

Curvature and the shape of the universe

Chikako Mese *

Department of Mathematics
Connecticut College
email: cmes@conncoll.edu

May 5, 2004

Have you ever looked up into the skies at night, gazed at the myriad of stars that speckle the darkness and wondered what is out there in the vastness which we call the universe? What will we find if we can take a space ship out into space, past our solar system, past our galaxy and explore immense regions of space? What is the shape and size of our universe? Will we reach an edge or a boundary where the universe stops? If you have ever asked these questions, you are not alone. For thousands of years, many human beings have looked up at the sky and done exactly the same. Due to our scientific and mathematical advancements, we are much closer today to discovering the reality of our universe. A crucial step in this quest is an understanding of the concept of curvature.

In this paper, we will not assume much background in mathematics beyond some basic high school geometry and a little bit of trigonometry. We hope to explain the notion of curvature, which is a very sophisticated mathematical concept, developed slowly by the finest mathematical minds of history. Our goal is to give enough of an appreciation of curvature so that you can get a better understanding of the physical nature of the universe.

Before we get into a discussion on the shape and size of the universe, let's start with a question a little more manageable. What is the shape and size of the Earth? Today, most of us have been taught that the Earth resembles a sphere. Additionally, we can look in the encyclopedia and find that the Earth has a circumference of approximately 24,900 miles. Our technological advancement allows us to know the answer to this question. But for thousands of years during the existence of the human race, the physical nature of the Earth was just as mysterious as our query about the cosmos.

During the course of the history, many people have believed that the Earth was flat and that we would fall off its face by wandering too far away. This is not a surprising conclusion. Take a casual look at the landscape around you, and the Earth does seem flat. But there is evidence all around that the Earth is spherical. For example, consider the lunar eclipse where the shadow of the Earth can be seen on the moon's surface or in the increase of the angle of altitude of the North Star as one travels north. Through these kinds of observations, some ancient civilizations came to the understanding that the Earth was spherical. Therefore, knowing that the Earth is a finite object, they contemplated its size. Without sophisticated instrumentations at their disposal, this was an extremely difficult question, one needing remarkable cleverness to obtain an answer.

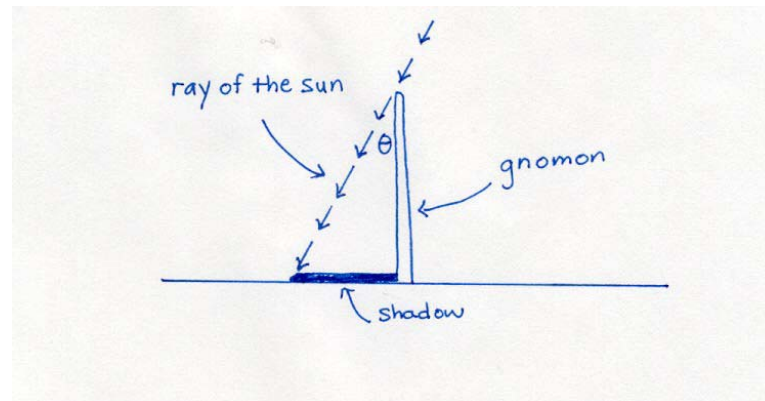
Eratosthenes, who lived in Alexandria around 3 B.C., proposed one of the most ingenious solutions to this problem. Amazingly, his main instrument was essentially a stick that is placed perpendicularly to a horizontal ground. This instrument, called the gnomon, had several purposes. One, it was sort of a primitive sundial; at noon when the sun was at its highest point in its daily trek across the sky, the shadow cast on the gnomon was the shortest. Furthermore, it also acted as a primitive calendar. If we measure the length of its shadow everyday at noon, the day that the shadow is the shortest is the day of the summer solstice (about June 21 in the northern hemisphere which is the day we receive the longest duration of sunlight) and the day the shadow is the longest is the day of the winter solstice (about December 21).

How did Eratosthenes use such a simple instrument to compute the size of the Earth? About 500 miles due north of Alexandria was an ancient city of Aswan (now called Syene). Aswan lies on the Tropic of Cancer, which means that at noon on the summer solstice, the sun lies directly above the city. Thus, at midday the gnomon in Aswan casts no shadow. Back in Alexandria, Eratosthenes measured the shadow cast

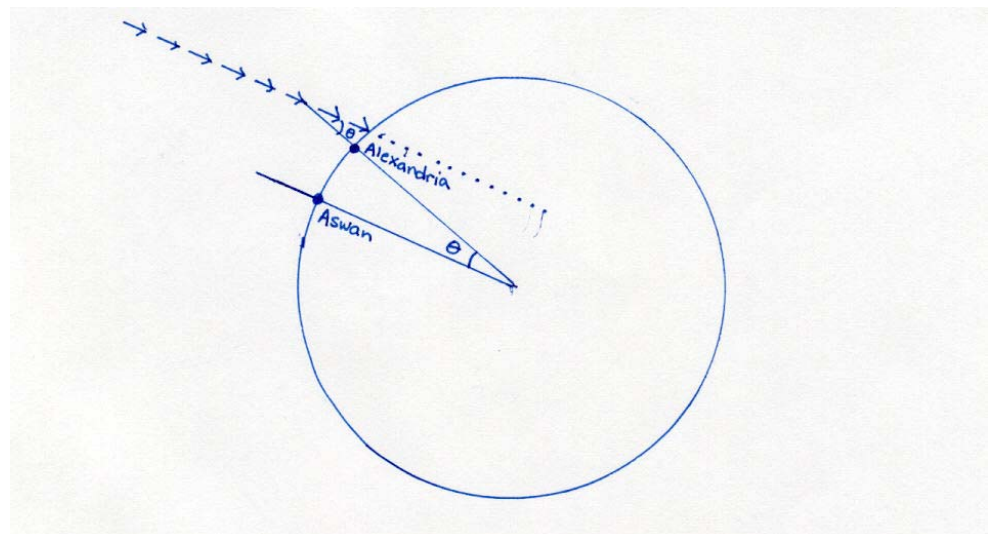
*Partially supported by grant NSF DMS-0072483. The author would like to thank the faculty of the University of Dayton Mathematics Department for their invitation to speak at their Undergraduate Mathematics Conference. Additionally, she would like use this opportunity to express her gratitude for the impact they have had on her career in mathematics.

