THE PROCESS OF SCIENCE: THE PLANET PARIDIGM

Frank Munley, http://munley.weebly.com; fjmunley@comcast.net

The purpose of this essay is to illustrate the so-called scientific method by the history of astronomy, in particular how we came to our present knowledge of motions in the solar system. This history shows that the actual method of science is nothing like the parody taught in primary and secondary schools, which emphasize hypothesis followed by experimental testing, leaving little or no room for switching gears to focus on the novel surprises of nature that experiment can encounter.

Our story starts over 2000 years ago with the Greeks, even though a complete treatment would go back farther to the debt the Greeks owed to the very first serious astronomers, the Babylonians. We will investigate the great achievements of the Greeks and then see how their system was overthrown about 500 years ago. Science is a dominant influence in modern life, and any modern person claiming to be educated should understand how science works.

A. Greek Astronomy

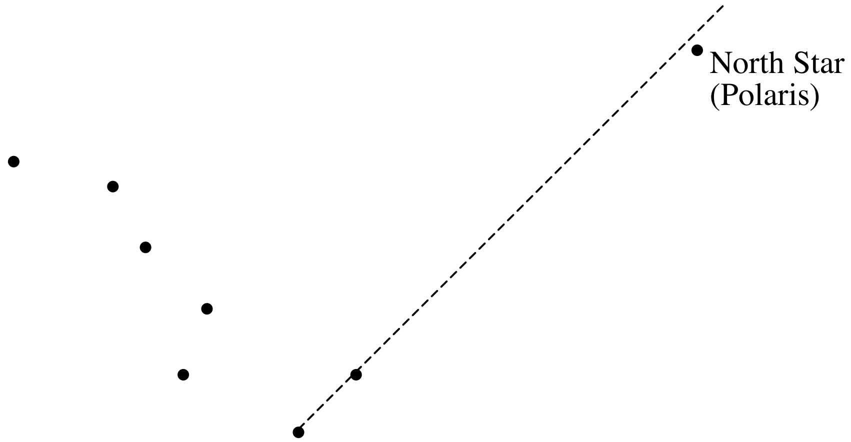
The ancient Greeks are renowned in the history of mathematics for, among other things, the development of geometry and the discovery of irrational numbers. The development of geometry received a particularly thorough treatment in Euclid's famous work in 300 bce, but its origins can be traced back several hundred years earlier. The rise of geometry went hand in hand with a scientific study of stars and their motions.

Just as the sun crosses the sky from east to west every day, so at night the stars cross the sky in the same direction, but they do not appear to move relative to each other. In any person's lifetime, indeed in a number of generations, the stars appear to stay in a fixed pattern relative to each other. Some of these fixed patterns are known as constellations. One off the best-known star patterns is the Big Dipper, the end stars of which point roughly to Polaris, the North Star. The Big Dipper and North Star are shown in Figure 1.

Very early in the stargazing game, the Greeks noticed that five "stars" wandered through the fixed stars, and they called them "planetes" which means "wanderer" and from which our modern term "planet" comes. These five planets, the only ones visible to the naked eye, are Mercury, Venus, Mars, Jupiter, and Saturn. They wander against the background of fixed stars, moving in a roughly west to east motion. There is an additional complication with the wanderers, called "retrograde motion," which will be discussed shortly. The wandering motion is rather slow, being detectable to the naked eye only if the planets are viewed night after night. (This is not to be confused with the entire sky's more rapid daily East to West motion relative to Earth's horizon, which results from Earth's rotation.)

The moon wanders too, but much more rapidly than the "planetes." The sun also wanders, but it and the moon are so much bigger than the stars and planets that they were considered to be a separate kind of thing. The most rapid wandering is carried out by the moon, which in two weeks' time will travel completely around a circle through the visible fixed stars.

We know today that the stars are not really fixed, but they have to be observed for many thousands of years for the motion to be noticeable. Consequently, the Big Dipper stars will no longer be in the shape of a dipper thousands of years from now, and Polaris will no longer be in the north direction.



<u>FIGURE 1</u>: The Big Dipper and North Star. Night after night and year after year, the Big Dipper retains its shape and its geometric relationship to the North Star. The end stars of the Dipper can be used as pointers to the North Star which is actually a slight bit off the guide line, as shown. We know today that stars have motions, and many thousands of years from now the Big Dipper will not have the shape it does now.

B. <u>Accomplishments of the Greeks</u>

The advance of geometry allowed the Greeks to discover many important things in astronomy. For example, they realized that Earth is a sphere, based on the following three pieces of evidence:

1. Ships disappear gradually as they get farther from shore—first the bottom of the boat disappears, then the body of the boat, leaving only the sails in view; and finally, the sails disappear too.

2. The sun's highest elevation (near noontime) is lower the higher the latitude, as shown in Figure 2.

3. The shadow of Earth on the moon during a lunar eclipse is always curved in a circular shape.

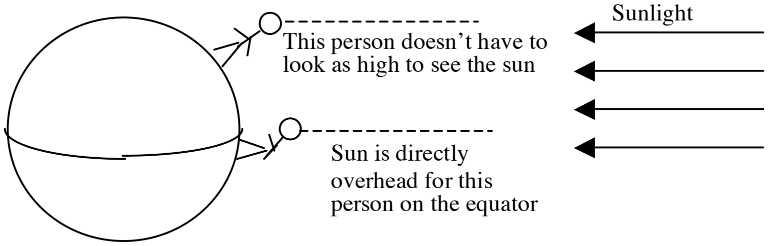


FIGURE 2: At the same time of day, a person farther north sees the sun at a lower elevation than a person directly south.

Aristarchus of Samos (c. 310-230 bce) used geometry to estimate the relative distance of the sun from Earth and the radii of the moon and sun. His estimates were rough due to the difficulty of measuring very tiny differences between angles. Letting d = distance of the moon or sun from Earth and R = radius of Earth, moon, or sun, here are his main results compared with reality:

1. $d_{sun} = 20d_{moon}$, but it is actually much bigger than the moon, being $390d_{moon}$.
2. $R_{sun} = 6.7R_{Earth}$. Again, this is a serious underestimate, the actual value being $100\ R_{Earth+}$.
3. $R_{moon} = 0.351R_{Earth}$, a slight overestimate because the moon is really $0.273R_{Earth}$.

In spite of the inaccuracies of these estimates, Aristarchus showed us how big the solar system is and that the sun is much bigger than Earth.

Shortly after Aristarchus's work, Eratosthenes (c. 276-194 bce) succeeded in measuring the size of Earth, so together with Aristarchus's work, the ancient Greeks had a rough idea of how far away the moon and sun are. Aristarchus's estimates involve complicated geometry, but we can understand the conceptual basis of Eratosthenes' estimate of Earth's radius from Figure 3.

In Eratosthenes' time, the Greeks controlled Egypt and Alexandria was Egypt's main city. Eratosthenes knew that on the first day of summer, when the sun is highest in the sky, the sun was directly overhead the Egyptian city of Syene, i.e., the sun's rays come down vertically so a stick held vertically in Syene would have no shadow. At Alexandria, 800 km to the north, on the other hand, the sun's rays come in 7 degrees to the vertical so a stick held vertically does have a shadow. This is illustrated in Figure 1.3, which illustrates how the length of the shadow depends on the radius. By working through the geometrical relationship between the length of the stick shadow and the distance between Alexandria and Syene, Eratosthenes was able to determine the diameter of Earth.

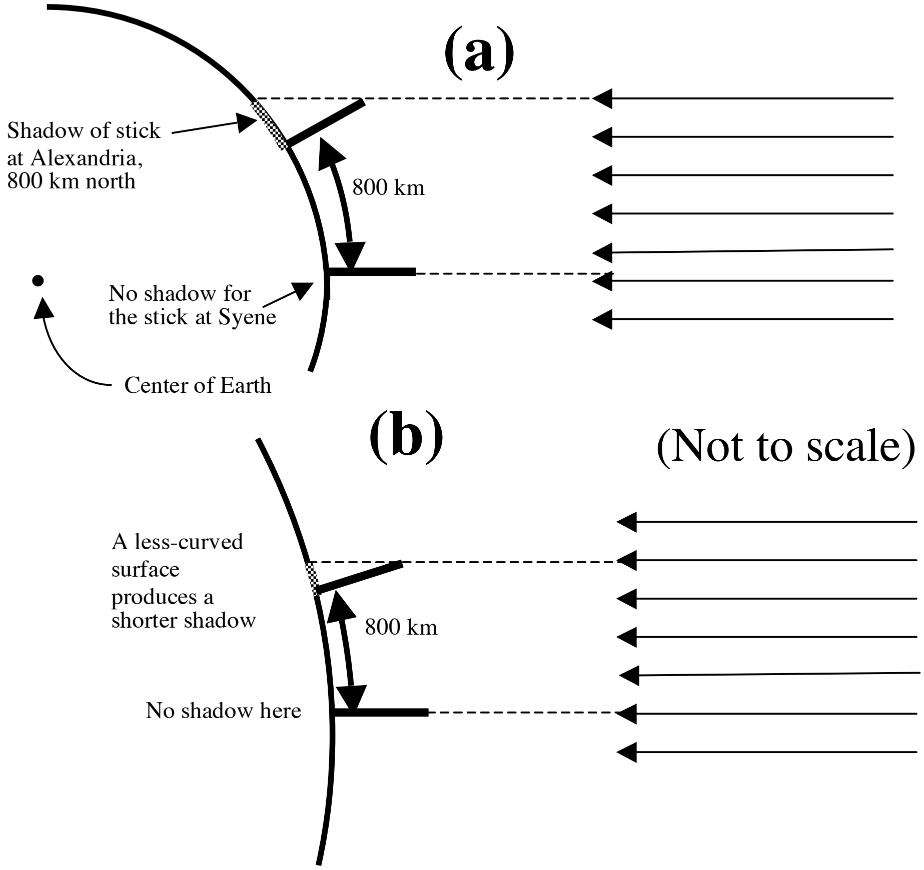


FIGURE 3: (a) Since Earth is round, the length of the shadow of a vertical stick at a given time of day depends on how far north or south it is. (b) Compared to (a), a less-curved Earth would result in a smaller shadow to the north or south. A flat Earth would produce the same length shadow everywhere.

C. Greek cosmology and planetary problems; retrograde motion

Every day we see the sun travel across the sky; the sun seems, to our senses, to be obviously in motion. It also seems obvious to the senses that Earth is at rest; we can't feel it moving at all, and if it were rotation, shouldn't we get dizzy? These and other "commonsense" observations led the Greeks and many before and after them to assume that Earth is at rest while the rest of the universe rotates around Earth. These assumptions may be obvious to the senses, but as we will see, they are incorrect.

