets all rise & set based on the same ideas shown above. When the Earth rotates around so that your Eastern horizon swings past them, they appear to rise, climb through the sky, and then set as your Western horizon swings away from them.

For example, in the picture below (taken from the North Pole) Savannah is represented by a red dot. The horizon is 90˚ away from Savannah in every direction. We can get a good approximation of Savannah’s horizon by pretending we are where the blue dot is (it’s easier to draw the horizon from there). In the top picture, the double-headed arrow represents the Eastern and Western horizons. As you can see, the arrow on the California side of the globe points to the West (where California is) and the arrow on the Atlantic Ocean side of the globe points to the East (where the Atlantic is from here).

Twelve hours later, the Earth has rotated around so that our horizons have switched around. In the top picture, the Moon was on Savannah’s Eastern horizon (so it was rising) and the Sun was on the Western horizon (so it was setting). After 12 hours, the bottom picture shows that the Sun is now rising (because it’s on the Eastern horizon) while the Moon sets in the West. Always remember: there is no rising and setting. It all boils down to the Earth turning and your horizon moving down to uncover some celestial object, and moving up on the other side of the sky to cover something else.
Conjunctions, Oppositions, and Quadrature – As we see the planets move through the sky, we notice that certain special arrangements are sometimes produced. These effects all depend on your position in the Solar system, by the way – a conjunction visible here could well be nothing special at all if you happened to be on Mars. When a planet is on the opposite side of the Earth from the Sun, we call that opposition. This can never occur for Mercury and Venus, because they are physically closer to the Sun than we are. There’s no way they could appear to be behind us (relative to the Sun) because we’re always outside their orbits. The Moon is in opposition when it is full.

A conjunction between two objects means they are both at the same Right Ascension. Notice that this doesn’t specify which thing is closer than the other or what’s going behind what. Also, if the second object isn’t specifically mentioned (“Mars is in conjunction tonight”) it’s assumed that you’re talking about the Sun as the other thing. The outer planets (also called superior because their orbits are larger than ours) can have only one kind of conjunction with the Sun, called a superior conjunction. This happens when the object moves behind the Sun. The inner planets (also called inferior because their orbits are inside ours and therefore smaller) can also have an inferior conjunction when they pass between the Earth and Sun. Therefore, we have three different kinds of alignments between objects and the Sun – opposition, inferior conjunction, and superior conjunction. The outer planets only have oppositions and superior conjunctions, while the inner planets only have inferior and superior conjunctions. Ask yourself about the rising and setting times (relative to the Sun) for planets which are in any of these three alignments.

An interesting question to ask would be “if Jupiter is at opposition tonight, how long until it reaches superior conjunction? Or its next opposition?” This doesn’t just depend on Jupiter’s motion. It also depends on the Earth’s motion. If Jupiter sat still, it would be 1 year between oppositions as Earth moved around its orbit and always found Jupiter in the same place. Jupiter doesn’t sit still, though, and it will have moved a little when Earth comes back around for the next opposition. After one year, Jupiter (sidereal period = 11.86 years) will have moved about 1/12th the way around the Sun (about 30˚). Earth has to keep chasing Jupiter a little longer to get back into position. Since Earth moves about 1˚ per day in its orbit, it can make up that 30˚ in about a month. Of course, Jupiter will have moved another few degrees by then, but Earth will make that up very quickly. All together, the time needed to go from opposition to opposition is (for Jupiter only) about one year + one month + a few days = 398.9 days, exactly. This amount of time is called Jupiter’s synodic period. It’s no coincidence that the Moon’s cycle of phases is called its synodic period – synodic means we’re worrying about something in relation to both the Earth and the Sun.

The time from opposition to superior conjunction is exactly ½ the synodic period. The same idea applies to Mercury & Venus – ½ of a synodic period between inferior and superior conjunctions. If the speed difference between two planets is large, the synodic period will be close to the sidereal period of the faster planet (this is because, for large speed differences, the slower planet almost sits still while the faster one
whizzes around in its orbit). For that reason, Pluto’s synodic period is 366.7 days (just a little longer than a year, because it’s barely moved at all during Earth’s year). Mercury (sidereal period of 88 days) has a synodic period of 115.9 days – it’s fast compared to Earth, but not very fast. The planets closest to Earth will, of course, have the largest synodic periods because it takes a long time to “lap” a planet when you’re not much faster than it is. Mars has a synodic period of about 780 days (more than 2 of our years).

There are other special positions. If a planet is 90˚ away from the Sun in the sky, it is said to be at quadrature (the difference between the planet’s RA and the Sun’s RA would be 6 hours at quadrature). The inner planets can never be at quadrature because that would mean they are more distant from the Sun than we are. Instead, they reach maximum elongation in their orbit when they appear to us to be as far from the Sun as they can get. This would obviously be the best time to observe the inner planets – the main reason they are so hard to see is that they spend much of their time close to the Sun (from our viewpoint). There is no simple relationship between the occurrences of these events (they don’t happen ½ synodic period apart, for example).

