Contents

2 Basics of general relativity **15**
2.1 The principles of general relativity 15
2.2 Curved 2D and 3D space 16
2.3 Metric of 2D and 3D space 17
2.4 4D flat spacetime 19
2.5 Curved spacetime 19

Contents

2 Basics of general relativity

2.1 The principles of general relativity

The theory of general relativity that was uncovered by Albert Einstein in 1915 (and nearly simultaneously by David Hilbert) is our current best theory of gravity. General relativity replaced the previous theory, Newtonian gravity, which can be understood as a limit of general relativity in the case of isolated systems, slow motions and weak fields. General relativity has been extensively tested during the past century, and no deviations have been found, with the possible exception of the accelerated expansion of the universe, which is however usually explained by introducing new matter rather than changing the laws of gravity [1]. We will not go through the details of general relativity, but we will give some rough idea of what the theory is like, and introduce a few concepts and definitions that we need.

The principle behind special relativity is that space and time together form four-dimensional spacetime. The essence of general relativity is that gravity is a manifestation of the *curvature* of spacetime. While in Newton's theory gravity acts directly as a force between two bodies[1], in Einstein's theory the gravitational interaction is mediated by spacetime. In other words, gravity is an aspect of the geometry of spacetime. Matter curves spacetime, and this curvature affects the motion of other matter (as well as the motion of the matter generating the curvature). This can be summarised as the dictum "matter tells spacetime how to curve, spacetime tells matter how to move" [2]. From the viewpoint of general relativity, gravity is not a force. If there are no forces (due to particle physics interactions) acting on a body, the body is in *free fall*, also known as *inertial motion*. A freely falling body moves along a straight line in curved spacetime, called a *geodesic*. Forces cause the body to deviate from geodesic motion. It is important to remember that the viewpoint is that of *spacetime*, not just space. For example, the orbit of the Earth around the Sun is curved in space, but straight in spacetime.

If a spacetime is not curved, it is called *flat*, which just means that it has the geometry of Minkowski space. In the case of space (as opposed to spacetime), flat means that the geometry is Euclidean. (Note the possibly confusing terminology: Minkowski space*time* is called simply Minkowski space!) In Newtonian theory, gravity is just an interaction between particles, but in general relativity, it has dynamics

[1] It is also possible to formulate Newtonian gravity in geometric terms, so that gravity is an expression of spacetime curvature, although this is less natural than in general relativity.

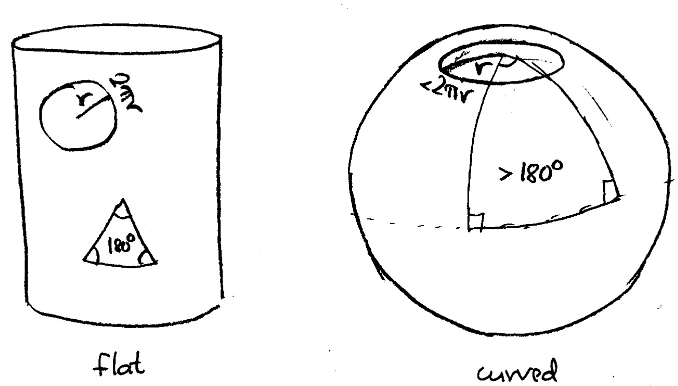

Figure 1: Cylinder and sphere.

of its own and its degrees of freedom are described by the *metric*. The equation of motion of the metric, sourced by the matter of the universe, is called the Einstein equation. We will below go through the concept of the metric; we will get to the Einstein equation in the next chapter.

2.2 Curved 2D and 3D space

To help to visualise a four-dimensional curved spacetime, it may be useful to consider curved two-dimensional spaces embedded in flat three-dimensional space.[2] Imagine there are 2D beings living in this 2D space who have no access to the third dimension. They can determine whether the space is curved by checking whether the laws of Euclidean geometry hold. If the space is flat, then the sum of the angles of any triangle is 180°, and the circumference of any circle with radius r is $2\pi r$. If by measurement they find that this does not hold for some triangles or circles, then they can conclude that the space is curved.

