Curved Space

The concept of curved space plays an important role in Einstein's general relativity. In this article, we introduce the concept of curved space without using too much complicated mathematical formulas.

1 Curved lines

First, we will begin with curved lines. In Fig. 1 we can see an example of two straight lines, and in Fig. 2 two curved lines.

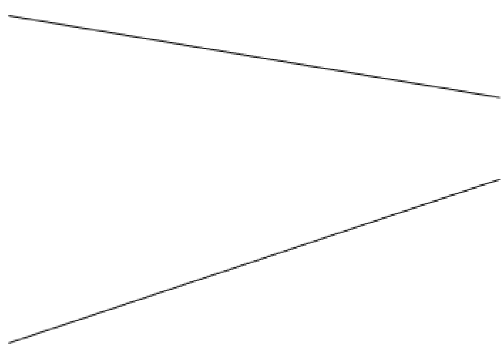

Figure 1: Representation of two straight lines

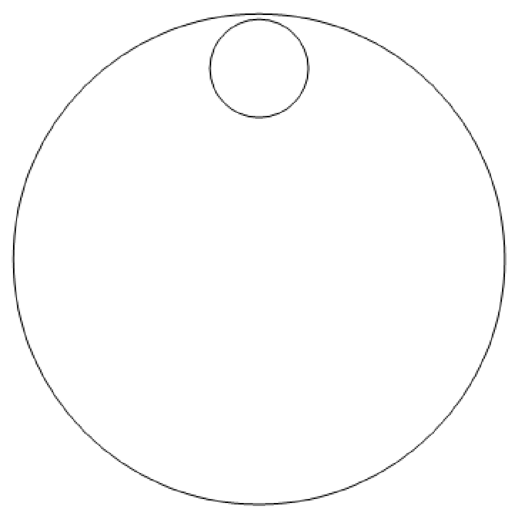

Figure 2: Representation of two circles with different radius. They are placed with their edges nearby to compare their curvatures.

Figure 1 doesn't need any explanation. It is obvious that the two lines are straight. They are not curved at all. In Fig. 2 you see two circles. Which one of them is more curved? The small circle or the big circle? For a better comparison, I placed the small circle near the big circle's edge. If you look closely at the region where they nearly touch (the upper part of the figure), it is obvious that the small circle is more curved than the big circle. So, if we assign a number called "curvature" that tells you how curved a line is, the curvature of the small circle must be larger than the one of the big circle. As a smaller circle has a smaller radius, if we define the curvature of a circle as the inverse of its radius, then the curvature of a smaller circle will be bigger than the curvature of a bigger circle. For Fig. 2, the radius of the smaller circle is 5 times smaller than the bigger one. Therefore, the curvature of the small circle is 5 times bigger than the curvature of the big circle. If the radius of the small circle is 2, then its curvature is 0.5. If the radius of the big circle is 10, which is 5 times 2, then its curvature is 0.1. The curvature of the straight lines in Fig. 1 is 0 because it's not curved at all.

Each point of the circle has the same curvature, namely the inverse of its radius. Nevertheless, not all lines have the same curvature at every point, like we show in Fig. 3.

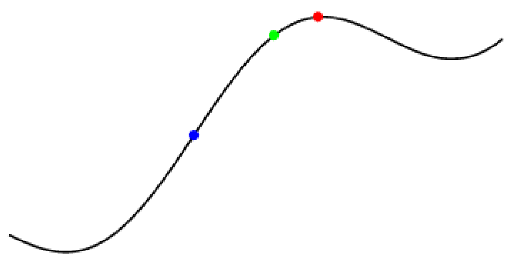

Figure 3: Schematization of a curved line. It has three arbitrary points marked along.

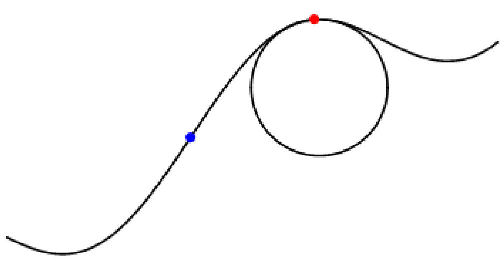

Figure 4: Curved line from Fig. 3. The circle on the red point represents its curvature.