Mechanics and Historical Models of the Solar System

Greeks – The Greeks seem to be the first culture that made detailed models of the Solar System in an effort to explain the appearance of the sky around them. Aristarchus (300 BC) determined that the distance from Earth to the Sun must be many times greater than the Earth-Moon distance due to the fact that the 1st and 3rd quarter phases of the Moon are equally spaced in its cycle of phases. In other words, we will see the these two phases of the Moon only when the Earth-Moon line and the Moon-Sun line make a 90˚ angle. If the Sun was close to the Earth, the time between 3rd & 1st quarters would be less than half of the cycle as shown below. On the right, the Sun is very far away compared to the Moon and the red lines are very close to parallel. On the left, where the Sun is much closer, you can see that when the Sun and Earth are 90˚ apart as seen from the Moon, the Moon and Sun are NOT 90˚ apart as seen from the Earth. This shows that the time from 1st to 3rd quarter would be longer than the time from 3rd back to 1st.
Aristarchus was also able to estimate the relative sizes of the Earth and Moon by precisely timing lunar eclipses. The time from the first appearance of the Earth’s shadow on the Moon until the Moon was completely in the shadow was proportional to the Moon’s size. The duration of totality of the lunar eclipse (time for which the Moon is completely within Earth’s shadow) was proportional to the Earth’s diameter. Finally, using the existence of eclipses he determined that the ratio of the Moon’s diameter to its distance was the same as the ratio of the Sun’s diameter to its distance. By using something like the small angle approximation, he was able to guess at the Sun’s size. This estimate depends on accurately knowing the distance to the Sun, though, and since Aristarchus thought the Sun was only 1/10 as far away as it is, his estimate of its size was 1/10 of what it is. Also, all of these sizes/distances were measured in units of Earth diameters. No one knew the Earth’s true diameter, so the knowledge about the Sun & Moon was incomplete.

**Eratosthenes** – In about 200 BC, Eratosthenes measured the Earth’s size in traditional units for the first time. He did this by using the fact that the Sun was known to shine directly down a well in Syene (near the modern-day Aswan dam) on one day of the year (if this only happened once a year, what’s the latitude of Syene if we know that it is somewhere in the Northern hemisphere?). On that same day, he measured the angle of the Sun in his hometown of Alexandria. He got an angle of 7˚, and he knew that the distance between the two towns was 5000 stadia (1 stadium – probably about 0.1 miles). This was enough to give him the circumference of the Earth in stadia.

**Hipparchus** – Hipparchus was the first to record the phenomenon of precession and accurately find its period of 26,000 years. He was able to do this because he had made high-precision observations of many stars and compared the changes in his coordinates and those of Babylonian astronomers. He also developed the magnitude system.

**Other work by the Greeks** – The Greeks also deduced which planets were closer to Earth and which were more distant by looking at the speed with which the planets moved through the background of stars. They correctly assumed that the planets which moved more quickly were closer (except for Mercury) and the slower planets were further away. They noticed that Mercury and Venus were never far from the Sun, that the outer planets generally moved towards the East, and that occasionally, the planets would stop moving & seem to turn around briefly & move the other way before stopping and reversing yet again.

**Retrograde Motion** – one of the larger puzzles for ancient astronomers was the fact that the planets seem to move smoothly in the sky from West to East (relative to the stars! *Everything* that’s not man-made moves East to West relative to land) and then suddenly reverse direction and move from East to West for a short time until they reverse yet again and go back to their West to East motion.
We know now that the reason for this strange motion is that the inner planets move through the sky and overtake Earth in its orbit. As they close in on us & pass us, they appear to briefly move the “wrong” way through the sky. Once they are past us, things return to normal. The same thing happens with the outer planets, except that we overtake them. A picture from [http://www-istp.gsfc.nasa.gov/stargaze/Ssolsys.htm](http://www-istp.gsfc.nasa.gov/stargaze/Ssolsys.htm) is shown below

**Geocentric Universe** – There is obviously relative motion between the Earth and the sky – one night of observation will determine that. Seeing the planets move around the Earth (apparently) and the Sun and the stars do the same thing through the year, it’s a somewhat obvious choice to place the Earth at the center of the Solar system (= universe 2,000 years ago). The main argument against placing the Sun in the center (favored by Aristarchus) was the lack of measurable parallax for the stars. In other words, if the Earth is moving around the Sun, why don’t we see some stars shift their positions relative to more distant stars through the year? The answer is that the stars are so incredibly far away (compared to the size of the Earth’s orbit) that this motion is very hard to observe, as we saw earlier.