by the gnomon there and found that the ratio of the shadow over the length of the gnomon was

$$\frac{\text{length of the shadow of the gnomon}}{\text{length of the gnomon}} \approx 0.13.$$



If θ is the angle between the gnomon and the ray of the sun, then trigonometry says $\tan \theta = 0.13$. This allowed him to compute $\theta = \tan^{-1}(0.13) \approx \frac{2\pi}{50}$. This angle is exactly the angle between the two lines drawn from Alexandria to the center of the Earth and from Aswan to the center of the Earth.



Using the fact that the distance between Aswan and Alexandria is 500 miles, he set up the following ratio,

$$\frac{C}{500} = \frac{2\pi}{2\pi/50},$$

where C is the circumference of the Earth. Solving for C , we get

$$C = 500 \times 50 = 25,000.$$

Eratosthenes' estimate is only 0.4% off today's estimate of 24,900 miles, an astounding feat considering the tools he had at his disposal!

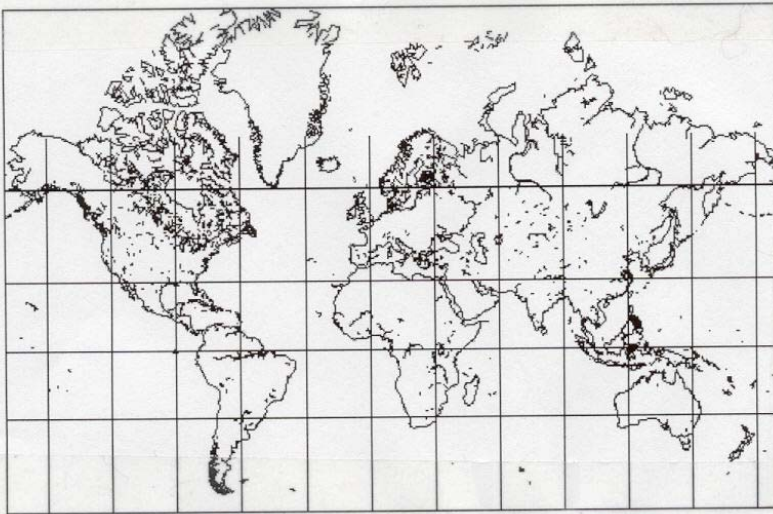
We now pose a harder question:

QUESTION: Can we measure the size of the earth without relying on an extraterrestrial aid such as the sun and the gnomon sticking out of the surface of the earth?

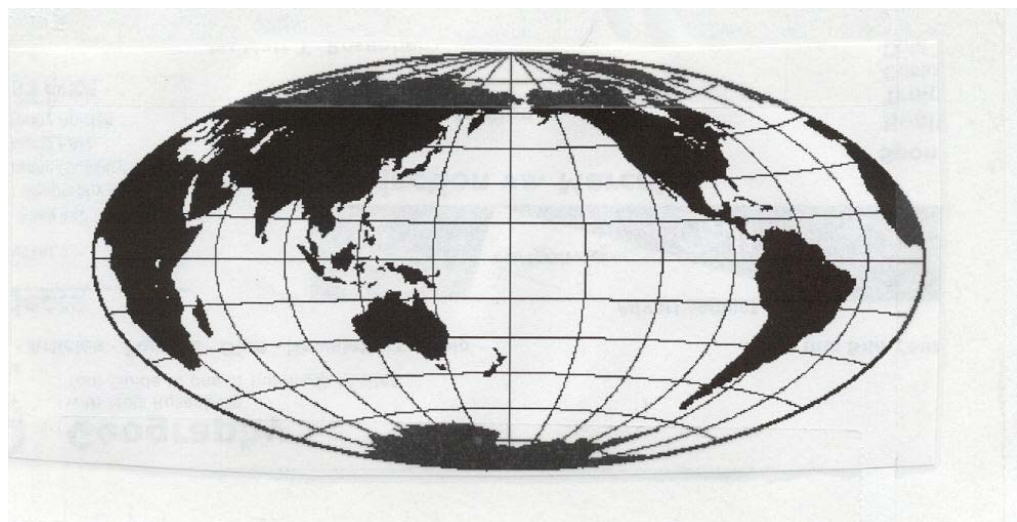
A rephrasing of this is questions is this: if we were a two-dimensional being living on a sphere, without any awareness of a three-dimensional space that contains the sphere, could we compute the size of the sphere without actually measuring it? The motivation for this question is this: we are a three-dimensional being living in a three-dimensional world. In order to understand the physical properties of our universe, it will be important to be able to answer the question about the size of the two-dimensional surface from a perspective of a two-dimensional being.

The key to answering this question is closely related to understanding cartography, the art of map making. The information that a cartographer wants to convey is the distance and direction from one location to another. In other words, a map should scale distances between two different places properly and preserve angles.

The importance of an accurate map became crucial during the Age of Exploration when ships sailed out of their ports to explore vast regions of the Earth. Accordingly, cartographers spent much energy and effort trying to give an accurate depiction of a large portion of the Earth on a piece of paper. Let's look at some examples. Here is a sketch of a very famous map called the Mercator projection.