Just looking at the sky, it is impossible to say much about the actual arrangement of the stars. The Egyptians believed the stars were arranged on a flat surface above Earth, but the Greeks had a different view. If the stars in the north are observed through the night, they are seen to travel in circular paths about Polaris. This suggested to the Greeks that the stars are arranged on a spherical surface which they called the "Celestial Sphere." The celestial sphere rotates about an axis passing through Earth and through Polaris, with Earth at the center of the celestial sphere.

Aristarchus's estimate of the distance of the sun from Earth showed that the stars must be much, much farther away compared to Earth's diameter than the sun. Figure 4 proves this in exaggerated form. Part (a) shows the case of a celestial sphere which is much larger. Then in this case the angular separation between two stars will be the same for any two observers on Earth. Compare this with part (b) of the figure, where Earth's radius is just a bit smaller than the radius of the celestial sphere. Now the angular separation between two stars is different for different observers, as shown. We know today that the angular separation between stars is much too small for the Greeks (who did not have telescopes) to detect, so (a) represents the situation, not (b), i.e., the celestial sphere, whatever its exact size, is much, much bigger than Earth. If the stars were only as far away as the sun, then the angular separation between two stars would be measurably different for different observers.

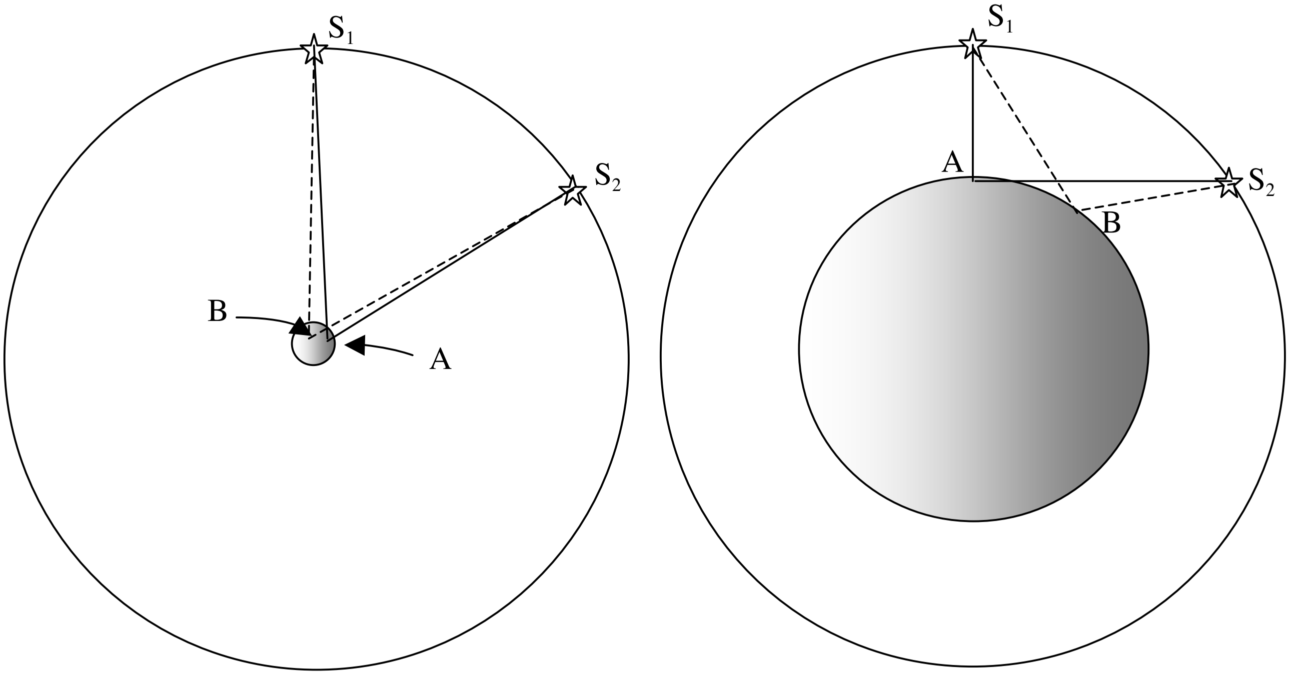


FIGURE 4: Two observers, A and B, measure the angular separation between two stars, S_1 and S_2. (a) If Earth is very, very small compared to the distance of the stars, A and B measure the same angle. (b) When Earth is comparable in size to the celestial sphere, A and B measure different angles between the same two stars.

The Greeks believed that stars are non-earthly in nature, made of a perfect, simple, pure substance called ether. A perfect "ethereal" substance should undergo perfect motion, and the "perfect" path of choice was considered by Greeks to be the circle. The choice of the circle was prompted also by the celestial sphere model mentioned above, in which stars undergo circular motions about Polaris. Circles are very elegant from a geometrical standpoint, since all points of the circle are the same distance from the center. All of these considerations led the Greeks to believe that circular motions were very special, and very appropriate for perfect ethereal objects. The sun and moon, unlike the stars, don't trace out circular paths about Polaris, but they do follow a circular path through the stars, one circle being completed in 27 days by the moon and one year by the sun. So in the Greek view, the moon and sun are both ethereal, unlike Earth.

The motions of the planets are more complicated because for short periods of time they reverse their direction of travel through the fixed stars, which is known as **retrograde motion**, as shown in Figure 5. Furthermore, unlike the sun, they vary significantly in brightness over months. The celestial sphere model fails to account for these facts, because the model assumes each star, and planet too, just goes around with one circular motion. Furthermore, the celestial sphere model fails to explain why the planets vary in brightness, since they should be on the celestial sphere at a fixed distance from Earth. These failures of the celestial sphere model required it to be modified. The most significant modification came from Appolonius of Perga (c. 225 bce), but he still used circles. His "epicycle model" is illustrated in Figure 6 with the planet Jupiter as an example.

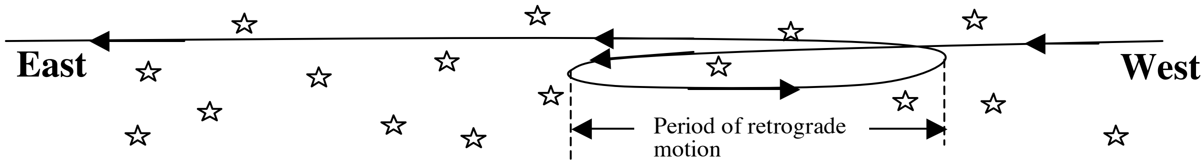


FIGURE 5: A planet's motion is shown to be generally from West to East, but for a period of time it reverses motion and goes westward relative to the background "fixed" stars.

Appolonius's epicycle model was refined significantly by the most famous of Greek astronomers, Ptolemy (c. 150 ce), 400 years later, **so we will refer to the epicycle model as the "Ptolemaic" model or system**. In the Ptolemaic model, a planet like Jupiter travels around a circle called an epicycle, but the center of the epicycle itself travels around a larger circle, called the deferent, centered on Earth. The combination of these two motions causes the planet to trace out a "loop the loop" or "circle within a circle" motion, as shown by the dashed and heavy solid line in Figure 6. (The inferior planets, Mercury and Venus, are not shown for simplicity.) The heavy solid part represents the period of retrograde motion.

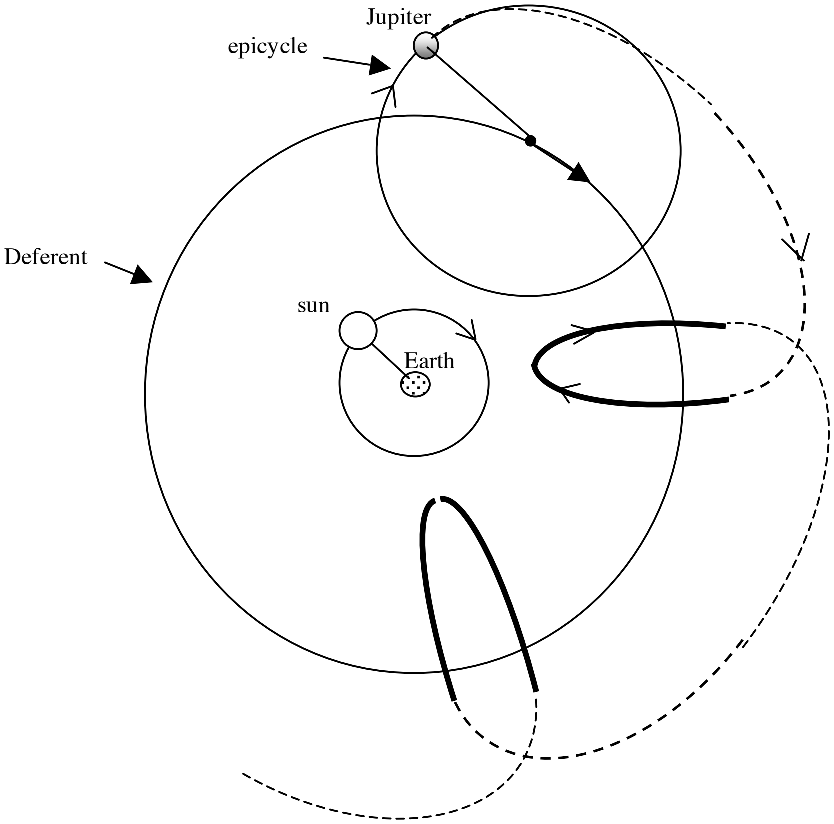
The epicycle model also changes the distance between Earth and planet, so the Greeks believed that this explains changes in the planet's brightness. Today, we know how brightness depends on distance, and in fact the Ptolemaic model really doesn't explain the observed changes in planet brightnesses.

D. Trouble with inferior planets: Aristarchus's challenge to Aristotle

The Ptolemaic model shown in Figure 6 works only for the "superior" planets of Mars, Jupiter, and Saturn. Note that they move rather independently of the sun, so that at times these planets can be in a completely different part of the sky, far from the sun as seen from Earth. However, the two "inferior" planets, Mercury and Venus, always stick close to the sun. Consequently, these two planets couldn't be represented by the simple model of Figure 6. Figure 7 shows the scheme Ptolemy imagined for the inferior planets; also shown is the moon. As with the superior planets, the two inferior ones go around in epicycles, but the centers of the epicycles are stuck on the line joining Earth and sun. By pegging their epicycles to this line, the planets are always close to the sun. A peculiarity of this system is that the orbit of Venus goes right through the sun. This didn't concern the Greeks, because the sun and Venus, being ethereal, did not have to behave like ordinary earthly objects.

It was just mentioned that the superior planets move "rather" independently of the sun. Actually, they are also constrained a bit, so that the lines joining the planets to the centers of their epicycles is always parallel to the line joining Earth and sun, as shown in Figure 6 for Jupiter. We might say that the superior planets are "peasants" of the sun.

Figure 6 shows that the superior planets do reverse their direction of motion, i.e., the epicycle idea explains retrograde motion. It also shows that their distances vary, being closer to Earth and hence brighter when they are undergoing their most rapid retrograde motion. This is exactly how the planets behave, so the Greeks took the epicycle idea very seriously.