Because of a belief in the importance of symmetry & simplicity, the Greeks believed that everything in the sky could only move in perfect circles. Therefore, the layout of the universe was that the Earth was in the center, surrounded by planets (and the Sun) and the stars which moved around it in circles. There was a need to explain
retrograde motion, however, which lead to the concept of the planets moving in circles around a point in space which moved (in a circle) around Earth. The planet was theorized to move on an **epicycle** which then moved in a larger circle (the **deferent**) around the Earth. If the time to complete the trip around an epicycle is shorter than the time it takes the epicycle to move around the deferent, you can have retrograde motion. Finally, if the centers of Mercury & Venus' epicycles are forced to always remain on the Earth/Sun line, they will always be near the Sun and their motions can be explained. This system was fleshed out by Ptolemy in about 150 AD and is therefore called the Ptolemaic system.

**The Middle Ages** – During the Middle Ages (approx. 500 AD to approx. 1500 AD), Western science didn’t make much progress. Middle Eastern astronomers kept Ptolemy’s work alive and it remained the dominant theory of how the universe worked for over a thousand years. The contributions to astronomy from the Middle Eastern & Chinese scientists in this era were primarily observational.

**Copernicus** – In the early to mid-1500’s, Copernicus, who was familiar with Aristarchus' idea of a Sun-centered universe, began to try to fit the heliocentric model to the observed behavior of the sky. The heliocentric universe explained many observations more simply than the geocentric model. The planets could now be put in order by distance from the Sun, and that information both explained planetary regression and provided a constant increase in sidereal period with distance. The problem with the Copernican theory was inability to abandon the idea that planets could move in anything other than perfect circles. This idea had persisted since Ptolemy, and by keeping it, Copernicus had to add epicycles and deferents to his model to explain existing planetary observations. The new model needed even more epicycles than the Ptolemaic model. The general principle in science (called Occam’s razor) is that the simplest explanation that fits the facts is the correct one. More epicycles made this theory more complicated. The theory was published in Latin, a language known only among the educated, in general. By this time the Catholic church had picked up the geocentric theory as fact, and the church was not interested in alternative explanations. Since his work was not published until shortly before his death, Copernicus didn’t have to go through the same trouble later scientists would see.

**Tycho Brahe** – Tycho was an observational astronomer who did most of his work for the King of Denmark in the late 1500’s. As Royal Astronomer, he had the funding necessary to build observational instruments of very high precision. The telescope had not yet been invented, so observational astronomy consisted of getting the positions (coordinates) of stars and planets to the highest accuracy possible (about 30 arc-sec. for Tycho) night after night. Tycho also observed a brilliant supernova and realized it must be very distant from Earth due to its lack of a shift in position relative to the distant stars as the year went on.

**Johannes Kepler** – Kepler was Tycho’s assistant at the Royal Observatory and inherited Tycho’s highly accurate data. He spent years trying to find a curve that
would accurately fit the observed positions of Mars. Eventually, he discovered that an ellipse would fit the data very well. This was a giant change – for many centuries, everyone had assumed that only perfect circles would be involved in the motions of heavenly bodies.

Kepler’s greatest work would be his discovery of three “laws” of planetary motion. These laws were really fits to the data he had; there was no theoretical basis to predict them at this point, since the laws of physics as stated by Newton were still many years off. The laws which Kepler discovered (and Newton later explained) were:

1) The planetary orbits are ellipses with the Sun at one focus.

An ellipse is a kind of flattened circle, and every ellipse has two foci (plural of focus). Just as a circle can be defined as “the set of all points exactly x meters from a center point”, an ellipse also has a definition. It is the curve traced out when a string is tacked to a piece of paper in two places (not stretched tightly between them) and a pencil is used to pull the string tight in every direction. In other words, you can go to every point on an ellipse, and if you add the distance from that point to one focus plus the distance from the point to the other focus, you’ll get a constant number. Below, call the distance from A to 1 → A1, B to 1 → B1, etc. If things are drawn correctly, A1 + A2 = B1 + B2 and that will be true for any other point around the ellipse.

![Ellipse Diagram](image)

As the ellipse becomes more and more like a circle, points 1 and 2 move toward each other. A circle is just a special case of an ellipse where points 1 and 2 are the same point. For most of the planetary orbits (nearly circular) the foci are very close together.

2) The line joining a planet and the Sun sweeps out equal areas in equal times.

What this means is that a planet goes more slowly (in a well-defined way) when it is farther from the Sun than when it is closer to the Sun. Look at the figure below.
The triangle made by A, B, and the Sun has the same area as the one made by C, D, and the Sun. Therefore, according to Kepler's 2nd law, it should take a planet as long to move from C to D as it does to move from A to B. Since the A to B distance is much larger, the planet must move faster when it is closer to the Sun and slower when it is farther from the Sun.