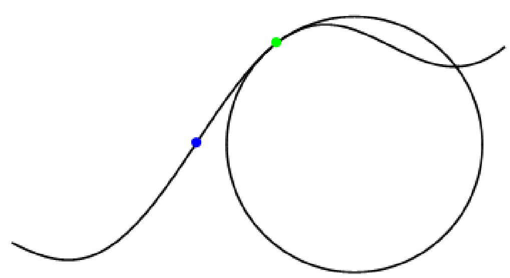

Figure 5: Same as in Fig. 4 but with a circle representing the curvature on the green dot.

You can see that, at the blue point, the line is not curved at all. However, it is curved at the green point and more curved at the red one. So, how can we calculate the curvature at these points? We can find the curvature at the red point by fitting a circle like we show in Fig. 4. Then, the curvature at the red point is given by the inverse of the radius of that circle. The same can be said about the green point (see Fig. 5). You see that the fitted circle is larger than the circle in Fig. 4. Indeed, we see that the green point has a smaller curvature than the red point. On the other hand, no circle would be able to fit a point like the blue point. The circle would have to be infinitely large, as only in such a case the arc of the circle would look close to a straight line. As the inverse of the infinite is zero, indeed the curvature at the blue point is zero. The same statement can be made for the two lines in Fig. 1.

2 Curved planes

2.1 Gaussian curvature

Now, let's talk about curved planes. In Figs. 6 and 7 we have two examples of (2-dimensional) curved planes. The surface in Fig. 6 is an example of a positively curved plane and the one in Fig. 7 is an example of a negatively curved plane. To explain why Fig. 6 is positively curved and Fig. 7 is negatively curved, look carefully the two thick lines in each figure.

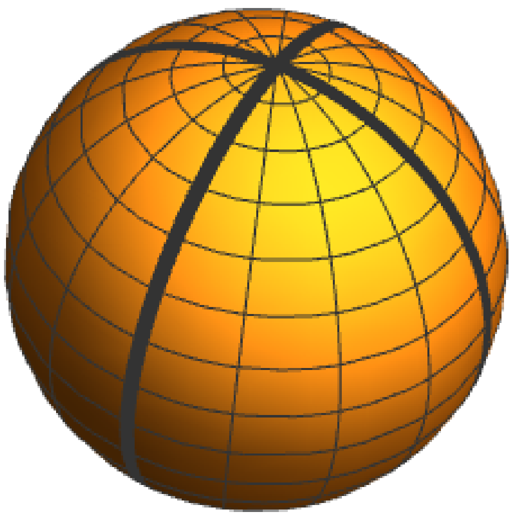

Figure 6: 2D plane curved to form a sphere. It has two curves (meridians) highlighted with a thicker trace.

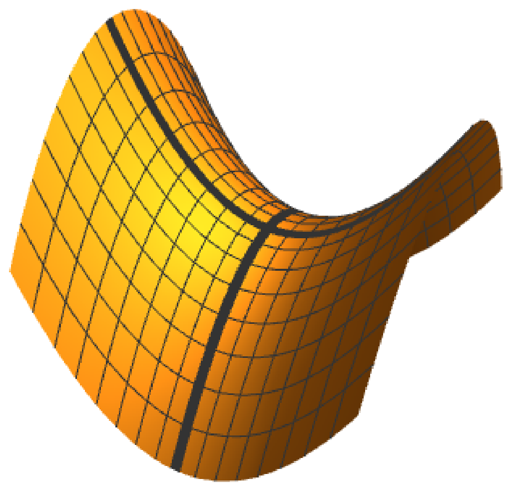

Figure 7: 2D plane curved as a saddle. As in Fig. 6, it has two curves highlighted with thicker trace.