A simple example of a curved 2D space is the sphere. The sum of angles of any triangle on a sphere is greater than 180°, and the circumference of any circle drawn on the surface of a sphere is less than $2\pi r$. (Straight lines on the sphere are sections of *great circles*, which divide the sphere into two equal hemispheres.)

In contrast, the surface of a cylinder has Euclidean geometry, i.e. there is no way that 2D beings living on it could conclude that it differs from a flat surface, and thus by our definition it is a flat 2D space. (By travelling around the cylinder they could conclude that their space has a non-trivial *topology*, but the geometry is flat.)

In a similar manner we could try to determine whether the 3D space around us is curved, by measuring whether the sum of angles of a triangle is 180° or whether a sphere with radius r has surface area $4\pi r^2$. The space around Earth is indeed curved by the gravity of the Earth and Sun, but the curvature is so small that very precise

[2] This embedding is only a visualisation aid. A curved 2D space is defined completely in terms of its two independent coordinates, without any reference to a higher dimension. The geometry is given by the metric (part of the definition of the 2D space), which is a function of these coordinates. Some such curved 2D spaces have the same geometry as a 2D surface in flat 3D space. We then say that the 2D space can be embedded in flat 3D space. But there are curved 2D spaces which have no such corresponding surface, i.e. not all curved 2D spaces can be embedded in flat 3D space.

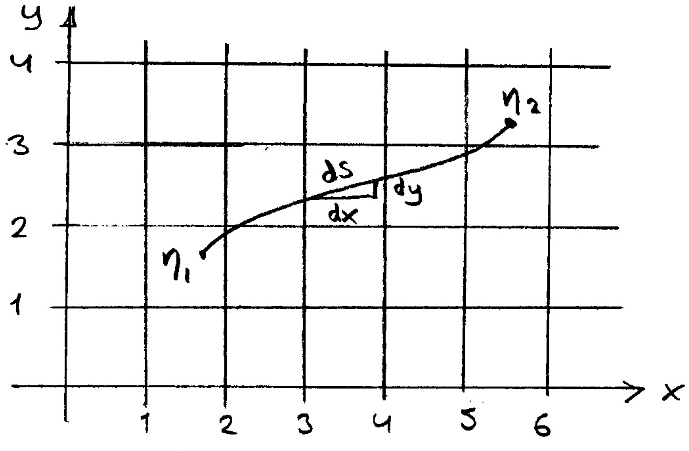

Figure 2: A parametrised curve in Euclidean 2D space with Cartesian coordinates.

measurements are needed to detect it. The effect of *spacetime* curvature, however, manifests itself as gravity, which is very easy to detect.

2.3 Metric of 2D and 3D space

The geometry of space is described by the *metric* that gives the distances between spacetime points. The metric is given in terms of a set of coordinates. The coordinate system can be arbitrary. The coordinates are a set of numbers that identify points in spacetime: they have no other identity a priori. For example, they do not necessarily directly correspond to physical distances or time intervals: that information is carried by the metric.

To introduce the metric, let us consider Euclidean two-dimensional space with Cartesian coordinates x, y. Take a parametrised curve $x(\eta)$, $y(\eta)$ that begins at η_1 and ends at η_2. The length of the curve is

$$s = \int \mathrm{d}s = \int \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \int_{\eta_1}^{\eta_2} \sqrt{x'^2 + y'^2}\mathrm{d}\eta \ , \tag{2.1}$$

where $x' \equiv \mathrm{d}x/\mathrm{d}\eta$, $y' \equiv \mathrm{d}y/\mathrm{d}\eta$. Here $\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}$ is the *line element*. The square of the line element is

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 = \sum_{i,j} \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j \ , \tag{2.2}$$

where δ_{ij} is the Kronecker delta defined as 1 when $i = j$, and 0 otherwise. In general relativity, the metric is the matrix of coefficients in front of the coordinate differentials. So in this case the metric is just δ_{ij}. Sometimes the line element itself is called the metric, as we can read off the metric from it, and we will adopt this sloppy language. On this course we will not need the matrix form of the metric, the line element will be sufficient.