3) The square of the period of a planet’s orbit (measured in Earth years) is equal to the cube of the planet’s average orbital radius (measured in AU’s). As a formula, \( P^2 = A^3 \) where P and A are the period and average orbital radius mentioned above.

This law appeared later than the other two. It’s obviously true for the Earth \( (1^2 = 1^3) \), but you can put in other numbers for the other planets. For Jupiter, for example, \( A = 5.2 \) AU. \( A^3 \) is then about 141. Jupiter’s year (sidereal period – NOT synodic) is 11.9 Earth years – \( P^2 = 141 \).

Why do the laws work? – These laws were derived from data – they describe what happens, but have nothing to say about why it happens. For that explanation, we move on to Galileo and Newton. Galileo was an early believer in the heliocentric hypothesis. He published his belief, and observations he had made with the telescope, in Italian rather than Latin, thereby making them accessible to a much larger number of people. This also enraged leaders of the Catholic church & Inquisition, who eventually placed him under house arrest and forced him to publicly accept the official view of the solar system rather than the Keplerian one. Galileo is widely credited with dispelling the ancient idea that heavier objects fall more quickly, using either logical arguments (1 heavy object = many light objects next to each other) or actual experiments. Among his astronomical observations were craters on the Moon, disks of the planets (rather than pinpoints of light, as the stars are), the Milky Way, the 4 largest moons of Jupiter, phases of Venus (proving that the Ptolemaic model, which would not allow phases > half, was wrong), and sunspots (probably why he died blind!).

Isaac Newton – Newton developed the laws of mechanics (classical physics) as well as calculus. In contrast to Kepler’s laws developed to match experiment, Newton’s laws are much more wide-ranging due to their basis in theory rather than experiment. There is NOT a direct correspondence between Newton’s three laws and Kepler’s
three laws – i.e., Newton’s 1\textsuperscript{st} law is not the specific basis for Kepler’s 1\textsuperscript{st} law. Instead, Newton’s laws provide a foundation for physics, the development of which proves Kepler’s laws and why they work. The three laws are:

1) A body at rest tends to stay at rest; a body in motion tends to stay in motion

This is in contrast to many centuries of thought (since Aristotle) that “rest” was the natural state of any object, and only the constant application of an external force could keep an object moving. This is a natural guess, because we live in a world full of friction; anything we encounter which is not powered or pushed by something \textbf{will} stop moving and come to rest. However, in a frictionless environment (space) it is just as natural for an object to move thousands of kilometers per hour (relative to something else) as it is for the object to sit still (relative to the same thing). Stopping \textbf{is} the unnatural thing – the application of a frictional force to a moving body.

2) Force = mass \times acceleration

This law gives the amount of force necessary to get a given mass moving with a given acceleration. It also shows that the same force will give a large object a small acceleration or a small object a large acceleration. As an example, some motorcycle engines can deliver the same force as small car engines. Since the motorcycle’s mass is probably only 20\% or so of the car’s mass, it will be able to accelerate \textbf{much} more quickly than the car. If you’d like to try this, you’ll probably find that the fastest car you’ve ever ridden in would lose a drag race to almost any motorcycle.

3) For every force one body exerts on another, the second body will exert an equal and oppositely-directed force on the first.

As an example, you’ve probably learned that you shouldn’t push someone while you’re both standing on a wet floor (or ice). You’ll both move away from each other. These forces will be equal, but the accelerations they will produce depend on the masses of the things involved (see 2\textsuperscript{nd} law). In other words, if you’re standing on ice & push a little kid, you won’t move much, but the little kid will go flying (not that you should try this with a little kid!). Try to push a car that’s stuck on the ice, however, and you’ll move a \textbf{lot} more than the car. Another example – if you fire a gun, the same force exerted by the on the bullet will be produced in the other direction on the gun. If it’s a shotgun, the kick in your shoulder will convince you of this quickly. Since the gun has a much greater mass than the bullet, it has a much smaller (but still noticeable) acceleration.

\textbf{Newton’s Law of Gravity} – Yet another of Newton’s laws. Newton postulated that the reason the Moon follows us around the Sun (for example) is that the Moon and the Earth attract each other through a force called gravity. He believed that the strength of this force depended on the mass of each body involved as well as the separation of the bodies. This is written below as
\[ F = \frac{G m_1 m_2}{r^2} \]

In this formula, the \( m_1 \) and \( m_2 \) are the masses of the two bodies (Earth and Moon, for example) and \( r \) is the distance between their centers (Earth-Moon distance). The \( G \) represents a constant (the same for all masses and distances), and as is often the case, the constant is just determined by experiment – there’s no way to derive theoretically what it should be. We know this force must exist if we believe Newton’s 1\textsuperscript{st} law (which Newton did!). If there are no forces acting between the Earth and the Moon, the Moon should fly off into space in a straight line.