Notice that on the Mercator projection, we can detect great distortions of distances near the North and South Poles. Observe also the distortion in the next map, called the Hammer projection.



There are many other types of maps invented by cartographers. But as hard as they tried, no one was able to come up with a map that precisely preserved information about distances. There was a reason for that. A mathematician by the name of Leonhard Euler proved the following theorem.

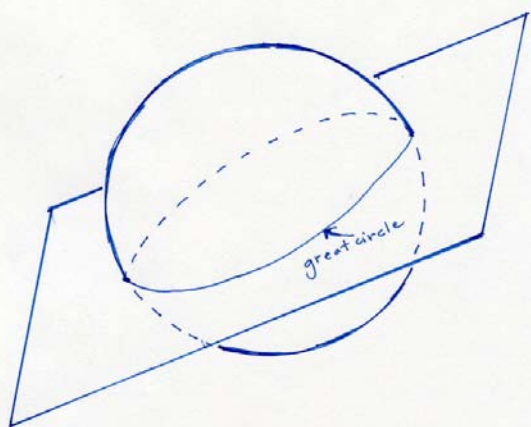
Theorem (Euler) *No map of any portion of the Earth scales properly.*

The key to understanding Euler's Theorem is to realize that the geometry of the plane and the sphere are quite different.

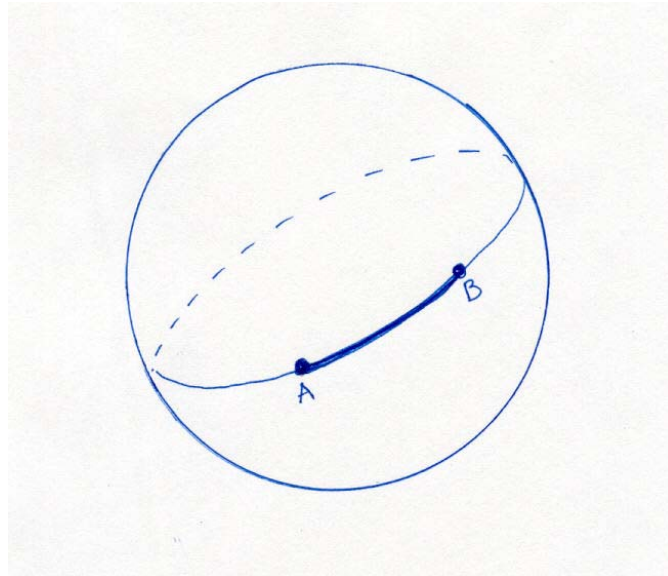
Recall what you learned in your high school geometry class. You probably learned that a circle of radius r has a circumference of $2\pi r$ and that the sum of interior angles of a triangle is equal to π radians or 180 degrees. You also probably memorized the Pythagorean theorem, which says that the sum of the squares of the lengths of the two legs is equal to the square of the length of the hypotenuse. Implicit in these statements is that this is the geometry of the plane or *planar geometry*.

Since the Earth resembles a sphere, let us study the geometry of the sphere or *spherical geometry*. What can we say about geometric objects that exist on the sphere?

First, we need some preliminary specifics concerning a sphere. A *great circle* or *the equatorial circle* is the curve of intersection of the sphere and a plane that cuts the sphere in half (that is, the plane that contains the center of the sphere).

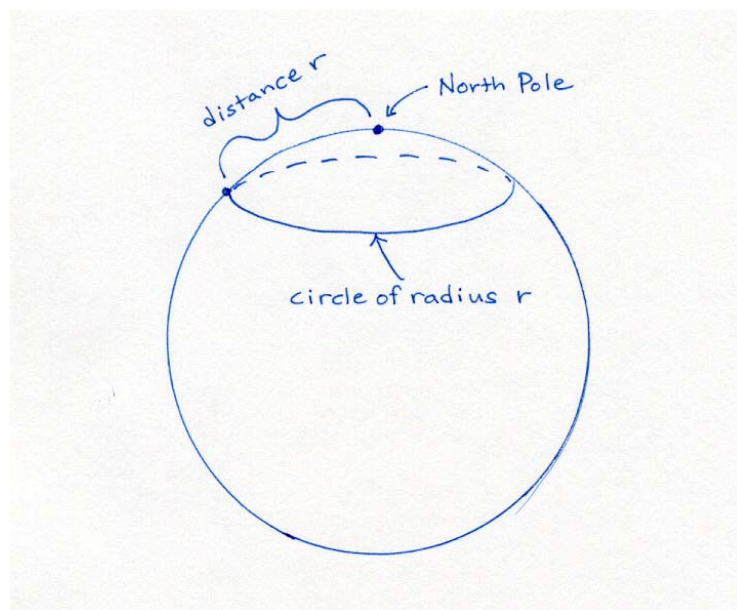


For example, the equator is a great circle. Additionally, note that longitudes are great circles, but the only latitude that is a great circle is the equator. Great circles are important when we consider spherical geometry because if you want to go from point A to point B on the sphere, the best way to get there is to follow along the curve of the great circle.