Precisely speaking, let me show you why the North Pole in Fig. 6 (where the two thick lines meet) has a positive curvature, and the saddle point in Fig. 7 (where the two thick lines meet) has a negative curvature. If we find two orthogonal (i.e., perpendicular) lines that descend from the North Pole and have the steepest slopes, they are both downwards, i.e., the same direction (actually, in the case of the North Pole on a sphere, the slopes of all the descending lines are the same, so it doesn't matter which ones you choose, as long as they are orthogonal). If they are curved in the same direction, the curvature at that point is positive. In the case of the saddle point, between the two orthogonal steepest lines, one line is curved upward while the other one is curved downward. As they are curved in opposite directions, the curvature at the saddle point is said to be negative. Actually, Fig. 6 is a sphere. So, regarding the curvedness, each point is equivalent, and therefore, every point has the same curvature. The same cannot be said for Fig. 7.

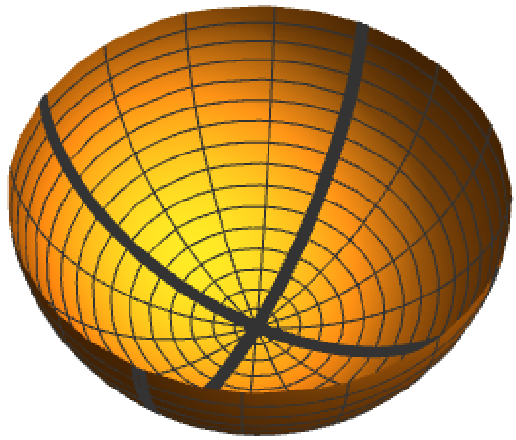

Figure 8: Same sphere as in Fig. 6, but only the lower hemisphere is represented. The two thicker lines from Fig. 6 meet again in the South Pole.

To better illustrate my point, please see Fig. 8. We want to calculate the curvature at the South Pole of the sphere. To see the two orthogonal lines clearly, I only drew the lower hemisphere. The two orthogonal slopes are both upwards, i.e., the same direction, so the curvature at the South Pole is, again, positive.

What I am talking about now is called "Gaussian curvature". The German mathematician Carl Friedrich Gauss, came up with this concept in 1827. The Gaussian curvature is given by the product of the curvatures of the two orthogonal steepest lines. For example, in the case of Fig. 6, if the radius of the sphere is R, the curvature of each line is $1/R$, so the Gaussian curvature is given by

$$1/R \times 1/R = 1/R^2. \tag{1}$$

Thus, we see that the bigger a sphere is, the less it is curved, which is precisely what we wanted. As the Earth is so big, we don't really feel that it is curved. This is exactly the reason why the crazy Flat Earth people say that the Earth is flat!

Actually, we can assign a sign to the curvature of line. In the case of calculating the Gaussian curvature at the North Pole, we assumed that the downward direction has a positive curvature. To be consistent with this assumption, if a line is curved upwards, it has to be a negative curvature. Now, if we see again Fig. 8, both lines are upward, so if we calculate the Gaussian curvature, we obtain

$$(-1/R) \times (-1/R) = 1/R^2 \tag{2}$$

Again, we see that it has a positive Gaussian curvature. In the case of Fig. 7, one of the two orthogonal lines has a positive curvature and the other a negative one. Thus, we conclude that it has a negative Gaussian curvature, as the product of a positive number and a negative number is always negative.

Now, let's move on to the case of zero Gaussian curvature. In Fig. 9, we have a flat plane, so it obviously has a zero Gaussian curvature and, in Fig. 10 we have a cylinder. Now, I claim that a cylinder has a zero Gaussian curvature. Why is it so? Look at the two thicker orthogonal lines. One is a straight line, so it has a zero curvature, and the other one is a curved line, which has a certain non-zero curvature.[1] So, if you multiply zero by a non-zero number, you get zero. Therefore, it has a zero Gaussian curvature.