**Weight vs. Mass** – Many people use these concepts interchangeably, but they’re actually different. Mass is a quantity that’s the same everywhere – here, space, the Moon, wherever. Weight is a measure of how strongly one mass (usually the Earth) attracts another one (whatever you’re trying to weigh). We tend to move freely back and forth between mass and weight, because we assume everyone knows that we mean \( m_1 \) is the Earth and \( r \) is the distance from the center of the Earth to the surface. Actually, mass and weight don’t even have the same units – they are as different as length and time. We convert back and forth between kilograms (mass units) and pounds (weight units) like they measure the same thing, but they really don’t. You can also think of mass as a measure of how many protons, electrons, and neutrons are in an object – it’s sort of a count, so it’s independent of where you are.

We know, then, that weight depends on having another body around where mass does not. The other thing that weight depends on is having somewhere to set your scale. For example, the astronauts in the space shuttle are weightless (not massless) not because there is no gravity, but because there is no “floor”. The space shuttle is falling out from underneath the astronauts just as fast as they are falling. We might not think of them as falling since they never “hit”, but that’s what’s happening. As an illustration, imagine firing a bullet horizontally. It will travel for some distance and then eventually hit the ground. As we increase the speed of the bullet, it will travel a greater distance before falling. If we keep making the bullet faster, the curvature of the Earth will start to be important; the ground will fall away from the bullet as fast as the bullet can fall! This is what orbiting a planet means.

**Other Physical Concepts** – In addition to ideas like mass, force, and acceleration, we also need to introduce the concept of momentum. Momentum is just the product of mass and velocity. A low-mass object moving quickly can have the same momentum as a high-mass object moving slowly. If there are no external forces on a group of bodies, the sum of their momenta (plural of momentum) is a constant. We can see this by realizing that it takes a force to change the momentum of anything (Newton’s 1\textsuperscript{st} law – if a force doesn’t act on a body, its velocity will be unchanged). If one of the objects in this group of bodies exerts a force on another one, it will have
an equal and opposite force exerted on it. One body will have an increase in momentum, the other will have a matching decrease. (It’s important to remember that momentum is a **directed** quantity – you can define a certain direction as positive, and that forces the opposite direction to be negative. Moving to the East requires positive momentum, and moving to the West requires negative momentum, for example). Momentum in a straight line is also called **linear momentum**. We usually use the letter \( p \) to represent momentum. We could write the formula for momentum as \( p = m v \) where \( m \) = mass and \( v \) = velocity.

The other kind of momentum we’ll discuss is called **angular momentum**. We can think of this as the momentum an object has when moving in a circle (it’s a little more complicated than that, but this is enough for our purposes). The formula for angular momentum is similar to that for linear momentum. We usually use the letter \( L \) to represent angular momentum. The formula for it is \( L = p r = m v r \) where \( r \) is the distance from the object which is moving to the center of the circle it’s moving in. As is the case with linear momentum, the angular momentum of an isolated system (one that’s not being acted on by external forces) is a constant. What this means is that if one of the quantities that make up \( L \) (\( m \), \( v \), and \( r \)) changes, the product of the other two has to change to keep the product of all three constant. The textbook example is a spinning figure skater. If the skater’s arms are out when the spin starts, the skater will spin much more quickly as she brings her arms in. In that case, \( r \) is getting smaller, so \( v \) has to get larger to compensate. When we look at the formation of the solar system, this will explain why a huge cloud of gas with only a very slow rotation will collapse (get a smaller \( r \)) into a smaller body with a very fast rotation (\( v \) increases to match the decrease in \( r \)).

**Newton’s Explanation of Kepler’s Laws** – Once Newton had developed his theory of force, momentum, acceleration, and the rest of classical physics, he was able to provide a theoretical basis for Kepler’s laws. By using his law of gravity and an expression for the “force” which seems to throw you to the outside of a curve as you move around it, he was able to find the formula below:

\[
(M_1 + M_2) P^2 = A^3
\]

This can be done with conventional metric units (meters, kilograms, seconds), but we’d have to add a few constants. We can actually use some units that are a little more appropriate for orbits. We got this by assuming that the force of gravity was keeping these two things in orbit around each other. Masses with significant gravity are usually many, many kilograms. If things are many kilograms in size, their orbits will be many meters in size (or else they’d hit each other). Gravity is weak, and it takes many seconds to complete these large orbits.