We can use another coordinate system to describe the same 2D Euclidean space, such as polar coordinates. Then the metric is

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\varphi^2 \, , \tag{2.3}$$

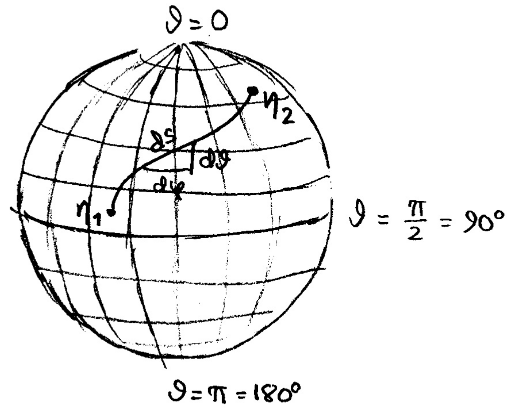

Figure 3: A parametrised curve on the two-sphere with spherical coordinates.

giving the length of a curve as

$$s = \int \mathrm{d}s = \int \sqrt{\mathrm{d}r^2 + r^2 \mathrm{d}\varphi^2} = \int_{\eta_1}^{\eta_2} \sqrt{r'^2 + r^2 \varphi'^2} \mathrm{d}\eta \,. \tag{2.4}$$

In a similar manner, in 3-dimensional Euclidean space, the metric is

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \tag{2.5}$$

in Cartesian coordinates, and

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2 \mathrm{d}\theta^2 + r^2 \sin^2\theta \mathrm{d}\varphi^2 \tag{2.6}$$

in spherical coordinates (where the r coordinate has the dimension of distance, but the angular coordinates θ and φ are dimensionless).

Now we can consider our first example of a curved space, the two-sphere. Let the radius of the sphere be a. As the two coordinates on this 2D space we can take the angles θ and φ. We get the metric from the Euclidean 3D metric in spherical coordinates by setting $r = a$,

$$\mathrm{d}s^2 = a^2 \left(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2 \right) \,. \tag{2.7}$$

The length of a curve $\theta(\eta), \varphi(\eta)$ on this sphere is given by

$$s = \int \mathrm{d}s = \int_{\eta_1}^{\eta_2} a \sqrt{\theta'^2 + \sin^2\theta \varphi'^2} \mathrm{d}\eta \,. \tag{2.8}$$

For later application in cosmology, it is instructive to consider the coordinate transformation $\rho = \sin\theta$. Since now $d\rho = \cos\theta \mathrm{d}\theta = \sqrt{1-\rho^2}\mathrm{d}\theta$, the metric becomes

$$\mathrm{d}s^2 = a^2 \left(\frac{d\rho^2}{1-\rho^2} + \rho^2 \mathrm{d}\varphi^2 \right) \,. \tag{2.9}$$

For $\rho \ll 1$ (in the vicinity of the North Pole), this metric is approximately the same as in (2.3), so the metric looks flat and the coordinates look like polar coordinates. As r grows, the deviation from flat geometry becomes more apparent. NWe run into a problem when $\rho = 1$. This corresponds to $\theta = 90°$, i.e. the equator. After this $\rho = \sin\theta$ begins to decrease again, repeating the same values. Also, at $\rho = 1$, the $1/(1-\rho^2)$ factor in the metric diverges. There is a *coordinate singularity* at the equator. There is nothing wrong with the space itself, but our chosen coordinate system covers only half of the space, the region north of the equator.

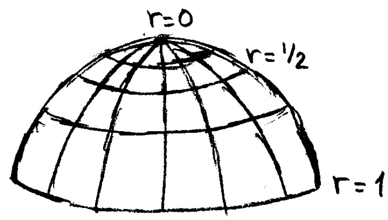

Figure 4: The part of the sphere covered by the coordinates in (2.9).

2.4 4D flat spacetime

The coordinates of the four-dimensional spacetime are (x^0, x^1, x^2, x^3), where $x^0 = t$ is a time coordinate. Some examples are Cartesian (t, x, y, z) and spherical (t, r, θ, φ) coordinates. The metric of the Minkowski space is

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \ , \tag{2.10}$$

in Cartesian coordinates. In spherical coordinates it is

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\,\mathrm{d}\varphi^2 \ . \tag{2.11}$$

The fact that time appears in the metric with a different sign reflects the special geometric features of Minkowski space. (We assume that the reader is familiar with special relativity, and won't go into details.) There are three kinds of distance intervals,

- timelike, $\mathrm{d}s^2 < 0$
- lightlike, $\mathrm{d}s^2 = 0$
- spacelike, $\mathrm{d}s^2 > 0$.