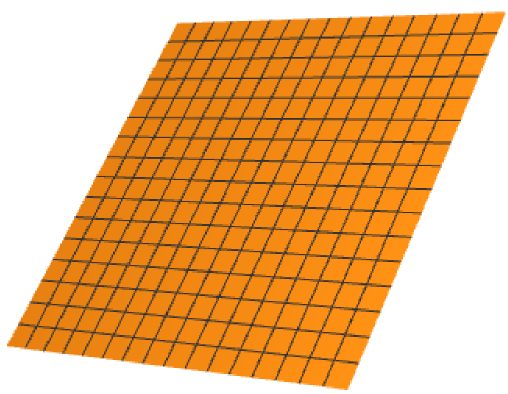

Figure 9: A representation of a flat plane.

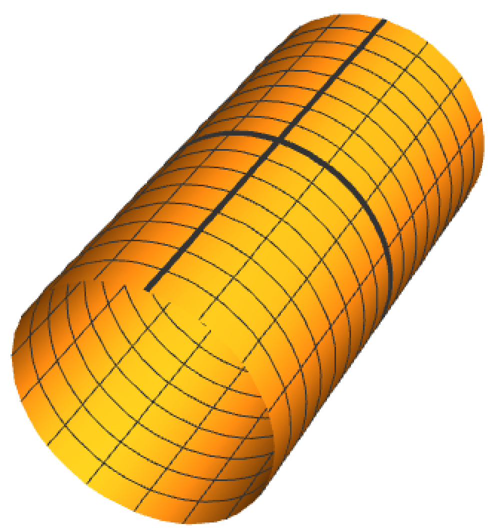

Figure 10: A cylinder with two thicker lines drawn.

This may sound strange considering that a cylinder is curved, but think about it this way. If you have a piece of paper, just like Fig. 9, then you know that it has a zero Gaussian curvature. Now, by rolling it, you can turn it into something like Fig. 10. If you can do so without cutting or pinching or wrinkling or crumpling, the Gaussian curvature remains the same. For example, note that you cannot turn a flat paper into Figs. 6 or 7 and vice versa. If we could, a flat world map would have accurately represented the Earth. You may be familiar with such a fact that Greenland is represented on the map much bigger than it

[1]You may note that the straight line is not the steepest slope, so it cannot be one of the two orthogonal lines that I talked about. Actually, my use of the word "steepest" line was not accurate. A more exact way of saying would be picking out the maximum and the minimum values of curvature among the lines that emanate from the point. The minimum is zero, and the maximum is $1/R$ if the radius of the circle in the cylinder is R.

is, because it is so close to the North Pole. A flat world map can never accurately represent the Earth.

So far, we only talked about planes that have positive Gaussian curvatures at every point, or negative Gaussian curvatures at every point, or zero Gaussian curvatures at every point, but it is also possible that a plane has positive curvatures at certain regions, and negative curvatures at others. A "torus" is such an example (see Fig. 11).

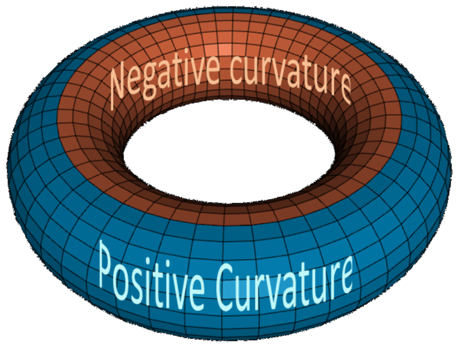

Figure 11: Graphical representation of a torus. The region with positive curvature is represented in color blue and the region with negative curvature in color red.[1]

2.2 Properties of curved surfaces

Now, let's view some properties of flat, positively curved, and negatively curved surfaces. For a positively curved surface, we will focus on the sphere as an example, because it is conceptually easy to deal with as it has constant curvature.

2.2.1 Flat surface

Parallel lines are defined as two straight lines that never meet. In the flat space, we know that there is only one straight line that is parallel to another given straight line and passes through given point (see Fig. 12 for an example). There is only one line that passes the green point and is parallel to the black line. That is the blue line. Notice that both lines stretch to infinity as the arrows indicate, but they never meet.