<u>FIGURE 6</u>: Appolonius's model: Jupiter's deferent and epicycle. This view looks down on Earth's north pole. The motions are generally clockwise, so the sun rises in the east and sets in the west, in conformance with reality. The diagram also shows the sun's orbit. This is a simplified version of the system of Ptolemy.

The need to tie the epicycles of Venus and Mercury to the line joining Earth and sun, thus making them in a sense "slaves" of th sun, cries out for explanation, but the

Greeks couldn't give one. They couldn't avoid treating the superior and inferior planets differently. This was a serious problem with the Ptolemaic model, because a good model should treat all of the planets in the same way. In addition, the moon's motion was all wrong. This serious problem will be discussed later.

Our old friend Aristarchus in fact did propose a model that treated all planets the same way. Aristarchus assumed that the planets and Earth too revolved around the sun, which sits stationary at the center of the circular planetary paths. He was led to this model when he calculated the distance of the sun from Earth. The only way the sun could appear as large as it does even at such a great distance is if the sun is much, much bigger than Earth. It didn't make sense to Aristarchus that a very large object should orbit a much smaller one, so he put the sun at the center. We know today that Aristarchus was perfectly correct that the planets (including Earth) revolve around the sun, but his views weren't taken seriously—he was too far ahead of his time. To understand why this eminently reasonable idea was rejected, we have to understand something about Greek philosophy.

The greatest Greek philosopher was Aristotle, who lived about a hundred years before Aristarchus. Aristotle believed that everything on Earth was composed of four basic elements: earth (e.g., rocks, dirt, and living things), water, air, and fire. Celestial objects, on the other hand, were made of a fifth element, ether, the pure and simple substance mentioned above. Each of these elements has a "natural" motion it tries to achieve. Earth's natural motion is towards the center of the cosmos, and this is why Earth is at the center of the celestial sphere. Hold a hunk of earth, e.g., a stone, up and let it go. It executes its natural motion by falling towards Earth's center, i.e., the center of the universe. Of course, a stone can undergo "forced" or "violent" motion if it is thrown, but Aristotle believed that the stone moves only as long as the original "impetus" of the throw continues to act on it. In Aristotle's view, when you throw a stone, the force you put on it continues to act on the stone even after it leaves your hand, and as it gradually diminishes, the stone falls "naturally" to Earth.

According to Aristotle, it is impossible for Earth to rotate on its axis or revolve around the sun without slowing down, because it would need a force to keep it moving. Since there is no conceivable force acting on Earth, Earth must be at rest. Note that these arguments against Aristarchus's basically correct idea cannot be applied to the planets, because they are made of ether and don't follow the same laws as Earth does.

E. <u>Ptolemy's further refinement of Greek astronomy and his failures</u>

One of the motivations behind Greek astronomy was to set up a calendar and time system which all Greeks could agree on. Prior to the development of astronomy, different Greek cities had their own calendar and time systems. By basing systems on star motions, everybody could agree on the time of the year. This project required that the motions of the stars, sun, moon, and planets could be accurately predicted. So over the centuries, Greeks "tweeked" the epicycle model by adding even more epicycles to each planet's motion.

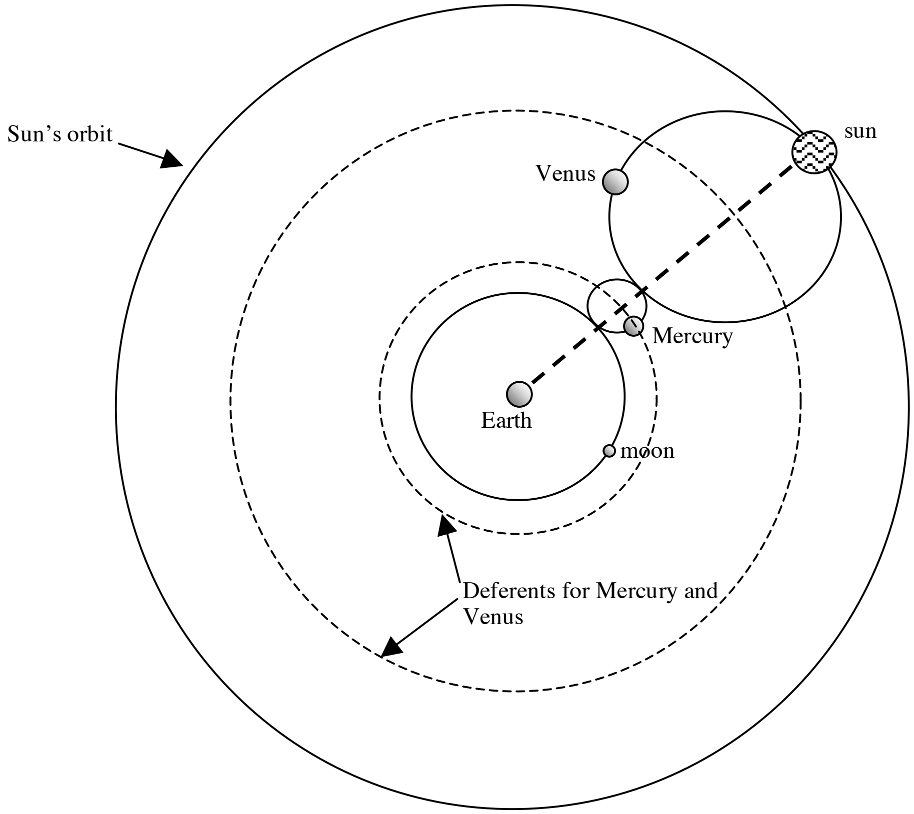


FIGURE 7: A simplified version of the system of Ptolemy, showing deferents and epicycles for the two inferior planets. Also shown are the deferents for moon and sun, which themselves have epicycles (not shown here) in the more complete version. Also, Ptolemy's "equant" points are not show.

A very important improvement was made by the most famous of Greek astronomers, Ptolemy, who was mentioned above. Ptolemy still assumed motion in circular paths, but his great innovation was to assume that the motion along the circle varied in such a way that it only **appeared** to have a constant motion at a special point called the equant. We shall not go into the details of Ptolemy's theory except to say that in its full-blown form, it was the most accurate predictor of celestial motions, given that positions of objects in the sky could be measured at best to about 0.2 degree. To get a feeling for what $0.2\,^{o}$ is, consider the fact that the moon has an angular width of just about $0.5\,^{o}$. Figure 8 shows the situation.

Ptolemy's model of the solar system and the cosmos must rank as one of the great achievements of the human mind. His attempt to describe the motions of stars and planets as seen from Earth worked admirably at the time, but we know today that his system is based on a number of major assumptions that are incorrect. His incorrect assumptions are as follows.

1. Earth is at rest neither rotating on its axis nor moving through space;
2. The motions of celestial bodies are combinations of circular motions; and
3. The inferior planets, Venus and Mercury, behave in a completely different way than the superior planets.

In addition, we must add another serious drawback in his model:

4. His orbit for the moon, required to account for its motion, makes the moon's distance from Earth vary by a factor of about two, so that the moon's diameter as seen by someone on Earth should also vary by a factor of about two.

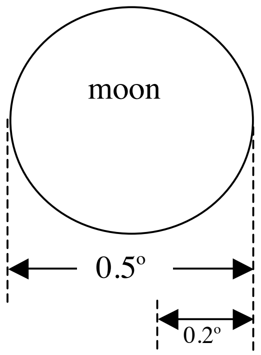


FIGURE 8: The angular width of the moon is 0.5 °. In Ptolemy's time, positions of objects could be measured to about 0.2 °.

This is certainly not the case, because the moon appears at all times to be about the same size. But if all we're interested in is how celestial bodies appear to move as seen from Earth, and if we are willing to accept a bit of inaccuracy in our predictions of where, say, Mars will be 5 years from now, then these errors are harmless. Consequently, the great Greek system of an Earth-centered universe, perfected by Ptolemy, lasted 1400 years.

F. BACK TO ARISTARCHUS: THE COPERNICAN REVOLUTION

Polish astronomer Nicholas Copernicus (1473-1543) was not just interested in the motions of celestial bodies, but also in the simplest way of describing their motions and their appearances too. His primary motivation was to come up with a model which was based on the real behavior of real bodies in the universe. For this purpose, Copernicus re-invented Aristarchus's model of a solar system and universe with the sun at the center and Earth in motion about it. But scientific revolutions oftentimes occur step-by-step, and Copernicus's revolution is no exception because he held fast to the Greek assumption of circular motion. In his model, all the planets follow circular paths, but the sun is not at the centers of these circles. The moon, as in the Ptolemaic system, orbits Earth, but doesn't change its distance by a factor of two so it stays pretty much at the same distance, i.e., its size as seen from Earth stays pretty much the same which is in fact the case.

Copernicus did not do away completely with epicycles, because he needed them to match the accuracy of Ptolemy's model. **But Copernicus's epicycles are much smaller than those of Ptolemy and are, therefore, a minor addition to Copernicus's simple circles.** In comparing Copernicus's model with Ptolemy's, it is therefore useful to distinguish between the **simple Copernican model**, shown in Figure 9, and the more **complicated Copernican epicycle model**. By focusing on the merits of just the simple model, we can better understand why Copernicus's idea of an Earth in motion about a stationary sun finally won the day.