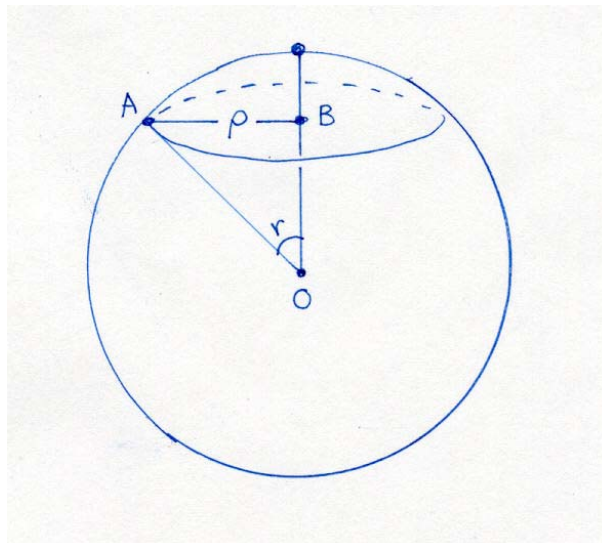


In other words, segments of great circles on the sphere (whose lengths are less than or equal to half of the circumference) are analogous to line segments on the plane in the sense that they are shortest paths between two points.

We will start by studying circles (not the great ones): a set of points equidistant from a given point) on a unit sphere (or sphere of radius 1). A circle of radius r centered at the North Pole looks like this:



In order to figure out the circumference of this circle, we need the value of ρ (the radius) in the diagram below.



Because the sphere has radius 1, the angle $\angle AOB$, measured in radians, is equal to r . Thus using some trigonometry on the right triangle $\triangle AOB$, we see that

$$\sin r = \frac{\rho}{1}$$

or that

$$\rho = \sin r.$$

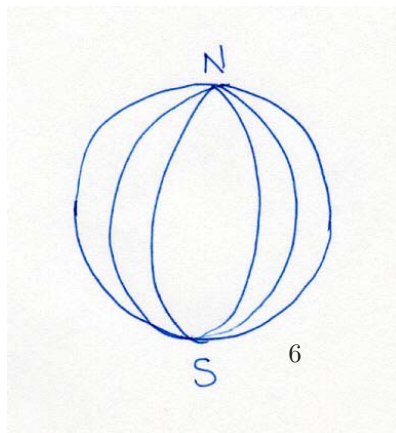
This means that:

The circumference of the circle of radius r on the sphere is $2\pi \sin r$.

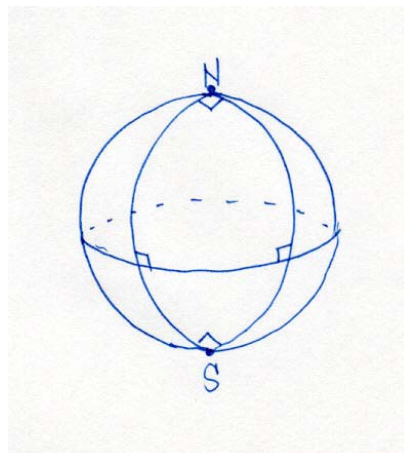
Now recall that the circumference of a circle of radius r on the plane is $2\pi r$ and since $\sin r < r$ for $r > 0$, we have

$$2\pi \sin r < 2\pi r.$$

In other words, the circumference of the circle of radius r is smaller on a sphere than on a plane. Visually, we can see that this makes sense since the radial lines from the North Pole (in other words, curves that are a part of the great circle that radiate from the North Pole) on the sphere converge closer and closer together until they meet at the South Pole.

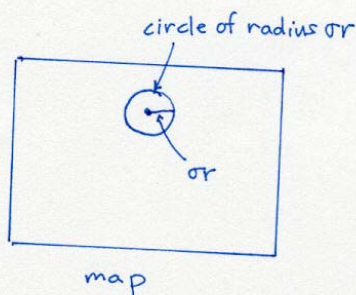
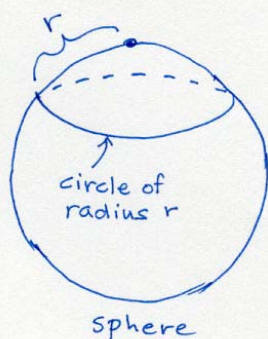


We also see that the two other geometric facts from planar geometry mentioned above are not the same in spherical geometry. In order to investigate this, consider great circles passing through the north and the south poles. We choose two such great circles that intersect at the poles in 90-degree angle.



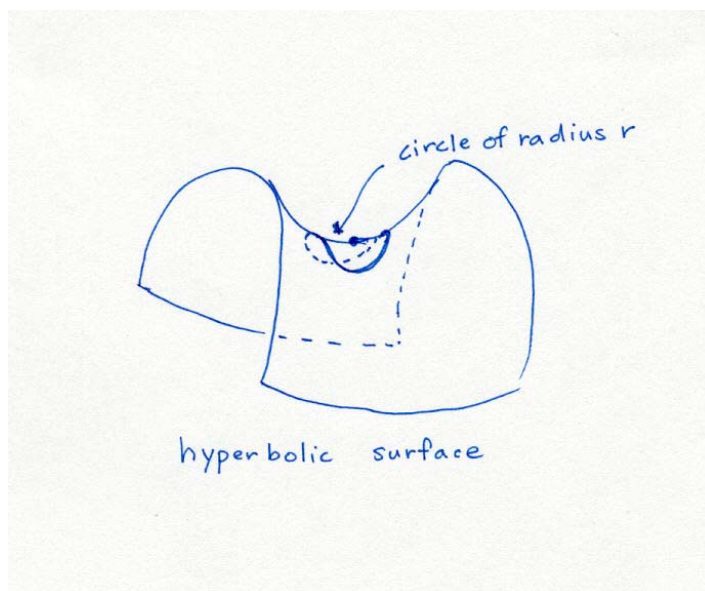
Consider a segment of these two great circles from the North Pole to the equator. These two segments, plus a portion of the equator, define a triangle on the sphere. (This triangle covers a fourth of the northern hemisphere.) Now notice that this is an equilateral triangle and the interior angles are all 90 degrees. In particular, this means that the interior angles do not sum up to 180 degrees and that the Pythagorean theorem does not hold.