2.2.2 Sphere, a positively curved surface

However, there is no such parallel straight lines on a sphere, as *any* two straight lines on a sphere always meet at two points. You can see this in Fig. 13. I will argue why it is always so in the next two paragraphs. I will assume that you are already familiar with the words

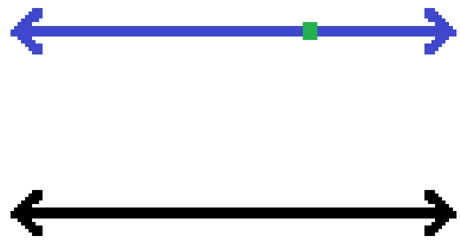

Figure 12: Graphic schematization of two straight lines. The arrows indicate that the lines extend through infinity. There is only one straight line (blue trace) that is parallel to a given line and passes through a given point.

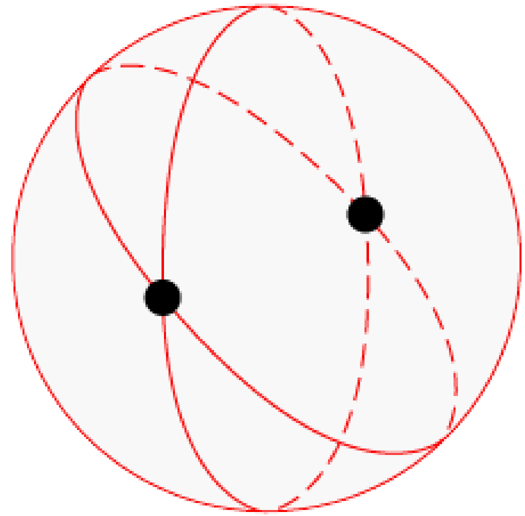

Figure 13: two straight lines on a sphere always meet at two points

describing the Earth, such as "the Northern hemisphere", "the Southern hemisphere" and "the equator". If you are not, you can first read our later article "the round Earth" or skip the next two paragraphs.

What would a straight line in a sphere look like? To find the answer, think of the sphere as the Earth, as the Earth looks approximately like a sphere. Let's say you choose a point on a sphere. Without loss of generality, let's find a coordinate system that defines the North Pole as the point you chose. Then, any direction from the North Pole is southward. If you keep moving, without changing the direction, you will always reach the South Pole. Once you reach the South Pole, if you keep going in the same direction, the direction will be northward, and you will come back to the North Pole. More generally, if you start at a certain point, and if you keep going straight, you will reach the "antipodal point" of the original point, then come back to the original point. The antipodal point of a point is the point directly opposite to it. For example, the antipodal point of the North Pole is the South Pole.

Given this, suppose you have a straight line. You know that this straight line divides a sphere into two equal parts. For example, the equator, which is a straight line, divides the Earth into the Northern hemisphere and the Southern hemisphere. Let's call one of the two such equal parts N, and the other S. For example, if the equator is the straight line that divides the sphere, the Northern hemisphere will be N and the Southern hemisphere will be S. Now, let's pick a point A. If the point A is in N, the antipodal point of A must be in S. Given this, recall that any straight line that passes A must pass its antipodal point. Since the antipodal point of A is in S, the new straight line from A has to pass the first straight line (i.e., in our specific example, equator) twice: when it crosses N to S and when it crosses from S to N. The same statement can be made when A is in S and its antipodal point is in

N. So, in a sphere, any two straight lines always meet. In fact, they meet twice. Therefore, there can be no parallel lines.

Another interesting property of straight lines in a sphere is that the sum of the angles of a triangle is always bigger than 180°. See Fig. 14 for an example.

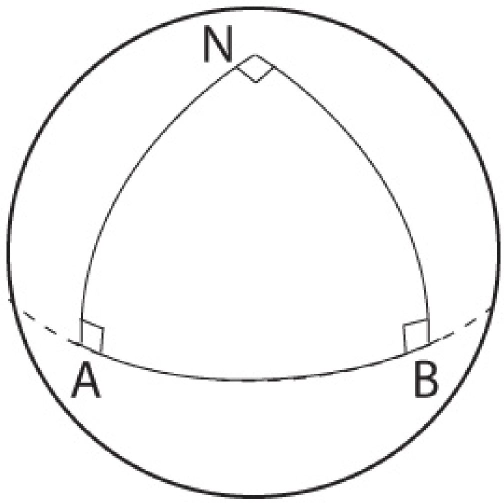

Figure 14: Schematic representation of a spherical surface. The sum of the angles of the triangle formed by the points A, B and N is 270°