The lightlike directions form the observer's future and past *light cones*. Light moves along the light cone, so everything we see with light lies on our past light cone, and we can receive signals slower than light from everywhere inside it. To see us as we are now (using light), the observer has to lie on our future light cone, and we can send timelike signals to everywhere inside it. As we move in time along our world line, we drag our light cones with us so that they sweep over the spacetime. The motion of a massive body is always timelike, and the motion of massless particles is always lightlike.

2.5 Curved spacetime

In general relativity, spacetime is curved, whereas Minkowski spacetime is flat. (Recall that when we say space is flat, we mean it has Euclidean geometry; when we say spacetime is flat, we mean it has Minkowski geometry.) As in special relativity, the (proper) length of a spacelike curve is $\Delta s \equiv \int \mathrm{d}s$. Light moves on lightlike curves with $\mathrm{d}s^2 = 0$, massive objects along timelike curves with $\mathrm{d}s^2 < 0$. In this course, we will only consider diagonal metrics of the form

$$\mathrm{d}s^2 = -A^2\mathrm{d}t^2 + B^2\mathrm{d}x^2 + C^2\mathrm{d}y^2 + D^2\mathrm{d}z^2 \ , \tag{2.12}$$

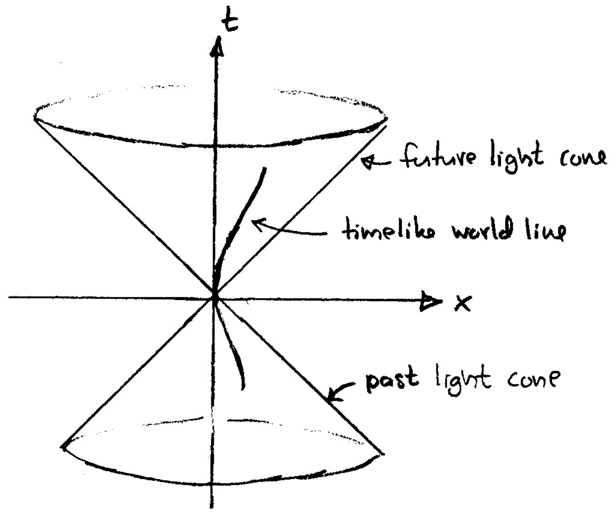

Figure 5: The light cone.

where the coefficients A, B, C, D are in general functions of t and x^i. This can also be the case in flat spacetime, as we saw with the spherical coordinates. However, in flat spacetime there exists a coordinate system (Cartesian coordinates) so that A, B, C, D are constant. In contrast, in curved spacetime there is no such coordinate system, and at least some of the coefficients vary across spacetime.

The three-dimensional subspace of spacetime, or *hypersurface*, defined by $t =$ constant called the space (or the *universe*) at time t, or a *time slice* of the spacetime. It is possible to slice the same spacetime in different ways i.e. make different choices of the time coordinate t. The volume of a region of space (given by some range in the spatial coordinates x, y, z) is

$$V = \int_V dV = \int_V BCD\mathrm{d}x\mathrm{d}y\mathrm{d}z \ , \tag{2.13}$$

where $dV \equiv BCD\mathrm{d}x\mathrm{d}y\mathrm{d}z$ is the *volume element*. Similarly, the surface area of a two-dimensional spatial region defined by $t =$ constant and $z =$ constant is

$$S = \int_S \mathrm{d}S = \int_S BC\mathrm{d}x\mathrm{d}y \ , \tag{2.14}$$

where $dS \equiv BC\mathrm{d}x\mathrm{d}y$ is the *surface element*. Continuing to one dimension, the proper length of the line of constant t, y, z is

$$L = \int_L \mathrm{d}L = \int_L B\mathrm{d}x \ . \tag{2.15}$$

References

[1] C. M. Will, *The Confrontation between general relativity and experiment*, *Living Rev. Rel.* **9** (2006) 3, [gr-qc/0510072].

[2] C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation.* W. H. Freeman, San Francisco, 1973.