We can use better units. Here, \( P \) and \( A \) stand for period (years) and orbital radius (AU) just as in Kepler’s 3rd law. The two masses, \( M_1 \) and \( M_2 \), are the masses of the two things involved in the orbit – the Sun and the planet orbiting it, for example. Because we know Kepler’s law works for the planets, we can figure out what units we

30
must be using to measure mass. From Kepler, for the Earth we have $1^2 = 1^3$. We have masses on the left side only, and this equation already balances, so the sum of the masses must be 1. (If the sum was 2.5, for example, we’d have $2.5 \times 1^2 = 1^3$ which is obviously not true). The two masses involved for the Earth are just the Earth and the Sun. We know that the Earth’s mass is almost nothing compared to the Sun’s mass. To a very good approximation, we can just say that the mass is the Sun’s mass. The Sun’s mass is about a million times larger than the Earth’s, so we can effectively say that $(M_{\text{Sun}} + M_{\text{Earth}})$ is the same as $M_{\text{Sun}}$.

For comparison, an 18-wheel truck (fully loaded) weighs about a million times as much as the mail you get in one day! What this also shows is that Kepler’s 3rd law is incomplete if we are extremely accurate in our measurements. Since the mass of the Sun + Earth’s mass is not the same as the mass of the Sun + Jupiter’s mass, we shouldn’t get that $P^2 = A^3$ for both Earth and Jupiter. The reason Kepler’s 3rd law works so well is that, for all the planets, their mass + the Sun’s mass is still almost the same as the Sun’s mass alone. Even Jupiter has only 1/1000th the mass of the Sun. We also gain something else by this new formula. Newton’s laws are not derived from planetary data – we can use the result above for any two things in orbit about each other. For example, because Jupiter is so much larger than its satellites, we could make a table of $P^2$ and $A^3$ for all of Jupiter’s moons, and we’d see that the ratio of $A^3/P^2$ was essentially constant for all of them and that constant is a measure of Jupiter’s mass. Therefore, whenever a small body orbits a very much larger body, observations of the orbit can give you the mass of the large body!

Newton was also able to explain Kepler’s 2nd law. It’s done by using the fact that the planet and the Sun are essentially an isolated system. The planet moves in a circle (ellipse, really) around the Sun and therefore has angular momentum relative to the Sun. By the law of conservation of angular momentum, if the planet gets closer to the Sun ($r$ decreases), it must speed up ($v$ increases). That’s why the planet sweeps out equal areas in equal times. Newton also provided the basis for Kepler’s 1st law, but it is too complicated to show here.

Newton’s laws were a tremendous success, explaining the motions of the planets so well that, when Uranus was discovered in 1781, careful observation of its orbit revealed that something else was out beyond it & pulling on it. Using Newton’s laws, the location of the new planet was predicted and very shortly after that, Neptune was found in the predicted position.

We can notice a few things about the many satellites orbiting Earth. There are basically two important speeds when we’re talking about space travel from Earth. The first is the speed needed to completely escape Earth (we can never really get completely away from Earth because of gravity’s infinite range, but once we’re at the point where another body pulls more strongly on our space probe than the Earth, we can say we’ve escaped it), called escape velocity. We can find this by finding the kinetic energy necessary to completely overcome the potential energy our space
probe has on the launch pad (surface of the Earth). We can then find the velocity associated with this kinetic energy:

\[
\frac{1}{2} m \frac{v_{esc}^2}{R_{Earth}} = \frac{G M_{Earth} m}{R_{Earth}} \quad \Rightarrow \quad v_{esc} = \sqrt{\frac{2 G M_{Earth}}{R_{Earth}}}
\]

Notice that the space probe’s mass doesn’t appear in our final formula for escape velocity. Also, remember that this is from the Earth’s surface. If we had started from an orbiting platform thousands of km above the Earth, our potential energy would have been less negative, so we could have escaped with a smaller velocity.

We can also look at the speed necessary to put something in an ordinary circular orbit. In this case, the centripetal force is provided by gravity, so we can set the two equal as we did while deriving Kepler’s 3rd law and get

\[
\frac{m v^2}{r} = \frac{G M_{Earth} m}{r^2} \quad \Rightarrow \quad v_{circular} = \sqrt{\frac{G M_{Earth}}{R_{Earth}}}
\]

For satellites which have altitudes small when compared to the Earth’s radius (like the Hubble telescope and the space shuttle), we could get a reasonably accurate number for the velocity by plugging in \( R_{Earth} \) for \( r \). We find that a satellite in very low Earth orbit would need a speed of about 7900 m/s. Knowing the circumference of the Earth, we can find that the shuttle should take about 84 minutes to orbit at this speed.