One of the vertices of the triangle is located at the North Pole and the other two are located at the equator. This is done in such a way that the length of the three sides is the same. In this case, the sum of the three angles is 270°. It is also easy to imagine that if the triangle is smaller, the effect of curvature is smaller, making the sum of the angles in triangle closer to 180°. Think it along this way. If you confine your movement in your city and draw a triangle, you will hardly notice that the Earth is round and the sum of the angles in your triangle will be very close to 180°. That's because the triangle is so tiny compared to the Earth. Actually, it is mathematically proven that the bigger the area of the triangle the bigger the sum of the angles of the triangle. Also, it is actually shown that if two triangles have the same area, then the sum of the angles inside them is the same, and vice versa, regardless of their shape. In our article "The Gauss-Bonnet theorem for triangle on a sphere," we will explicitly express the area of triangle as the sum of its angles. Of course, as we will see, when the sum of its angle is exactly 180°, the area of the triangle is exactly zero. Only when the effect of curvature is completely off is the sum of the angles of triangle is exactly 180°.

2.2.3 Negatively curved surface

On the contrary, the sum of the angles of a triangle on a negatively curved surface is always less than 180°. We show an example in Fig. 15. Each angle in the triangle is not broad as was the case in Fig. 14, but quite narrow. Also, in such a surface, there are infinitely many parallel lines to a given line and passing a given point. It is because the distance between two straight lines get farther and farther as you go farther and farther, actually in both directions. This is shown well on the right bottom corner of Fig. 15. Remember that, in the case of sphere, two straight lines always meet at two points, which mean that they get closer and closer and eventually meet as you go in both directions.

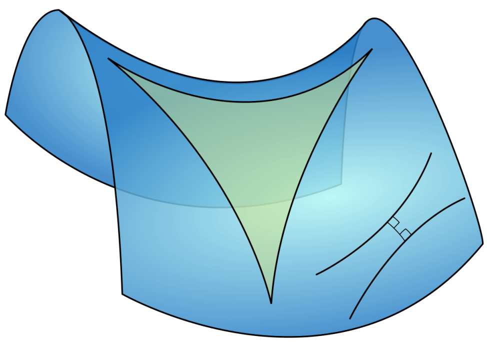

Figure 15: the sum of the angles of the triangle in a negative curved space is always less than 180°.[2]

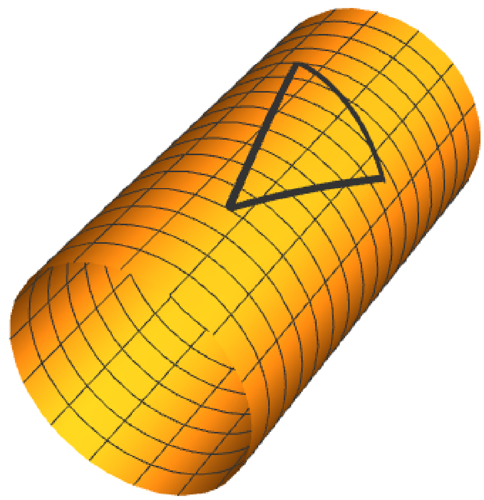

Figure 16: the sum of the angles of the triangle in a cylinder is always 180°

2.2.4 Cylinder

In the case of a flat space, you already know that the sum of the angles of a triangle is exactly 180°. This is also true for a cylinder which has a zero Gaussian curvature (see Fig. 16). If you draw a triangle on a flat paper and roll it to make a cylinder, the angles in the triangle won't change. Therefore, the sum of the angles remains the same, namely 180°.

3 Comments

Before we reach the end I'd like to highlight three comments.