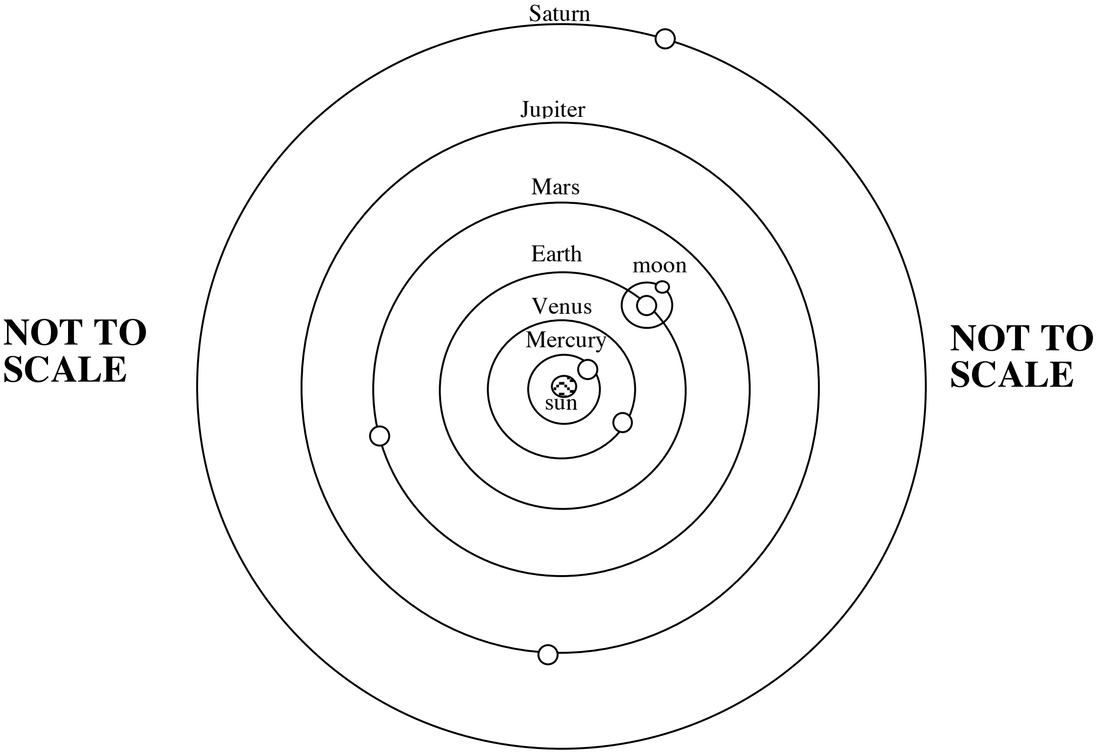
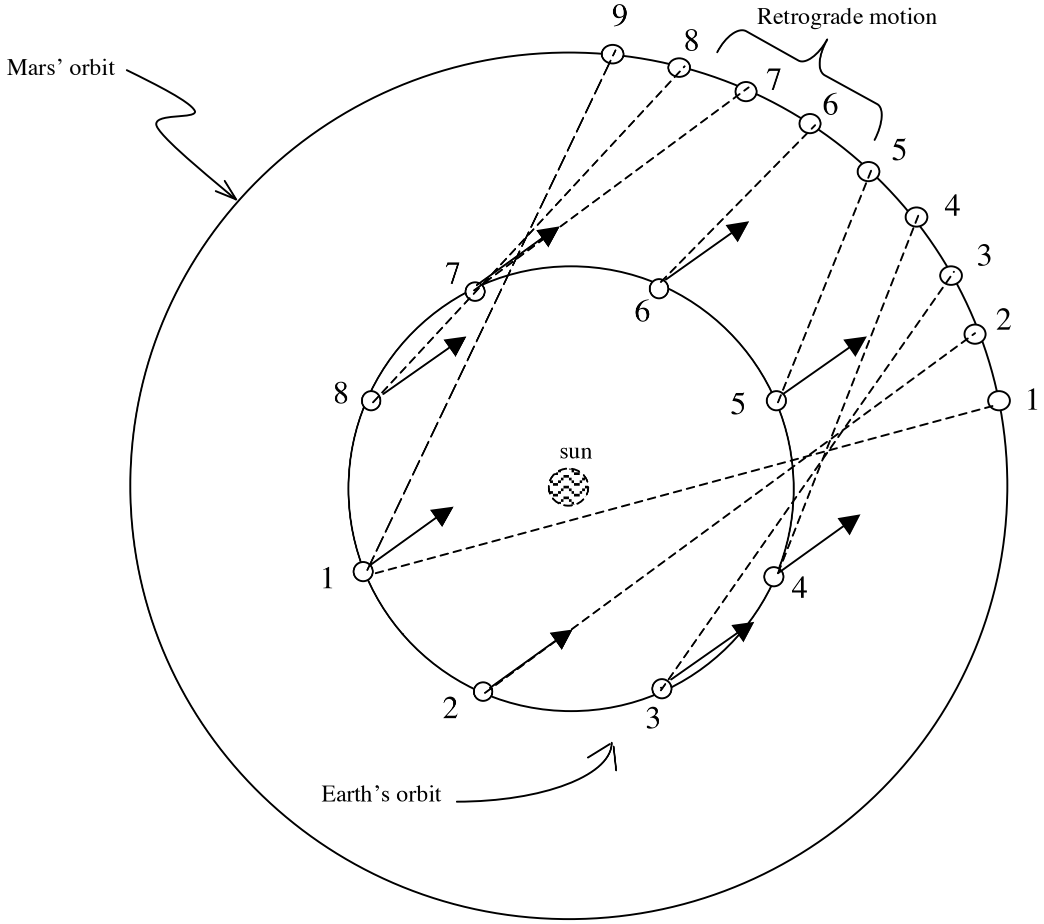


FIGURE 1.9: Copernicus's simple model of the solar system, which has the sun at the center and the planets in circular orbits.

Consider Copernicus's explanation of retrograde motion and variations in planet brightnesses. In his simple model, the bigger the orbit of a planet, the slower it moves. Figure 10 shows how Mars would be viewed from Earth under these circumstances. In this view, we are looking down on Earth's North Pole. The figure shows Earth at 8 equally spaced times as it makes a complete orbit. For each position of Earth, a dashed line from Earth to Mars indicates the corresponding position of Mars. Earth and Mars are both seen to orbit counterclockwise. Also shown for each Earth position is an arrow pointing in a fixed direction in space, say to a particular fixed star in space.

Just looking at position 1 of Earth, we see the two positions of Mars at the beginning and end of Earth's complete orbit. Clearly, Mars has drifted to the left relative to the reference direction over a year's time, i.e., it is drifting counterclockwise. So Mars' normal direction of motion is counterclockwise, and any motion back to the right is, by definition, retrograde motion. Now look at positions 4 and 5. Mars is just about in the same direction at these two point, i.e., it has stopped its normal motion and is about to begin retrograde motion. Comparing positions 5 and 6, we see that Mars has indeed swung a bit clockwise relative to the reference direction. This clockwise swing persists until we compare positions 7 and 8. So we can say that the retrograde motion takes place between positions 5 and 8.

Note that the period of retrograde motion occurs when Earth is closest to Mars. Looked at another way, Mars drifts off to the right, i.e., its direction from Earth relative to the distant fixed stars swings clockwise, because Earth is moving faster and between these times it is overtaking Mars. It is just like driving on the highway and passing another car, which will seem to bo backwards not just relative to you but also relative to a distant background. Copernicus himself used this motion analogy (but of course he didn't phrase it in terms of cars, which hadn't been invented yet).



<u>FIGURE 10</u>: Orbital positions of Earth and Mars over time in Copernicus's model. The solid arrows all point to a particular direction in space, e.g., to a particular star. The diagram shows that retrograde motion occurs roughly between positions 5 and 8.

G. Is Copernicus's model really so much better than Ptolemy's?

Given what was known in Copernicus's time, what made it appealing? Definitely not its central assumption that Earth revolves around the sun. This seems quite appealing to use today, because we're taught that that is the case, but it was a most controversial assumption at the time. The appeal was, rather, the consequences of this assumption. In both models, Mars (or any other superior planet) is closest to Earth when undergoing retrograde motion. So they both require that Mars should always be brightest when undergoing retrograde motion, and there is a draw between the two on this point. But Copernicus made a great improvement by coming up with a scheme which abolished "slavery" for Mercury and Venus and treats all planets the same.

Brightness of the moon is also a deciding factor between the two competing models. We mentioned above that Ptolemy's model has the distance between Earth and moon varying by a factor of 2, which means that the moon should get bigger, and therefore brighter when it is full at the same time as its size is twice the minimum. This is not observed, of course; the moon's size barely changes throughout time, and so the brightness of the full moon is almost perfectly constant

Finally, Copernicus's simple model closely approximates Ptolemy's much messier model, and the addition of small epicycles serves to make the former equivalent to the latter. . For all these reasons, we could look more favorably on Copernicus if we were alive at that time.

H. MORE EVIDENCE AGAINST PTOLEMY: GALILEO'S FINAL "NAIL IN THE COFFIN"

Galileo Galilei (1564-1642), the eminent Italian scientist about whom we will hear more of later when we discuss the modern theory of motion, built himself a telescope and discovered many amazing things about celestial objects. Here are three crucial discoveries he made with his telescope which were incompatible with the Ptolemaic view of the universe.

1. The sun has spots. This means it is not perfect as celestial bodies are supposed to be, but has blemishes.

2. Jupiter has moons revolving about it, so not everything revolves around Earth.

3. Venus shows a full phase; this is impossible in the Ptolemaic model.

The first two discoveries strike at the philosophical assumptions surrounding the Ptolemaic model. After all, if Jupiter and the sun, and probably other celestial objects, are nothing special, maybe they are made of the same stuff as Earth. But if the importance of Ptolemy's model is a description of the motions of the planets and stars, then (1) and (2) could be accepted without throwing out the Ptolemaic model. It is the third discovery which is totally incompatible with Ptolemy's model and which represents the "final nail in the coffin." As stated, a full phase of Venus is impossible in the Ptolemaic scheme, and this is shown in Figure 11. This figure, like Figure 7, shows that

Venus's orbit goes right through the sun. When it is "inside" the sun, it can't be seen, so this special position of Venus would **not** correspond to any phase at all. So at no point will Venus show a full phase to someone on Earth. Figure 12 shows how Venus **can** have a full phase in Copernicus's model. So Copernicus wins hands down on this point.

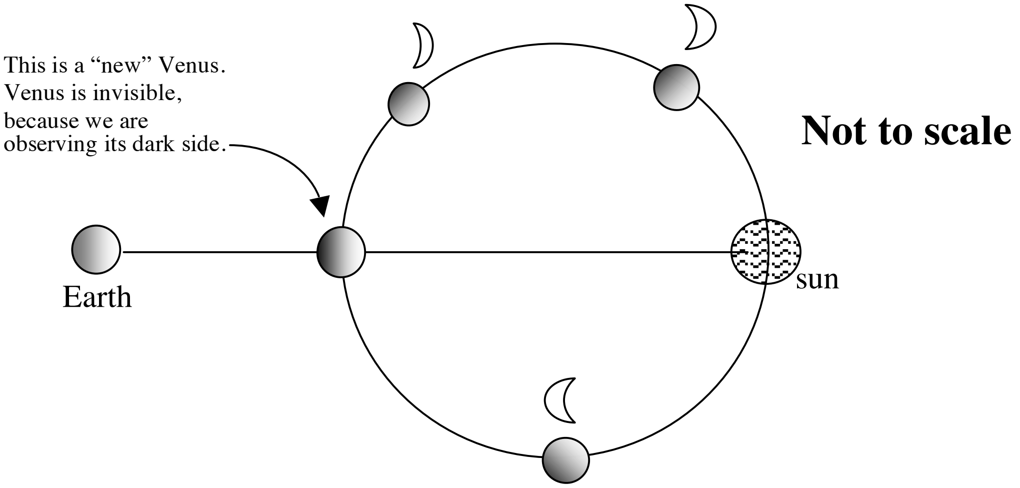


FIGURE 11: The possible phases of Venus in the Ptolemaic model. Venus can show crescent or new phases, but never a full phase. The crescent shapes show how Venus would appear to an observer standing on Earth and oriented perpendicular to the plane of this sheet of paper.