Why are astronauts in the shuttle weightless? We know it’s not that there is a lack of gravity out there – Earth’s gravitational pull is almost as great where the shuttle is as it is here on Earth. The reason astronauts are weightless is because they are really **falling** around the Earth. The shuttle is basically a projectile, but it is moving so quickly that the Earth’s curvature becomes important as it travels. In fact, the surface of the Earth curves away from the shuttle exactly as fast as it falls, so that the shuttle is really just falling the entire time it’s in orbit. If the floor is falling out from under your feet, it can’t provide a normal force (there’s nothing for it to push on). Being in a falling elevator (if it were freely falling) would provide the same effect. If you’re ever in an elevator and you start to float, you’re about to have a bad day.

**Heliocentric Evidence** – Starting in the 1700’s, evidence for the heliocentric idea started to accumulate. This was in addition to the fact that models with the planets moving around the Sun in elliptical orbits were so much simpler. In 1729, stellar aberration was observed for the first time. This is an effect shown by a change in the direction a telescope must be pointed to see the same star at different times.
throughout the year. The telescope has to be shifted a very tiny amount in the direction of the Earth’s motion compared to the way it would be pointed if Earth were not moving. It’s because the Earth is moving as light travels towards it; it’s the same thing that happens when you try to run in the rain with an umbrella – you don’t hold it perfectly vertical if you want to stay dry. You have to angle it into the direction you’re moving.

It makes sense if you think of yourself as standing still and the raindrops as doing all of the moving. To make the picture match what happens when you run, the drops have to move not just down from the sky, but towards your face. The faster you run, the faster the drops seem to move towards your face. The velocity of the drops then becomes a combination of downward & towards you, which means they come in at an angle, so you have to tilt your umbrella. This effect for starlight amounts to only 20 arc-seconds during a year’s travel around the Sun. That’s why it took so long to be noticed. It also provided a measurement of Earth’s velocity around the Sun in terms of the speed of light (about 0.0001 times the speed of light).

The parallax we’ve already discussed is of course another piece of evidence.

**Evidence for Spinning Earth** – One obvious demonstration of Earth’s rotation can be made by imagining a pendulum (basically a rock on a string) swinging at the North Pole. Assume it’s swinging back & forth in the same plane as the 90˚ W longitude line and the 90˚ E longitude line. As time passes, the Earth will rotate under the pendulum, but the pendulum will keep swinging in the same plane. This can also be demonstrated at other latitudes, but it’s simplest at the poles.

Another piece of evidence was noticed by the German artillery in WW I. If a shell is fired (at high speed, of course) to the South or North of the gun firing it, it will seem to curve away from the target. This effect (called the Coriolis effect) has a simple explanation. All parts of the Earth rotate 360˚ in an hour, but different latitudes rotate at different speeds. For example, the equator is 40,000 km around, so the Earth moves at about 1700 km/hour there. If you make a circle about 4 km in radius and center it on the North Pole, points on it will only move 1 km/hour. Anything on the ground will be moving at the same speed as its local ground. If you fire from the Equator to a point directly North of it, the shell will already have a 1700 km/hour velocity from West to East just because of where it started from. The ground under it will have a smaller West to East velocity, so it will seem like the shell is moving both North (as fired) and East. In the same way, a shell fired from somewhere in the Northern hemisphere directly to the South will have a smaller West to East velocity than the ground under it. It will therefore seem to deflect towards the West. The way to remember this is to realize that in either case (Northern hemisphere), the deflection is to the shell’s **right**. The deflection in the Southern hemisphere is to the left.

**Tides** – One of the most obvious ways the Moon affects us on the Earth is the existence of tides. Every part of the Moon is pulling on every part of the Earth (and
vice versa) with a force that depends on the masses involved and the distances between them. We generally ignore the difference in the distances from the Moon to the near side of the Earth vs. the far side. Likewise, we ignore the difference in the direction to the Moon from two different parts of the Earth. We can’t do this when we talk about tides because tidal forces are differential forces.

These are forces remaining after the addition of two or more forces which are different in either direction or magnitude. If we imagine an ideal ocean which completely covers a spherical Earth, it’s easier to see what’s happening. The Moon raises two tidal bulges on the Earth – one on the near side and one on the far side. On the near side, we can explain the bulge by realizing that the water closest to the Moon is being pulled more strongly than any other water on the Earth, so it seems reasonable that a bulge should exist there.

These tides pull on the oceans as well as the Earth itself (solid-body tides) You can imagine the daily tides (two low and two high in most places) as being caused by the rotation of the Earth underneath these bulges pointing towards and away from the Moon. Similar (but much smaller) bulges appear in the Earth itself as a result of the solid body tides. Rotating under these bulges (or dragging them around the Earth if you look at it from the bulge’s point of view) tends to slow the Earth’s rotation on its axis. This slowing amounts to a loss of angular momentum (something with mass is spinning slower than it was). Since we can mostly consider Earth & the Moon to be an isolated system, that angular momentum has to go somewhere, and there’s nowhere left for it to go except into the Moon. This tends to increase the distance between the Earth and the Moon to make up for the decreasing angular momentum in Earth’s spin.