With this, we can now give a proof of Euler's theorem. Suppose that there is a map of the Earth that scales distances. That is, any two points of distance d apart on the Earth is represented on a map by two points of distance σd on the map. (Here σ is the scaling factor.) Then a circle of radius r on the Earth is represented by a circle of radius σr on the map. Let C_{Earth} be the circumference of the circle on the Earth and C_{Map} be the circumference of the circle on the map.



Furthermore, because this map scales distances by a factor of σ , C_{Map} would equal σ times the circumference of C_{Earth} ; in other words, $C_{\text{Map}} = 2\pi\sigma \sin r$. This is a contradiction since a circle of radius σr on a plane has a circumference of $2\pi\sigma r$.

The key feature of this proof is that it exploits the fact that spherical and planar geometry are qualitatively very different. There is another type of geometry that is different from spherical and planar geometry and that is the hyperbolic geometry. Hyperbolic geometry can be thought of as the geometry that is the opposite of spherical geometry. In this geometry, radial lines from a point diverge (as opposed to converge in spherical geometry) and thus circles of radius r have a circumference of more than $2\pi r$ and the sum of the interior angles of a triangle is smaller than 180 degrees.



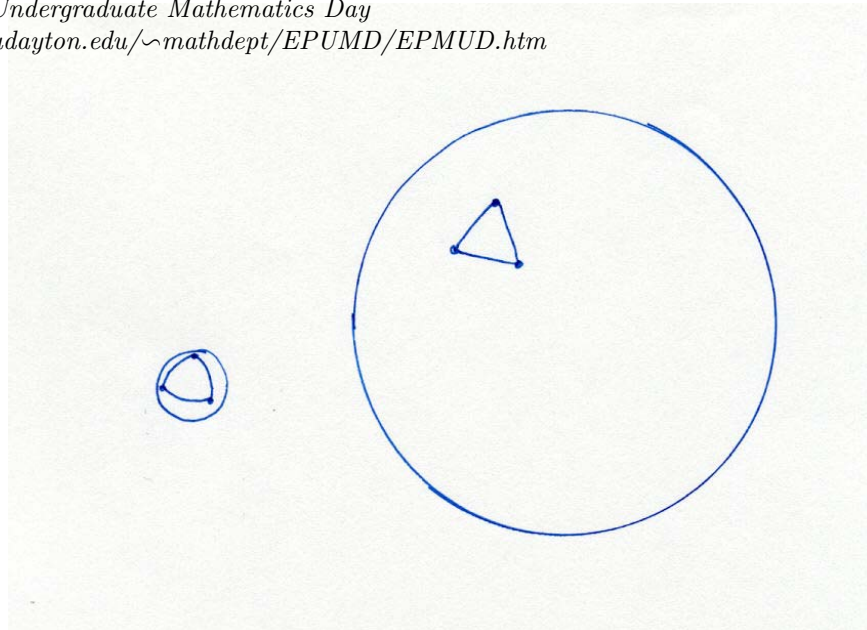
We use these qualitative differences to define curvature of a surface as positive, zero, or negative. A surface that has the same geometry as a plane is said to have zero curvature and a surface in which all circles of radius r have circumference smaller than (or bigger than) $2\pi r$ is said to have positive (or negative) curvature. We note here that the choice of positive for spherical geometry and negative for hyperbolic geometry is by convention. In other words, we could have easily defined spheres to have negative curvature and hyperbolic planes (two dimensional spaces with hyperbolic geometry) to have positive curvature. Furthermore, we can also define positive and negative curvature in terms of the sum of interior angles of a triangle.

Here's question to test your understanding of the concepts above.

Can we construct a map of a negatively curved surface without distorting distances?

The answer is no. Once again, we can prove this by exploiting the fact that the geometry on a negatively curved surface is fundamentally different from the geometry of a plane.

Getting back again to spheres, we notice that the geometry in spheres of different radii is not the same. That is, consider one sphere of small radius, R_{small} , and another sphere of large radius, say R_{large} , and study a triangle of fixed side lengths on each of the respective spheres. We know that sum of the interior angles of each of the triangles is greater than 180 degrees. But moreover, we note this difference is much greater on the small sphere than on the large sphere.

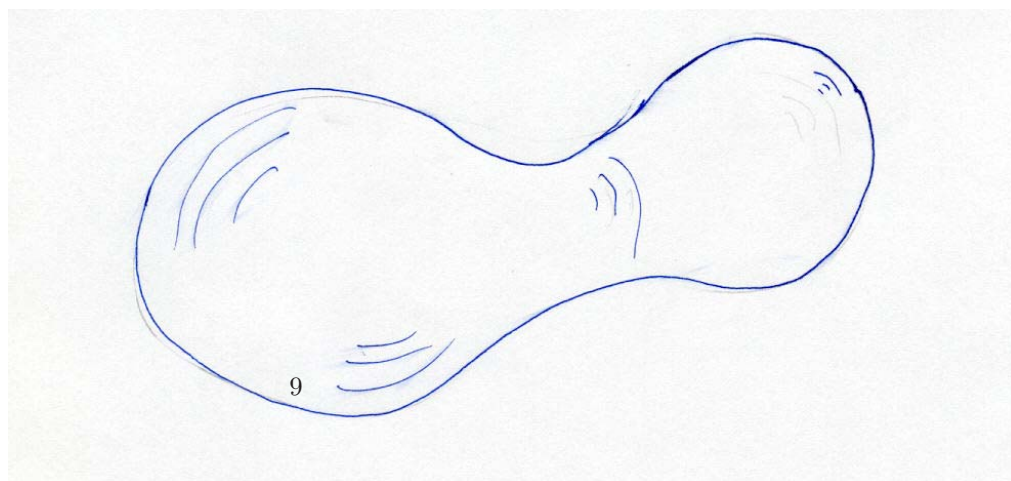


In other words, there is a quantitative difference between the geometries, even amongst different spheres.