3.1 Relation between the finiteness and the sign of curvature

If an object has a positive curvature everywhere, it necessarily has a finite size. This can be proven mathematically rigorously, but it is also intuitively clear. If we draw the two orthogonal lines that we have mentioned at each point, they are curved toward the same side to have a positive Gaussian curvature. It means the two orthogonal lines cannot "run away" from each other, as was the case in Fig. 7 with a negatively curved surface. Speaking of an everywhere negatively curved surface, it has an infinite size if it doesn't have a boundary (i.e., if the surface is discontinued at certain points). The two orthogonal lines always run away from each other. Notice that an everywhere positively curved surface has a finite size, even if it doesn't have a boundary; the surface closes itself.

3.2 Euclid's fixth axiom

A book that is published in the greatest number of editions is definitely the Bible. The Bible has been translated into many languages, including the ones that only small number of people speak. Then, what is the book that is published in the second greatest number of editions? That is the *Elements* written by the Greek mathematician Euclid around 300 BC. Over thousand editions have been published. The *Elements* is the most successful textbook in human history and deals with several branches of mathematics such as geometry and number theory.

In mathematics, one begins with "axioms." Axioms are a set of rules that one assumes as basis in mathematics to construct theorems. The *Elements* also lists axioms to construct theorems such as Pythagorean theorem or the theorem that says that the two angles of an isosceles triangle are equal. Here are the five axioms in the *Elements*.[3]

> Let the following be postulated:
>
> 1. To draw a straight line from any point to any point.
>
> 2. To produce [extend] a finite straight line continuously in a straight line.
>
> 3. To describe a circle with any centre and distance [radius].
>
> 4. That all right angles are equal to one another.
>
> 5. That, if a straight line falling on two straight lines make the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which the angles are less than two right angles.

The fifth axiom is often called "the parallel postulate." If you find it hard to understand, let me explain it by showing Fig. 17. As $\alpha + \beta$ is smaller than the sum of two right angles (i.e., 180°), the two lines meet on the side where the two dotted lines are, instead of the other side (i.e., the left side in the figure, where the two lines are moving farther apart).

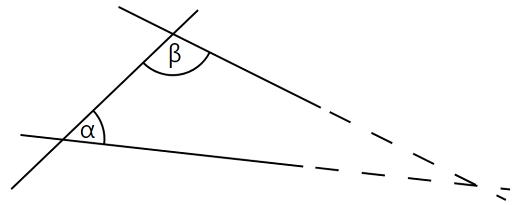

Figure 17: Two straight lines which form angles α and β such that α, $\beta < 180°$.[4]

Many people doubted that one needed the parallel postulate as a separate axiom, and thought that one could probably "derive" (i.e., "prove") this axiom from the other four axioms. Usually, we take certain statements as axioms, as we all can easily agree with them. However, Euclid's fifth axiom didn't seem self-evident, while all the other four axioms did. If something is not self-evident, it should not be taken for granted, but it should be explained and follow from others by logic.

In other words, many mathematicians thought that the parallel postulate was redundant. They thought that the parallel postulate must be automatically satisfied, without the need to be assumed as a separate axiom, as long as all the other four axioms are satisfied. However, none of them succeeded in proving the parallel postulate from the other four axioms.

If you read this article so far, you will understand why they failed. This axiom is true only in flat space. Think about two straight lines on the sphere. They meet at both sides, no matter whether the sum of the two angles is larger or smaller than 180°. Think about two straight lines on a negatively curved surface. They do not necessarily meet even though the sum of the two angles are smaller than 180°.

The geometry which satisfies the parallel postulate is called "Euclidean geometry". Such a geometry is necessarily on a flat surface or flat space. If the parallel postulate is not satisfied, we call it "non-Euclidean geometry". Geometries on curved spaces are such cases. We will talk about non-Euclidean geometry quantitatively in our article "Non-Euclidean geometry".