A similar effect long ago locked the Moon’s rotation so that its periods of rotation and revolution are equal. Many years in the future, the Earth will keep the same face towards the Moon at all times – at that point, the month and day will have the same length (47 of our current days!). There’s evidence in the fossil record of corals that the day was 22 hours long (so the year was 397 days long) 500 million years ago – the Earth really is slowing down. This is also why a “leap-second” needs to be added to the almost every year. The day is about 2 milliseconds longer than it was in the early 1800’s when it was exactly 86,400 seconds long. Those 2 ms per day amount to about a second over 365 days.
There are refinements to this simple story, however. First, the bulge does not stay directly under the Moon – it actually leads the Moon slightly because Earth rotates so much faster than the Moon revolves. The slowing of Earth’s rotation can be thought of as the pull of the Moon on the bulge. The pull of the bulge on the Moon moves it away from Earth. Also, the time of high tide depends very strongly on the local underwater topography. These effects can be so large that only one set of tides a day may occur rather than two. Around the Tybee lighthouse, high tide happens around when the Moon rises and when it sets. An additional complication is the fact that the Sun also affects tides on the Earth. It’s about half as effective as the Moon at raising tides, and we’ll see why in a minute. When the Moon and the Sun work together, though, the tides can be especially strong in each direction. When this happens, it’s called a Spring Tide. When the Moon and the Sun are pulling at a 90° angle to each other, those are called Neap Tides.

Now, back to the earlier question – why does the Sun have a smaller tidal influence than the Moon? We know tides are related to gravity, and gravity depends on mass, and the Sun’s mass is millions of times as great as the Sun’s. The reason that the Moon has so much stronger an influence is that the tidal force depends strongly on distance.

We can find the expression for this dependence easily enough. If we want to know about the tides raised on the Earth by the Moon, we can consider the Moon to be a point (zero size, but it keeps the same mass it really has. It sounds weird, but it’s OK mathematically, as long as we don’t care about the tides the Earth raises on the Moon). We then look at the gravitational force between the Moon and the side of the Earth nearest it, as compared to the force between the Moon and the side of the Earth furthest from it.

If we wanted to find the difference in the gravitational forces, we’d have to use Newton’s law of gravity for two distances – the first one would be from the Moon to the near side of the Earth (red arrow). Since we’ve drawn the center-to-center distance as \( r \) and the size of the Earth as \( x \), the distance would be \( (r - x/2) \). The distance between the Moon and the other side of the Earth (blue arrow) would be \( (r + x/2) \). The only question left is what masses we’ll use in the formula. One mass will be the Moon’s – the whole Moon is working on the Earth. It’s not so clear what the other mass is; we know the whole Earth isn’t at either place. We can dodge this question by looking at the tidal acceleration due to the Moon, which just means we’re going to
divide the formula for the force of gravity by the mass we’re not sure of. We’re left with this:

$$\text{Accel.} = \frac{GM_{Moon}}{\left(r + \frac{x}{2}\right)^2} - \frac{GM_{Moon}}{\left(r - \frac{x}{2}\right)^2}$$

You can put some numbers in this yourself. What you’ll get for each of these two terms is a very large number, and the difference between them is not very large. It’s hard to do math like this – your calculator is not good at trying to get very tiny differences between huge numbers. This is one of the benefits of calculus (we’re not going to do calculus). If we were going to use calculus, we’d find that the formula below is an easier version of the one above.

$$a = \frac{2GM}{r^3}x$$

In this formula, where M is the mass of the tide-raising body (Moon or Sun, if we’re looking at Earth), r is the distance between the centers of the bodies, and x is the diameter of the body on which tides are being raised (Earth’s diameter). The $r^3$ part means that the formula is very sensitive to distance – even though the Sun’s mass is so great, its distance from Earth is about 400x the Moon’s. If we cube 400, we get $400^3 = 64,000,000$. The Sun is only about 27,000,000 times the mass of the Moon. That means that the relative importance is $27/64 = \text{about 42\%}$. The Sun’s tides are therefore only about half the strength of the Moon’s. By the way, if you calculate the Sun’s density / Moon’s density, you also get 0.42. This is not quite the coincidence it seems (and the reason is related to eclipses!).

**Bibliography**

Discovering Astronomy (Shawl, Robbins & Jefferys)
UNIVERSE (Freedman & Kaufmann)
21st Century Astronomy, 4th ed., (Kay, Palen, Smith, and Blumenthal)
Astronomy Today, 8th ed., (Chaisson, McMillan)