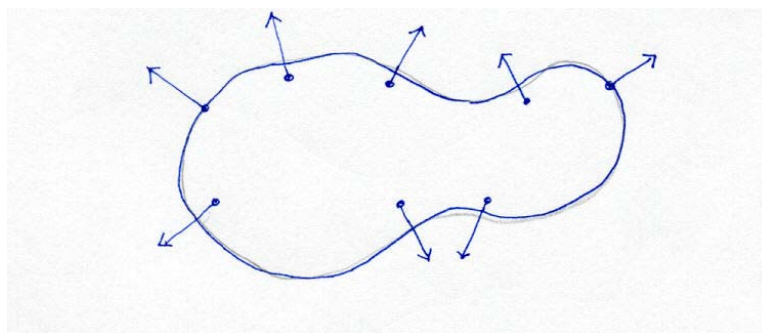
At this point, let's get back to our question about the size of the Earth. Consider an equilateral triangle of fixed side length on the Earth. Let's say, for the sake of argument, that the side length is 500 miles. Now measure the interior angles of this triangle. Now, take a sphere that is reasonably sized, say with circumference $1/100$ miles, and find an equilateral triangle with the same size angles. Because this sphere is much smaller than the Earth, the triangle on this sphere is much smaller than the corresponding one on the Earth. For the sake of argument, let's say we find this triangle to have side lengths of $\frac{1}{50000}$ miles. Now notice that all spheres look alike; in other words, they can be scaled up and down in size. Thus, the fact that the equilateral triangle on the Earth and the equilateral triangle on the reasonably sized sphere have the same size angles means the ratio of size of the Earth and this reasonably sized sphere is $\frac{500}{\frac{1}{50000}} = 2,500,000$. Thus the size of the Earth is this ratio times the size of the reasonably sized sphere, or $2,500,000 \times \frac{1}{100}$ miles = 25,000 miles. In this calculation, we are exploiting the quantitative difference between a big sphere and a small sphere.

Mathematician Carl Fiedrich Gauss (1777-1855) studied these quantitative differences of geometries (and not just the qualitative differences that define positive, zero and negative curvature) and was able to give a numerical value for curvature at a point in a surface in a mathematically precise way. In understanding this, it is helpful to consider yourself riding your bicycle in a straight line, around in a big circle, and around in a small circle. Experience will tell you that its easiest to ride in a straight line, harder to ride in a big circle, but hardest to ride in a small circle. So it make sense to assign a larger numerical value of curvature on the small circle than on the big circle. Analogously, it should make sense to assign a larger numerical value of curvature on the small sphere than on the large sphere.

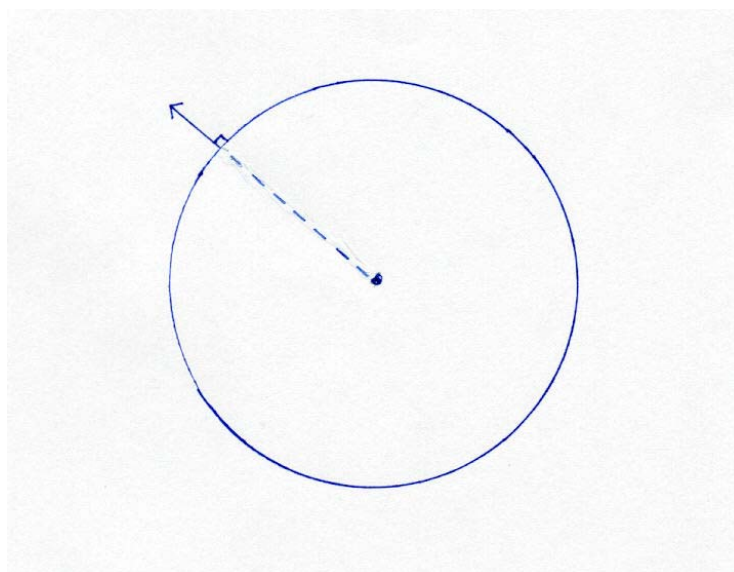
Now note that a sphere is a special type of a surface where the geometry near a given point of a surface is exactly the same as the geometry near another point of the surface. This is an example of what is known as a *homogeneous space*. In other words, two circles of the same radius centered at different points on a homogeneous space look exactly alike. Now notice that a general surface, such as the one pictured below, does not have this property.



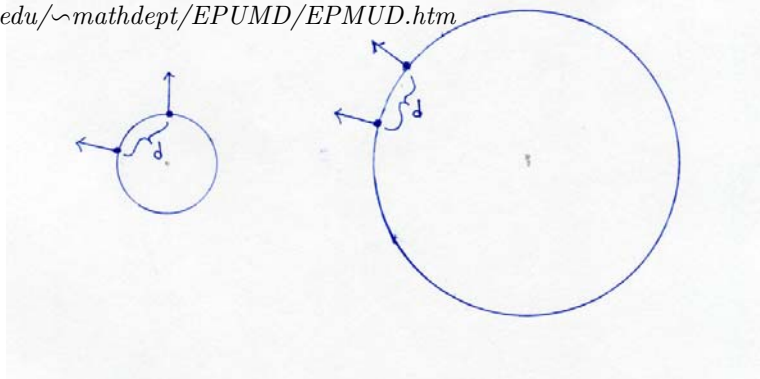
How does one precisely assign a numerical value of curvature to a point on a surface? One method is to see how the surface curves into the three-dimensional space. To do this, consider arrows placed perpendicularly at different locations along a surface with arrows oriented so that they point outward from the surface. These arrows are called *normal vectors* to the surface. Note that these normal vectors exist in the three-dimensional space that contains the surface.