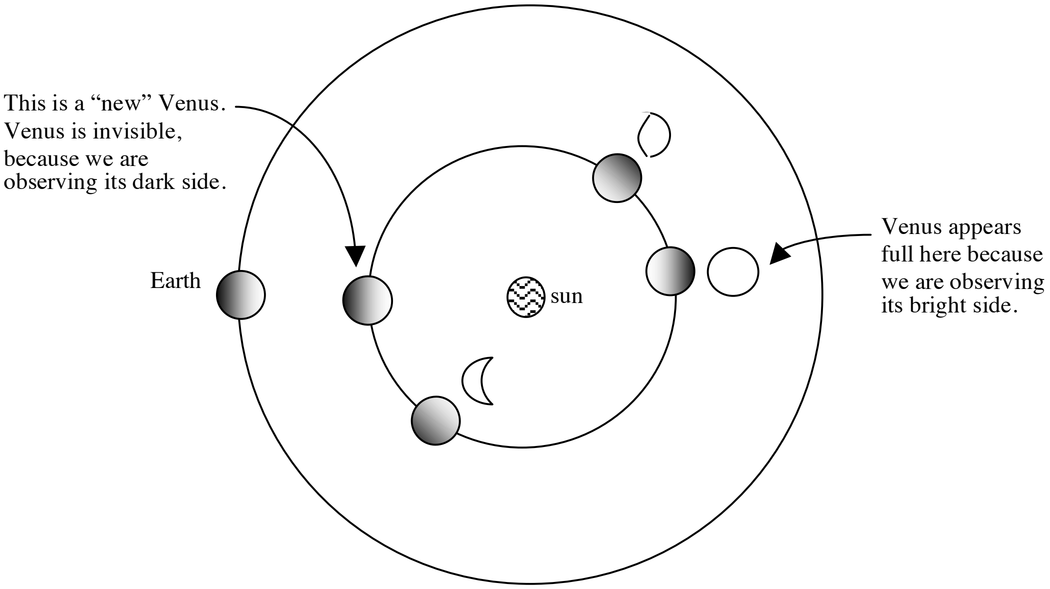


FIGURE 12: The phases of Venus in Copernicus's model. Now a full Venus is possible when Venus is on the other side of the sun and close to it.

I. Does Earth really have to move?

While Ptolemy's model is definitely wrong and Copernicus's model seems much better, we can always wonder if there isn't some way to explain things assuming Earth is really at rest, as our senses seem to suggest. In fact, it is possible to avoid the "slavery" of the inner planets, and construct a consistent and highly accurate model of planetary motion with Earth at rest if we use sufficiently complicated motions for the planets. But such an Earth-centered model would not be nearly as simple as the Copernican model. Furthermore, as we will see shortly, there are a number of experiments (see Section S, item (3) below) which are best explained on the presumption that Earth rotates on its axis and revolves around the sun. We could build an entire science of mechanics, optics, etc. to accommodate an Earth-centered solar system, but the result would be incredibly messy. It is much simpler to give up the idea of Earth at rest and go with Copernicus. In other words, like many things in life, we ascribe reality to the planetary model that makes the most convenient and simple sense of the world around us. Simplicity is of central importance of in ascribing reality to our beliefs.

J. What did Copernicus accomplish for the future of science?

We have already mentioned Copernicus's achievements for understanding planetary motion, but there are deeper philosophical implications of his work which helped set the stage for modern science. Specifically, he put Earth with the heavens by showing that it is part and parcel of the "ethereal" planets formerly considered to be of a completely different nature than Earth. From another perspective, he brought the heavens down to Earth by showing that the planets are in the same situation as Earth. Perhaps most important from a philosophical and even a religious standpoint, Copernicus uprooted Earth from the center of the universe so we could no longer consider Earth to be a special in any sense. By "dethroning" Earth, Copernicus made it possible to imagine that the laws of nature we discover here apply throughout the universe. This is a major achievement.

K. BRAHE, KEPLER, NEWTON, AND EINSTEIN COMPLETE THE STORY

The history of science is a constant interplay between experiment and theory. As new and more accurate observations of nature are made, the successes or shortcomings of previous theories or models become apparent. Galileo's use of the telescope to see sunspots, Jupiter's moons, and the full phase of Venus, all of which validated Copernicus's model, is an excellent example of such interplay.

It was mentioned above that Copernicus's complex model matched Ptolemy's in accuracy, and if that were the only consideration there would be no reason to choose between the two. There are, of course, other considerations, such as Ptolemy's contrived slavery of Venus and Mercury and the reality of the full phase of Venus, which tilted the balance in favor of Copernicus. But more accurate observations carried out by the Danish astronomer **Tycho Brahe** (1546-1601) brought new problems. Brahe built sighting instruments, constructed without lenses, which allowed him to mark planetary positions to an accuracy of 1/30th of a degree. Remembering that the moon has an angular width of about half a degree, Brahe's accuracy is about 1/15th of the moon's diameter. This exceptional accuracy disclosed a serious discrepancy between Brahe's observations of the actual motions of the planets and the motions predicted by either Copernicus or Ptolemy . Clearly, Copernicus's model had to be modified. The man up to the task was the German astronomer **Johannes Kepler** (1571-1630). Kepler took Copernicus's sun-centered theory of the universe very seriously, but eventually realized its deficiencies. He worked with Tycho Brahe until the latter's death, and had to bargain for years with Brahe's family for his precious planetary data. Then the real struggle began. After years of tedious attempts with circles and circular motions of various sorts, Kepler slowly came to the realization that the actual motions are not circularly shaped at all but are elliptical (oval, or egg-shaped), with the sun occupying a special place inside the elliptical path called the focus as shown in Figure 13 with a greatly exaggerated ellipticity for the planet's orbit.

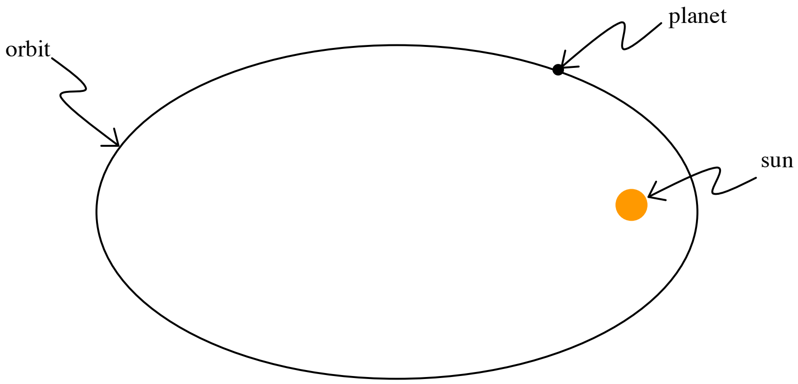


FIGURE 13: A planet orbiting the sun. (Ellipticity greatly exaggerated—in fact, all planetary orbits are much closer to circles.) The sun is at the focus of the ellipse.

Kepler's break with circular motions was an act of intellectual courage. When great thinkers use only circles for over a thousand years to explain planetary motions, it is quite an achievement to realize that they were wrong and that the motions are more simply explained by something rather different. Actually, a circle is a special case of an ellipse, but the key point is that Kepler had the courage to search and search until he came up with the right thing, even though it contradicted the centuries-old assumption about the special role of circles.

Kepler discovered three laws of planetary motion, the first of which is the elliptical shape of the orbits. We know today that Kepler's laws describe not just planetary motion, but the motion of any two bodies orbiting each other, such as a comet orbiting the sun, a satellite orbiting Earth, and two stars orbiting each other**. And all of this can be done without any epicycles at all!**

Everything we have discussed so far, from Ptolemy to Kepler, is purely descriptive in nature. Kepler's two remaining laws are quantitative, and marvelously account for planetary motions. But why do Kepler's laws work? There must be some deeper reason for this behavior. This question remained unanswered until Isaac Newton (1643-1727), at the tender age of 25 years, turned his attention to the problem. An appreciation of Newton's accomplishments returns us to the old objection against Earth's motion: What provides the huge force needed keep it moving? The ancients, except for a few like Aristarchus, couldn't imagine any such force and concluded that Earth must be at rest. But Galileo discovered what we might call "the law of coasting": the "natural" motion of a body that has no forces on it is to keep on coasting along in a straight line. The planets, of course, don't travel in straight lines, and Newton discovered that the force of gravity, provided by the sun which plays such a key role in

Kepler's laws, is responsible for curving their motions. To simplify, assume a planet is in a circular orbit with the sun at the center (which is very close to the truth for a planet like Earth). Gravity pulls towards the center and all by itself would cause a planet to plunge into the sun. On the other hand, if there were no gravity, the sideways motion of the planet, i.e., its motion perpendicular to the radius from sun to planet, would by Galileo's law of coasting cause the planet to coast away from the sun. The two influences together are practically in balance for Earth, so it goes about the sun in an almost circular orbit, the only force acting on it being the sun's gravity directed towards the sun.

Newton's "law" of gravitation, which describes in precise mathematical detail exactly how gravity behaves, can be used to prove that Kepler's three laws result from the existence of gravity. In fact, we should refer not to Newton's "law," but to Newton's "theory" of gravitation, because properly speaking, a law only describes behavior but explains nothing. Explanation is provided by theories, and no science worth the name can be taken seriously unless it strives to provide the deeper insights and illuminations provided by theories. Newton made a giant leap beyond the specifics of moon-Earth-planet motion and generalized his concept of gravity to all matter in the universe.

We said that planets travel in "almost perfect" ellipses, i.e., they "almost" obey Kepler's first law. The deviations are very small, and his first law can be used if the deviations can be ignored. The beautiful thing about Newton's theory of gravitation is that is clarifies the condition under which Kepler's laws are obeyed very precisely. The condition is that we have only two bodies orbiting each other, with one much larger than the other. In other words, Kepler's laws apply not just to planets, but to any two objects with one much larger than the other. For example, Kepler's laws apply pretty well to satellites orbiting Earth. When we have more than two bodies, Newton's theory shows that each will be affected by all of the others. Now the sun is much larger than any of the planets, but the planets do influence each other by their gravitational pulls. In particular, Jupiter, the largest planet, tugs on other planets with the result that the planetary orbits are not exactly elliptical. A careful application of Newton's theory to the interaction between planets accounts nicely for the precise shapes of the orbits.

Suppose we have only two bodies, but they are comparable in mass, e.g., a binary star system in which two stars orbit each other. In this case, Newton's theory shows that Kepler's laws again hold but with slight modifications. For example, each star travels in an elliptical orbit, but the focus of each ellipse is not at the other star but at a point called the "center of mass" situated between the two.