A side note. Before non-Euclidean geometry was established, many mathematicians indeed worked on the proof of the parallel postulate. For example, in the 17th century, John Wallis thought that he had proved the parallel postulate by assuming that, for any given triangle, there always exists another triangle similar to this triangle (i.e., having the same shape), but with any arbitrary size. In other words, his assumption was that two triangles can be similar, even though they are not congruent. Note that, if two triangles are not congruent, but similar, their sizes are necessarily different. However, his assumption is not true if the parallel postulate is violated. As we have briefly mentioned earlier in this article, two triangles with the same sum of the angles inside each of them have the same area, if these two triangles are on a sphere instead of a flat space. Therefore, if two triangles are similar, which means they have the same angles, the two triangles have the same area, which implies that they are congruent. In other words, on a sphere, two triangles are similar, *only* if they are congruent; if two triangles on a sphere have the same shape, they have the same size.

3.3 Riemannian geometry

What we have talked about, things such as curved space is treated in a branch of mathematics called "differential geometry", which Gauss first developed. However, notice that his treatment is confined to 2-dimensional surfaces in a 3-dimensional space. In mathematics, one has the freedom to think about higher dimensions, and higher dimensional surfaces, as we mentioned in our earlier article "Manifold". The German mathematician, Bernhard Riemann, who received his doctoral degree under Gauss, succeeded in treating curved higher-dimensional surfaces. Thus, not only a 2-dimensional surface can be curved, but also the 3-dimensional space we live can be curved. If our 3-dimensional space has a positive curvature, if we draw a big triangle in our Universe, the sum of its angles will be larger than $180°$, and if it has a negative curvature, it will be smaller than $180°$. It's hard to imagine, but it must be true.

Einstein had to learn Riemannian geometry from his mathematician friend Marcel Grossmann before he developed general relativity. General relativity is written in the language of Riemannian geometry. In our later article "An Introduction to General Relativity," we will teach you Riemannian geometry first, then general relativity, as is the case with any general relativity course at university or graduate school.

When I took a general relativity class at Harvard, a student who apparently knew some Gaussian geometry asked the professor about the connection between the Gaussian curvature and the curvature in Riemannian geometry. The professor dismissed the question and noted that he never mentioned the Gaussian curvature in the course. Actually, it is possible to study the Riemannian geometry with zero knowledge on Gaussian geometry and, in fact, virtually all general relativity books do not mention anything about Gaussian geometry.

After taking the general relativity class, I took a class titled "Lorentzian geometry and Riemannian geometry." It was almost like a general relativity class but taught by a mathematician. There, the math professor mentioned the connection between the Gaussian curvature and the curvatures in the Riemannian geometry, but I forgot the exact connection because I never had a chance to use it again.

Summary

- The curvature of a circle is given by the inverse of its radius.
- The curvature of a line at a point is given by the curvature of a circle that fits at that point.
- The Gaussian curvature of a surface at a given point is given by the product of the curvatures of the two orthogonal lines which have the steepest slopes. If the two orthogonal lines are curved toward the same side, the Gaussian curvature is positive,

and if the two orthogonal lines are curved in opposite directions to each other, the Gaussian curvature is negative.

- The Gaussian curvature of a cylinder is zero because one of the two orthogonal lines that we just mentioned have zero curvature; zero multiplied by any number is always zero.
- An Euclidean space is the space in which Euclid's fifth axiom ("the parallel postulate") is satisfied.
- In the Euclidean space, there is only one straight line that is parallel to another certain straight line and passes a certain point.
- On a sphere, which has a constant positive curvature, there is no such parallel line.
- On a negatively curved space, there are infinitely many such lines.
- On a positively curved surface, the sum of the angles of a triangle is always larger than 180°.
- On a negatively curved surface, the sum of the angles of a triangle is always smaller than 180°.

References

[1] Adopted from `https://commons.wikimedia.org/wiki/File:Torus_Positive_and_negative_curvature.png`.

[2] `https://commons.wikimedia.org/wiki/File:Hyperbolic_triangle.svg`

[3] Heath, Thomas L. (1956). *The Thirteen Books of Euclid's Elements* (2nd ed. [Facsimile. Original publication: Cambridge University Press, 1925] ed.). New York: Dover Publications. Vol.1, p195-202, reproduced in `https://en.wikipedia.org/wiki/Euclidean_geometry`

[4] `https://commons.wikimedia.org/wiki/File:Parallel_postulate_en.svg`