For example, on the sphere, the normal vectors lie along the ray radiating out from the center of the sphere.

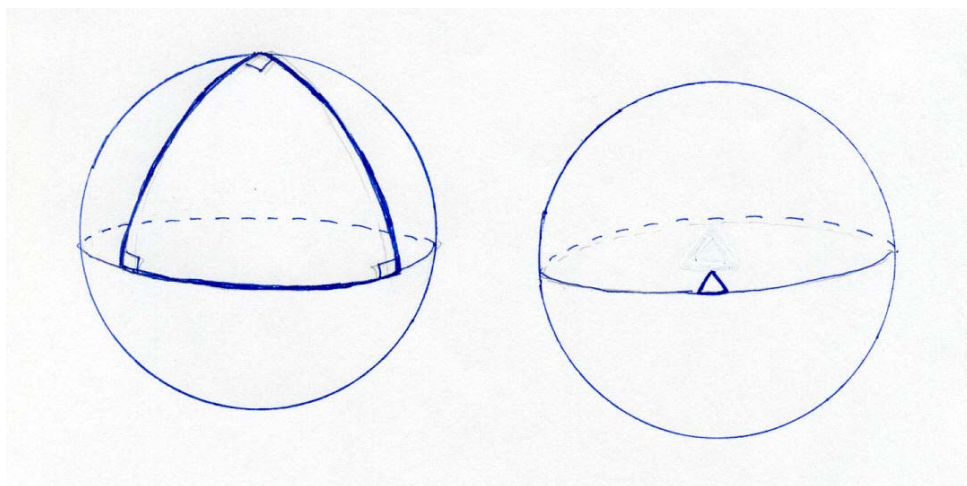


Now an important observation is that if we take a pair of normal vectors of some fixed distance apart, say d , on a small sphere and also on a large sphere, the difference in how the two vectors of the pair lie in space is much greater on the smaller sphere than on the larger sphere.



This is also the reason it makes sense to define a larger numerical value of curvature to a small sphere than to a large sphere. It is also the same idea that was utilized by Eratosthenes in his computation of the size of the Earth. Recall that the key to his clever solution was to measure the gnomon's shadow at different places on the Earth's surface. Thus, he is using the sun to measure the change in the way the gnomon (which we can interpret as a normal vector) sits in the three-dimensional space. The method of using the three-dimensional space that the surface lies in to measure its curvature is called the extrinsic measurement of curvature.

Gauss' great contribution to the geometry of surfaces is the observation that we can measure the curvature *intrinsically*; that is, we can compute the curvature of a surface without relying on the three-dimensional space that the surface lies in. (This observation is known as Theorem Egregium, or the Remarkable Theorem.) How can we do this? The answer again lies in our previous observation about the differences in geometry that we see in spherical, planar and hyperbolic geometry. Take, for example, a point on a sphere of radius R and consider a triangle that contains the point in the interior. As we learned earlier, if we measure the sum of the interior angle, we know that it is more than 180 degrees (or π radians). Now how much this is different from 180 degrees depends, not only on the size of the sphere, but also on the size of the triangle; if we take an equilateral triangle with side lengths 1 then the sum is 270 degrees or $3/2 \pi$ radians, and if we take an equilateral triangle that is very small, the sum is only slightly more than 180 degrees.



In fact, because of this dependence on the area of the triangle, the useful quantity to consider is

$$\frac{\text{sum of the interior angles of a triangle measured in radians} - \pi}{\text{area of the triangle}}.$$

On a plane, the ratio above is always equal to 0 since the sum of the interior angles of all triangles is equal to 180 degrees. It turns out that in a sphere of radius 1, the limit of this quantity as we take smaller and smaller triangles is equal to 1; and more generally, on a sphere of radius R , the limit is $\frac{1}{R^2}$. We define the curvature at a point on the arbitrary surface as the limit of this ratio. For example, every point on a sphere of radius R has a curvature of $\frac{1}{R^2}$. We note that we can also use the circumferences of circles of small radii to define curvature.

Let us summarize two important points that are made in the discussion above:

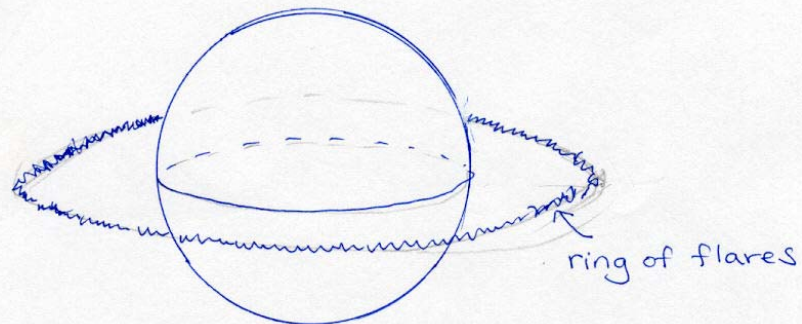
- (1) The curvature of a surface is not a vague notion indicating a contrast to flatness. In fact, there is a way to precisely assign a numerical value of curvature at a point in a surface.
- (2) This measurement does not depend on the way the surface curves into the three-dimensional space containing it. We can make this measurement intrinsically on the surface, without relying on extrinsic measurements outside of the surface.

These two points are important as we consider the curvature of higher dimensional spaces and were instrumental in Bernhard Riemann's (1826-1866) development of a branch of geometry now known as Riemannian geometry. We will restrict ourselves to trying to understand three-dimensional spaces, which we model our universe on. (Here, I should point out that in order to better grasp some widely accepted theories of the universe such as Einstein's General Relativity and the Big Bang Theory, scientists and mathematicians study a four-dimensional space-time continuum consisting of three spatial dimensions, with time counting as the fourth dimension. This is beyond the scope of this paper, and we will restrict ourselves to the study of the three spatial dimensions.)