Newton's theory of gravitation was used for over two hundred years to understand and explain exquisite details in planetary motions, but one problem stubbornly escaped its grasp. Mercury, the planet closest to the sun, travels fastest in its orbit. The orbit is elliptical to a high degree of approximation, but the orbit rotates, or "precesses," over time. Over one century, it rotates 0.1594(≈ 0.16) degrees shown in Figure 14, which greatly exaggerates the extent to which Mercury's orbit is egg-shaped and the size of the angle of rotation. If Newton's theory of gravitation is used to calculate the gravitational effect of other planets on Mercury's motion, then a rotation of

0.1475 degrees is expected—about 0.01 degree short. Many people would be content to ignore such a small difference, but true scientists cannot be content with such an imperfection. But try as they might, nobody could account for the extra 0.01 degree.

The strange behavior of Mercury remained unexplained until 1915 when Albert Einstein created his General Theory of Relativity, which is a theory of gravitation based on the strange ideas of three-dimensional curved space and even curved time. Einstein's theory predicts significantly different results than Newton's only when masses are very, very large—substantially larger than our sun's—or when objects travel close to the speed of light. But at lower speeds, Einstein's more general theory transforms back into Newton's. In this situation, we say that Newton's theory is **incorporated** into the more general Einstein theory. As mentioned, Mercury travels fastest in its orbit, about 48 km/s (compared to 30 km/s for Earth). This is not very big compared to the speed of light (300,000 km/s), so there should be only a slight difference between Einstein's and Newton's descriptions of Mercury's motion. To Einstein's great achievement, his theory predicts that Mercury's orbit should indeed rotate by the extra 0.01 degree per century, as observed. Einstein's theory also predicts a similar but even smaller excess precession for other planets, and modern accurate measurements have again verified Einstein's theory.

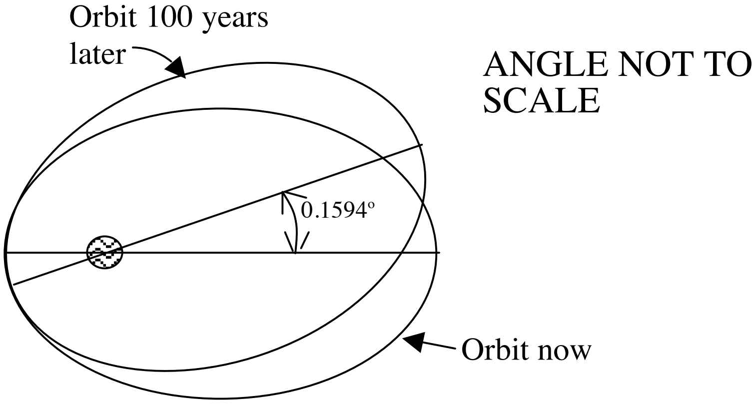
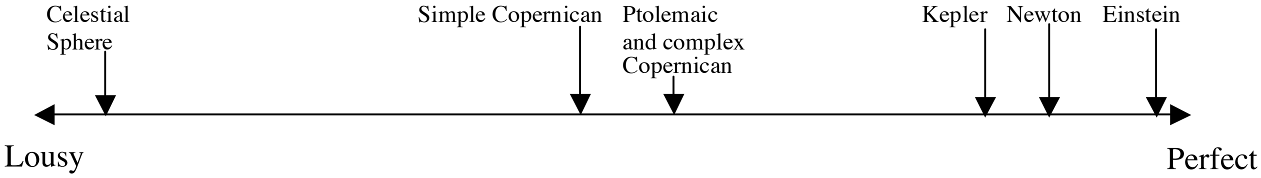


FIGURE 14: Precession of Mercury's orbit. Over a 100-year period, the orbit slowly rotates. The rotation amounts to about 0.16 degrees per century, but Newton's theory of gravitation comes up 0.01 degree short.

Einstein's new theory of gravitation gives a deeper insight into what gravity is than Newton's theory does. But as mentioned, Einstein and Newton's theories are practically identical for objects moving throughout our solar system at moderate speeds. Even for space probes traveling faster than Mercury (and with flight times much less than 100 years), it is fine to use Newton's theory. It would be foolish to throw out Newton's theory just because it isn't perfect, especially since Einstein's more perfect theory is more mathematically complex and difficult to use.

The following scale summarizes the accuracies of the various models and theories of planetary motions, from the ancient celestial sphere to Einstein's theory. As you might expect, the celestial sphere model is pretty lousy while Einstein's is, as far as we can tell, perfect. This scale shouldn't be taken as a mathematical measure; it is only qualitative.



As we know, Ptolemy's model is definitely unrealistic because it puts Earth at the center, while the simple Copernican model correctly has Earth going around the sun. Nonetheless, as the accuracy scale shows, Ptolemy's is more accurate than the simple Copernican model. So accuracy doesn't count for everythings; conformance with reality is even more important. Note also that the Ptolemaic and complex Copernican models are essentially equivalent in their accuracies.

L. FACT, LAW, AND THEORY IN SCIENCE

Science is sometimes misunderstood to be a collection of facts, e.g., particular observations such as "Mars was 30 degrees above the horizon in the constellation Orion on the night of September 4, 2000." Facts are very important in science, but if that's all science is, it would be very boring. In fact, science is much more. Specifically, it is the search for interesting patterns in observations and an explanation of why such patterns occur. The patterns in the observations are referred to as "laws," and the explanations of the patterns are "theories."

Our study of the Copernicus-Kepler-Newton-Einstein revolutions in our understanding of the solar system provides a good example of how facts can be generalized into laws and how theories are used to explain the laws. An example from the kinetic theory of matter, discussed in more detail in the next chapter, is also useful.

Here are four examples of "facts," followed by generalizations of the facts (i.e., "laws,") and theories which explain the laws.

M. Facts

1. The sun went across the sky yesterday, the day before, two days before, etc., from east to west.

2. Mars had the following positions and brightnesses in the sky at 2 am of the day indicated. (The units of position and brightness aren't important, but "dec" stands for declination; RA stands for right ascension; declination and right ascension are "celestial coordinates," much like latitude and longitude, and are used to locate an object in the sky.)

Date	Position		Brightness
September 1	dec 30 degrees	RA 18h20m	-1.84
September 2	dec 29.8 degrees	RA 18h05m	-1.82
September 3 …….....	dec 30.1 degrees	RA 18h10m	-1.83

Etc., etc.

3. Venus had the following phases at 3 am on the date indicated:

Date	Phase
October 23	10 percent crescent
October 24	14 percent crescent
November 15	100 percent (i.e., full)
Etc., etc…………………………………………………	

4. The pressure and volume of a gas are measured as a gas at constant temperature is slowly compressed, yielding the following values:

Pressure, P (N/m^2)	Volume, V (m^3)
0.3	0.51
0.375	0.41
0.5	0.29
0.2	0.75

N. Laws inferred from the above facts

1. The sun always rises in the east, and sets in the west.

2.1 Mars usually travels across fixed stars from west to east.
2.2 Mars reverses direction (“retrograde motion”) for about 70 days in the year.
2.3. Mars is brightest when in retrograde motion.

3. Venus goes through all phases, including full.

4. The volume is inversely proportional to the pressure. To put it another way, the product of pressure and volume is a constant at constant temperature.

Let’s take a closer look at how this last law is obtained. Figure 15 shows a plot of the data, and the following table reproduces the data above and shows the product too. By fooling around with the data, it can be seen that the product of P and V is close to 0.15. Is the product really constant? No, but all of the values are clustered around 0.15 or so. A sophisticated “curve fitting” process shows that the best value for the product is 0.1508, as shown in Figure 15.

Pressure, P (N/m^2)	Volume, V (m^3)	PV
0.3	0.51	0.153
0.375	0.41	0.154
0.5	0.29	0.145
0.2	0.75	0.150

Given that the product of P and V from the data table is not constant, why should we conclude that there exists a law which says that the product is constant? Anybody who has measured anything, from temperature to the length of a board, knows quite well that measurements are never perfect. In addition, the experiment which gave rise to the data was supposed to be carried out at constant temperature. It is impossible to make perfectly accurate measurements, and it is impossible to keep temperature perfectly constant, and this is why the product is not exactly constant. But the data strongly suggests that if more accurate measurements were made, and if the temperature could be controlled better, then the product of P and V would be even more constant than the above data shows. Indeed, the history of modern experimental science consists to a significant degree in better and better control and measurement of physical quantities such as temperature, pressure, and volume, and when this is done, the data conforms even more closely to the ideal law. This history of closer approximation to the ideal shows that we must look past the unwanted errors to see the underlying simplicity of nature. We construct our physical laws from the ideal (but impossible) measurements we imagine we could make if there were no errors.

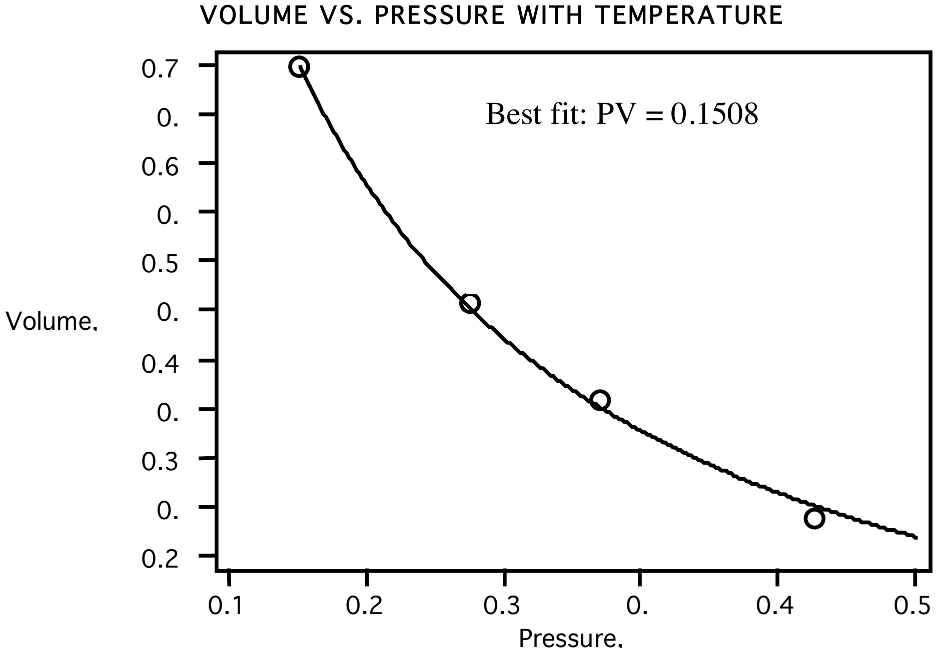


FIGURE 15: A plot of the volume of a gas as a function of pressure (with temperature held constant) shows that the product of the two quantities is a constant. Note how closely the data points are to the curve associated with the ideal equation PV = constant = 0.1508.