Our intuitive feeling about the world around us is that the universe is flat; i.e. that the geometry of the universe resembles the geometry of the plane. But is this really an accurate description of the universe? Recall that for centuries upon centuries, the prevailing assumption was that the Earth was flat. Of course, the reason for that can be explained by the sense that, on the small scale, the Earth really does seem flat and there is no danger in assuming this. For example, if you take out a map of Dayton, Ohio, we can navigate around the city by trusting that this map, which is drawn on a flat piece of paper, is giving us an accurate description of the city. It is only when we try to navigate around a region covering a large portion of the Earth that a map gives inaccurate information about the location of places that you want to visit. For example, if you were piloting a plane from Chicago's O'Hare International Airport to Tokyo's Narita Airport and followed the path given by drawing a straight line on a Mercator projection map, you would be wasting a lot of time and fuel. A more efficient route (though that also depends on factors other than the geometry of the sphere such as wind currents and the rotation of the Earth) would take you more toward the North and over Alaska.

In a similar way, the flat universe is a reasonable assumption for living our everyday lives. In playing a game of basketball or in planning the construction of a new science and math building on the Dayton campus, the assumption that we live in a flat universe will not create any problems. Notice that these activities take place in a very small region of the universe. What happens if we use our flat universe assumption when we get on a spaceship and begin exploring vast regions of space? Analogous to using the flat Earth assumption when exploring a large portion of the Earth, you should expect that you would run into many problems. Moreover, from our discussions above, you should expect that a more accurate description of the universe involves the understanding of its geometry.

Riemann proposed that we should be able to measure the curvature of space. His idea can be explained by imagining the following experiment. Line up thousands and thousands of rockets around the equator of the Earth. These rockets will take off perpendicularly into the air, all at the same time, and travel for a fixed distance, say distance r , whereupon they fire off a flare that makes a ring around the Earth.



Now if the universe is truly flat, the circumference of the ring should be $2\pi r$. If not, we can measure the deviation of the circumference from $2\pi r$ to detect the curvature of the space around the Earth. Riemann proposed that, in this way, we could compute the numerical value of curvature at every point of our universe. (Technically speaking, this definition involves a plane contained in the three-dimensional space on which all the rocket flares lies in. Thus, curvature in space whose dimension is greater than 2 is measured at a point with respect to a chosen plane containing the point.) Notice that an important feature of this description of curvature is that we do not have to rely on our three-dimensional universe curving into some unknown fourth-dimensional space.

In Riemannian geometry, we study (two, three or higher dimensional) spaces in which we can compute curvature at every point. As can be seen from above, the curvature provides a local description (i.e. the geometry nearby a point of interest) of the universe that we live in. From this description, we can try to find a way to understand the global shape of the universe. As an example, let's try to understand the shape of the universe if we assume that it is spherical or like a three-dimensional sphere; in other words, we assume that the universe is a homogeneous space of positive curvature. This is very much analogous to the model of the Earth as a sphere; though the Earth might deviate slightly from a sphere on a local scale (such as mountains and valleys that make it not exactly a sphere), on a large or a global scale, the Earth looks much like a sphere. Recall that rays which propagate along the great circle from a point on a sphere converge to a point on the other side of the sphere. For example, two rays starting at the North Pole meet again at the South Pole. A point that is on the opposite side of a sphere (for example, the South Pole is opposite to the North Pole) is called its antipodal point. The implication of this is that if you travel straight ahead from a point on the Earth you eventually come back to the same point. That is, if you start from the North Pole and travel south, you will eventually reach the South Pole. But once you pass the South Pole, you start traveling northbound until you eventually get back to the North Pole.

Similarly, if the universe is of constant positive curvature, then we should expect that traveling from a given point on the universe would eventually lead back to the starting point. More precisely, suppose we take a space ship and travel into space in a straight line. Assuming that we do not run into any obstacles, we will eventually reach the antipodal point to our starting point and, once we pass by that, start heading back until we eventually end up back at the starting point again. Notice that this model gives us a picture of the universe as a finite object, yet it is a model that is devoid of any edges or boundary (just as the Earth does not have edges from which we can fall off). A finite universe seems to be consistent with some experimental data we currently have and with theoretical assumptions that is generally accepted in the scientific community, such as the Big Bang Theory.

Riemann proposed much more than a description of the universe as a spherical one. He went a step further in saying there is no reason we should believe that the curvature of space is constant at all points; in fact, the curvature of a point near our solar system may be quite different than the curvature at a point in a solar system of a galaxy far away. Still, the understanding of curvature can give us information about the general shape of our universe. In fact, trying to deduce the global shape of space from its local geometry (i.e. curvature) is a very important study in Riemannian geometry. Its understanding will shed light on the true nature of the universe that we live in and answer questions such as about its shape and its size.

In summary, curvature can be defined using the local geometry of space, such as the circumference of a circle of radius r and the sum of the interior angles of a triangle. This is known as the intrinsic definition of curvature, one that can be made from measurements within the space under consideration.

We hope that the information presented above has given you more of an the understanding of curvature and its impact on our understanding of cosmology. For a more detailed discussion, refer to Robert Osserman's delightful book, *The Poetry of the Universe*. This book is intended for a general audience and many of the topics discussed here were inspired by it. For those who want to study more geometry, we recommend Manuel DoCarmo's text book *Differential Geometry of Curves and Surfaces*. You will need to feel comfortable with material from a multivariable calculus course as a prerequisite.