O. Theories which explain the above laws

The word "theory" comes from the Greek word *theoria*, meaning "mental viewing." To view something mentally is to grasp an understanding of it. In science, the essence of a theory is to explain, i.e., to help us understand, laws and the observations from which the laws come. A theory can explain what we can see by reaching **beyond** experience and imagining what the world is like at the very tiny atomic level, too small to be seen by a microscope, or at very large levels, beyond even our most powerful telescopes. A theory can also explain by going beyond our experience of the everyday world. Take friction, for example. Friction is always present; without it you couldn't walk across a floor, and with it you must burn gasoline to keep your car moving along a smooth level road. The ability to imagine what the world would be without friction was instrumental in Galileo and Newton's great theories about motions of ordinary bodies, as we will see in Chapter 4.

Laws 1, 2.1, 2.2, 2.3, and 3 are all explained to a high degree of approximation by Copernicus's sun-centered theory of the solar system. When coupled with Kepler's laws, Copernicus's sun-centered theory works marvelously well to explain planetary and solar motions and planetary brightnesses. This theory leads to an understanding that Law #1 is not very robust, because if Earth were hit by a huge hunk of matter, its direction of rotation could change so the sun would rise perhaps in the west and set in the east. In other words, the fact that the sun rises in the east depends on the accidental circumstance that Earth has been undisturbed by huge impacts throughout human history.

Law 4 is neatly explained by the *Kinetic Theory of Matter*, which states that matter is made of particles, ie., molecules and atoms, which are in constant motion. The pressure of a gas is caused by the impacts of the molecules on the sides of the container. When the gas is compressed, the density of the molecules is increased, so more of them hit each unit area of the container and this increases the pressure. A precise analysis of this process leads to the conclusion that the volume is indeed inversely proportional to the pressure.

The kinetic theory of matter explains heat (or more precisely, thermal energy) as the energy of motion of molecules. This is in stark contrast to the old caloric theory of heat, in which heat was considered a fluid which passed from hot bodies to cooler ones. The caloric theory is incorrect, as a number of experiments have shown.

P. Contrast of the everyday meaning of "theory" with the scientific meaning

In everyday language, the word "theory" is taken to mean a hypothesis or a speculation. This is clearly not what is meant in science, since some theories, e.g., Newton's theory of gravity (often referred to as a "law"), Einstein's theory of special relativity, and the atomic theory are so well-established that it would be foolish to have any doubt about them. In science, the essence of a theory is its ability to explain certain laws. Theories intellectually satisfy our human sense of curiosity about why things are the way they are. Facts, and even laws, are not the ultimate goal of science; theories and the understanding they provide are.

Q. The fate of theories

Scientists are constantly investigating the world, and this often involves taking more precise measurements under new conditions. Occasionally, the new measurements will discredit a theory in which case the theory has to be discarded. But sometimes it is possible to keep the old theory by modifying it or by recognizing its limitations. Here are some examples.

1. Ptolemy's theory is dead wrong in its assumption that the sun goes around the Earth. It was **discarded** and replaced by Copernicus's theory.

2. Copernicus's theory is absolutely right that the planets (including Earth) go around the sun. It is wrong in details, e.g., in its assumption that the orbits are circles or combinations of circles and epicycles, but the more important thing is that Earth goes around the sun, as Copernicus said. Furthermore, although the planetary orbits are not perfect circles, they are close to circular, which again lends credence to Copernicus's theory. His theory has been **modified**, but not in any real sense discarded. Today, we are all Copernicans even though we know his theory is not correct in all details.

3. Kepler's laws modified Copernicus's simple model by introducing elliptical orbits. With Kepler's modifications, the positions of the planets can be predicted with great accuracy without any need for epicycles. But most of the planetary orbits, though elliptical, are very nearly circular, and circles are mathematically a special form of ellipse. So Copernicus's simple model is **incorporated** into Kepler's more general and accurate theory.

4. Sir Isaac Newton developed the theory of gravity and used it to explain Kepler's laws of planetary motion. The theory of gravity can be used to explain deviations from Kepler's laws. In this way, Newton's theory of gravity shows the limitations of Kepler's laws, but also shows that they still work very well under a variety of circumstances. Kepler's laws are **incorporated** into Newton's theory because his laws of planetary motion are a special case of the variety of motions predicted by Newton's theory.

5. Albert Einstein developed a very deep theory of gravity which explained Newton's theory and which also showed that Newton's theory is not exactly correct. Newton's theory was not falsified, but was shown to be a special case of Einstein's more general theory, i.e., shown to hold only when planetary velocities are much smaller than the speed of light. Newton's theory, therefore, is **incorporated** into Einstein's theory.

6. The caloric theory of heat was **falsified** by a number of experiments which were totally incompatible with the theory. Caloric theory was **discarded** and replaced by the kinetic theory.

In summary, then, theories can be falsified and discarded, or they can be modified and incorporated into a better theory which shows the limitations of the original one.

R. The "quality spectrum" of theories

As we have seen, some theories may not be perfect, but that doesn't mean they are thrown out; they can be modified or their limitations recognized. Other theories have been discredited and are completely worthless, while others are, as far as we can tell, completely consistent with all observations. Generally, theories can be classified in order of their certitude as superb, useful, speculative, or worthless. Here are some examples of each, some of which will be met in later chapters:

Superb: Copernicus's sun-centered theory; quantum mechanics; Newton's laws; the kinetic theory of matter; Einstein's special theory of

Relativity; quantum mechanics; parts of the theory of black holes.

Useful: Copernicus's simple circular orbit model; the big bang; the standard model of elementary particles; Bohr's model of the atom; and parts of the theory of black holes.

Speculative: String theory of subatomic particles; parts of the theory of black holes; the inflation model of the big bang.

Worthless: Caloric theory of heat, Ptolemy's theory, Lamarck's theory.

Note that Copernicus's theory appears twice, depending on what aspect of it is of interest. The same is true of the theory of black holes and the big bang theory of the universe's origin.

S. Three conditions satisfied by a good theory

All theories except worthless ones must satisfy a certain number of conditions if they are to be seriously considered at all.

(1) Better understanding. First, a good theory must explain laws by allowing the laws to be logically deduced from the theory in such a way that it helps us to better-understand them. For example, Copernicus's simple theory shows that retrograde motion can only occur when Earth, which moves faster in its orbit than the superior planets, is overtaking a superior planet, say Mars. Since Earth can pass it by only when it is closest to it, it follows logically and simply that Mars must be brightest when it is in retrograde motion. Copernicus's simple theory also explains how Venus can have a full phase and allows the motions of the planets—which appear to be so complicated when viewed from a moving platform like Earth—to be simply understood to a good degree of approximation as combinations of a few simple circular motions. In all cases, there is a beautiful simplicity in Copernicus's model which is lacking in the much more complicated Ptolemaic model.

The great poet Samuel Coleridge defined beauty as "unity in variety," and that's exactly what we see in the ability of Copernicus's model to show how a great variety of observations stem from one simple idea. Indeed, many scientists throughout the years have used the criteria of simplicity and beauty to distinguish between competing theories. Regarding simplicity, note that a simple explanation gives a more immediate and satisfying sense of understanding than a complex one, which is why simple theories are always chosen over equivalent but more complex theories.

(2) Falsifiability. Second, a good theory must be "falsifiable," i.e., it must be possible for experimental observations to test the theory and possibly show that the theory is wrong. Again, let's look at Copernicus's theory. This theory said that the planets travel in circular orbits. But we know from Kepler's analysis of Tycho Brahe's data that the orbits are not circles, or even circles with epicycles; they are ellipses. So this detail of Copernicus's theory was falsified by the evidence. At the same time, the heart of Copernicus's theory--that the planets go around the sun--has been verified again and again so there is no doubt about it.

Could Copernicus's theory have been completely falsified to the point where it would have to be thrown out? Yes. If we never observed Venus to have a full phase, then Copernicus's theory would be completely falsified, because it would have incorrectly predicted that Venus <u>must</u> have a full phase.

Biological evolution provides another example of a theory which was falsified. Prior to the modern version of Darwinian evolution, a Frenchman by the name of <u>Lamarck</u> theorized that acquired characteristics could be inherited. For example, if a man and his wife both exercise their right arms so that these arms are extremely muscular, Lamarck predicted that the offspring could inherit very large, muscular right arms. A wealth of evidence has shown that this is impossible.

In the mid-1800s, the famous scientist Lord Kelvin estimated the minimum age of Earth to be about 100 million years on the assumption that Earth has been continuously cooling down by heat conduction from a temperature initially just above the melting point of rocks. His theory was decisively refuted when Kelvin's associate, Perry, pointed out that a molten Earth would also cool by convection, i.e., by the internal flow of material from the deep interior to regions closer to the surface. Given the rate at which Earth's temperature increases with depth (a measurement known to Kelvin), this mixing of Earth's interior slows the cooling of the interior by keeping the surface warm, and this implied an age for Earth much greater than Kelvin's 100 million years. Further evidence against Kelvin's theory surfaced when the phenomenon of radioactivity was discovered in the late 1800s, though this is a minor effect compared to Perry's theory. Radioactive materials throughout Earth's mass contribute heat, and further extends the time needed for Earth to cool. Modern estimates of Earth's age on the basis of Perry's rate of cooling and radioactive decay indicate that it is about 4 billion years old. (An excellent history of this intellectual adventure can be found in the July/August 2007 issue of *American Scientist.*)

The steady state theory of the origin of the universe provides us with another example of a scientific theory falsified by evidence. According to this theory, the universe is infinitely large and has existed for an *infinite* amount of time. Yet, we observe galaxies (which are huge rather stable clumps of stars) to be moving away from each other. If this expansion process were really going on for an infinite time, there would be an infinite distance between galaxies, i.e., we couldn't see any galaxies outside of our own. In fact we can see many nearby and distant galaxies. The steady state theory accounts for the observed density of galaxies by assuming that matter is continuously created as the universe expands. This new matter forms new galaxies at just such a rate as to keep the density of galaxies constant. One consequence of this theory is that any large group of galaxies anywhere in the universe at any instant of time--today, or billions of years ago--should have the same distribution of ages, i.e., a fixed fraction of very young ones, a fixed fraction of middle age ones, etc. The main competing theory, the big bang, assumes that the universe is not infinitely old but came into existence about 15 billion years ago. Because it takes light from distant galaxies a long time to get to us, looking at more and more distant galaxies is equivalent to going back in farther and farther in time, which means that very distant galaxies should appear as they were shortly after the big bang, i.e., they will be “young” in the sense that they have a different appearance than older galaxies closer to us. In recent years, astronomers have been able to observe in some detail extremely distant galaxies, and they do indeed appear to be younger than the ones closer to us. These observations, therefore, support the big bang theory and falsify the steady state theory.

The big bang provides simple and reasonable explanations of a number of other properties of the universe, and for this reason it is the favored explanation for the nature and origin of the universe. Does this mean that the steady state theory should be tossed into the garbage dump of history? Not quite. Cosmology, the study of the origin and properties of the entire universe, is a very young science, and the big bang itself has difficulties accounting for certain phenomena, particularly the clumping of large numbers of galaxies over large volumes of space. It is rather unlikely, but perhaps some day new data will decisively invalidate the big bang and a clever modification of the steady state theory will offer a better explanation of the new data.

The Ptolemaic and Copernican theories of the solar system, the Lamarckian theory, the steady state and big bang theories of the universe, and the various theories for the ages of Earth are all scientific in the sense that they are falsifiable, i.e., they can be put to an experimental test to determine if they are right or wrong. In this regard, the hallmark of science is to assume that natural phenomena have natural explanations, because only natural explanations can be experimentally tested:

“SCIENTIFIC MEANS TESTABLE AND HENCE FALSIFIABLE;
NATURAL PHENOMENA HAVE NATURAL EXPLANATIONS.

In this sense, we can consider a person who forms a scientific theory to be courageous because their ideas can be put to an experimental test which will either prove or disprove the theory. In the worst case, the theory winds up being so wrong it must be thrown out. But more often, the theory is modified or recognized to have a limitation which doesn't completely wipe out its usefulness.

The "falsifiability" aspect of science can be appreciated by contrasting it with non-scientific ideas. Consider, e.g., the statement that "There are unknowable, invisible alien creatures among us." This rather strange claim may be true, but how can we every know if it is true or false? Since these creatures are unknowable, we can never do anything to prove they aren't "among us." Therefore, this statement is unscientific, even though it is possibly true or false! As another example, consider the creationist claim that God created the world 6,000 years ago, but included in the creation fossils and other kinds of things which make the world look much older. Since the creationist God is not of this world, and since it there is no conceivable evidence that would show that the fossil evidence is really millions of years old but made by God 6000 years ago, there is no way to disprove the creationist claim. For this reason, the claim is unscientific even though it is possibly true or false.

(3) Broad consistency ("consilience"). The third condition a good theory must satisfy is that it must be consistent with a <u>broad</u> range of evidence, laws, and theories.

Again, Copernicus's theory serves as a good example of this condition, because it is consistent with:

(a) the rotation of Earth, as disclosed by the famous Foucault pendulum experiment (in which the direction of swing of a pendulum changes as Earth rotates underneath it);
(b) Newton's theory of gravitation, according to which Earth must be in motion around the sun to avoid being drawn into the sun;
(c) the "dipole" nature of the cosmic background radiation left over from the big bang which gave birth to the universe ("dipole" means that Earth's motion produces annual variations in the big bang radiation); and
(d) annual parallax of stars.

The current buzzword for this broad consistency is **consilience**. Another example of consilience is provided by the atomic theory, which requires that equal volumes of different types of gases (e.g., oxygen and hydrogen) at standard temperature and pressure must contain the same number of molecules. This leads to the concept we call Avogadro's number (named after the 19th century Italian physicist), which is the number of molecules in a particular volume of gas at standard temperature and pressure. It is common to take the volume to be 22.4 liters, in which case Avogadro's number is astonishingly large, 6.025×10^{23} which reminds us how small gas molecules and atoms are! Avogadro's number appears throughout physics and chemistry, and there are approximately fifty different independent experiments involving various properties of matter, from the behavior of gases to electrical properties of matter to the blueness of the sky, which allow Avogadro's number to be determined. All of these various methods give the same value within the experimental measurement errors, and thereby provide overwhelming evidence that the material world around us is truly atomic in nature and that Avogadro's number reflects an important property of the real world. Atoms are not an imaginary artifact, as many chemists and physicists thought in the 19th century, but are as real as you and I are even if we can't pin them down to being either waves or particles. This sense of robust reality characterizes other superb theories of physics, such as the theory of EM waves, Newton's laws, and Einstein's special theory of relativity.

Another example of a theory, or rather an application of a theory, which is consistent with a wealth of data from separate fields of science, takes us back to the age of Earth. As mentioned above, the rate of cooling of Earth, properly taking account of loss of heat by convection, indicates that it is about 4 billion years old. Independent measurements of radioactive decay products in stable rocks indicate that Earth's age is approximately 4.5-billion years old, i.e., it is in the same ballpark as the age determined from rate of cooling. That the two methods are as close as they are is truly amazing, because estimation of something as far removed from everyday experience as the age of Earth is a tremendously difficult thing to do. If the estimate from cooling were far removed from the age of rocks due to radioactive decay, then one or the other method of estimation would be suspect and we could not say, even roughly, what the age of the Earth is.

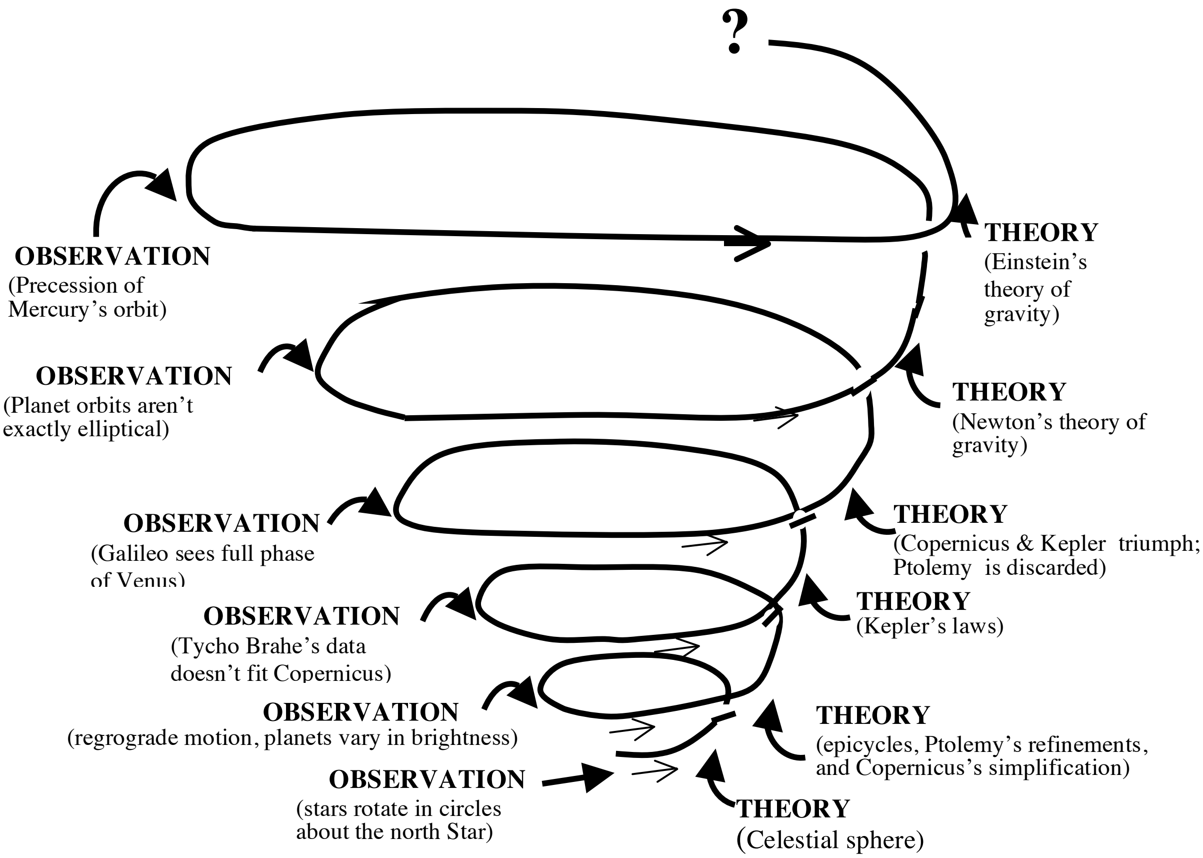
A number of other age estimates point in the same direction. For example, the estimate of the likely age of the sun, based on rather well-established theories of nuclear reactions, point to a figure of about 5-billion years old, again in the same general ballpark as Earth's age by the other two techniques. It would appear, then, that Earth formed about half a billion years after the sun's nuclear reactions ignited. Finally, we should expect the age of the universe to be greater than the age of the oldest stars we can observe. Again, both figures are in the tens of billions of years, with the universe being 12 to 15 billion years old and the oldest stars being about the same. Actually, observations of stars in "globular clusters" indicated that their oldest stars are older by a billion or so years than the universe they inhabit! This, of course, is as impossible as a child being older than a parent. Both estimates are somewhat uncertain, so the discrepancy wasn't serious enough to warrant throwing out one or the other one. But good science requires that more and more precise estimates be made of both things, and astronomers have studied the latest observations from the Hubble space telescope and the Hipparchos satellite. Indications are that the globular clusters are actually farther away than thought, and the age estimates depend crucially upon how bright the stars are. Since the clusters are farther away than previously thought, their stars are brighter and hence younger, so the age contradiction appears to be resolved.

A recognition of the broad consistency and interrelationship of theories and laws can be achieved only by a disciplined study of many branches of science. For example, the age of Earth provides us with an example involving theories of heat transfer, radioactive decay, and stellar evolution, all of which indicate that Earth is between 4 and 5 billion years old. Broad consistency strengthens our confidence that the scientific way of thinking can indeed provide us with reliable knowledge of the world around us.

T. The interaction between theory and observation

We cannot have science without experimental observation, because only observation can test and possibly falsify an otherwise logically consistent theory. But the interaction between theory and experiment is a two-way street. Sometimes a new theory predicts the outcome of an experiment, and when the experiment is done, the theory's prediction is either verified or falsified. Other times, an experiment appearing at first to have little or nothing to do with a particular theory will give a result which is found to be inconsistent with that theory, causing it to be modified or discarded. We have seen a number of examples of the interaction between theory and observation. For example, Einstein's theory of gravity suggested that light bends in a gravitational field, and when this experiment was done, Einstein's theory was verified. On the other hand, Tycho Brahe's very accurate observations of planetary observations, intended to verify his own Earth-centered model of the solar system, conflicted with both the Copernican and Ptolemaic models and presented a challenge to Kepler who successfully met it.

Figure 16 shows schematically how theory and experimental observations of planetary motion interact in such a way as to kick science (and theory in particular) along an upward spiral to higher, more accurate, and more intellectually satisfying levels. In any natural science, it is the experimental observations which "juice up" the system in unexpected ways so it can go to higher levels. Nature is full of surprises, and without this "juice," science wouldn't be the fascinating and challenging enterprise we know today. At the same time, theory suggests new experiments to be done, so there is a fruitful interaction between theory and experiment. In this way, our scientific understanding of the world is continuously refined.



<u>FIGURE 16</u>: The rising spiral of scientific progress showing how new observations spur scientists to formulate more accurate theories. Planetary motions are used to illustrate the interplay between observation and theory, which has culminated in Einstein's theory of gravity. The observations and theories kick the spiral ever higher; the spiral itself represents the strenous mental and physical efforts of experimentalists and theorists to figure out nature. What follows Einstein's theory remains to be seen.