Lesson 1: The equivalence principle and tensor analysis

Notes from Prof. Susskind video lectures publicly available on YouTube

Introduction

General relativity is the fourth volume in the collection *The Theoretical Minimum.* The first three volumes were devoted respectively to classical mechanics, quantum mechanics, and special relativity and classical field theory. The first volume laid out the lagrangian and hamiltonian description of physical phenomena, and the principle of least action which is one of the fundamental principles underlying all of physics (see volume 3, chapter 7 on fundamental principles and gauge invariance). They were used in the first three volumes and will continue in the present and subsequent ones

Physics uses extensively mathematics as its toolbox to construct formal, quantifiable, workable theories of natural phenomena. The main tools we used so far are trigonometry, vector spaces, and calculus, that is, differentiation and integration. They have been explained in volume 1 as well as in brief refresher sections in the other volumes.

We assume that the reader is familiar with these mathematical tools and with the physical ideas presented in volumes 1 and 3. The present volume 4, like its predecessors, except volume 2, belongs to classical physics in the sense that no quantum incertainty is involved.

We also began to make light use of tensors in special relativity and classical field theory. Now in general relativity we are going to use them extensively. We shall study them in detail. As the reader remembers, they generalize vectors.

Just like vectors have different representations, with different sets of numbers depending on the basis used to chart the vector space they form, this will be true of tensors as well. The same tensor will have different components in different coordinate systems. And the rules to go from one set of components to another will play a fundamental role.

Tensors were invented by Ricci-Curbastro and Levi-Civita[1] to develop work of Gauss[2] on curvature of surfaces, and Riemann[3] on non Euclidean geometry. Einstein[4] made extensive use of tensors to build his theory of general relativity. He also made important contributions to their usage: the standard notation for indices and the Einstein summation convention.

In *Savants et écrivains*, 1910, Poincaré[5] writes that "in mathematical sciences, a good notation has the same philosophical importance as a good classification in natural sciences." In this book we will take care to always use the clearest and lightest notation possible.

The equivalence principle

Einstein's revolutionary papers of 1905 on special relativity deeply clarified and extended ideas that several other physicists and mathematicians, Lorentz[6], Poincaré and others,

[1] Gregorio Ricci-Curbastro (1853-1925) and his students Tullio Levi-Civita (1873-1941) are Italian mathematicians. Their most important joint paper is "Méthodes de calcul différentiel absolu et leurs applications", in Mathematische Annalen 54 (1901), pp 125-201. They did not use the word tensor, which was introduced by other people.

[2] Carl Friedrich Gauss (1777-1855), German mathematician.

[3] Bernhard Riemann (1826-1866), German mathematician.

[4] Albert Einstein (1879-1955), German, Swiss and finally American physicist.

[5] Henri Poincaré (1854-1912), French mathematician.

[6] Hendrik Antoon Lorentz (1853-1928), Dutch physicist.

had been working on for a few years. Einstein investigated the consequences of the fact that the laws of physics, in particular the behavior of light, are the same in different inertial reference frames. He deduced from that a new explanation of the Lorentz transformations, of the relativity of time, of the equivalence of mass and energy, etc.

After 1905, Einstein began to think about extending the principle of relativity to any kind of reference frames, frames that may be in acceleration with respect to one another, not just inertial ones. Remember: an inertial frame is one where Newton's laws, relating forces and motions, have simple expressions. Or, if you prefer a more vivid image, and you know how to juggle, it is a frame of reference in which you can juggle with no problem – for instance in a railway car moving uniformly, without jerks or accelerations of any sort.

After ten years of efforts to build a theory extending the principle of relativity to frames with acceleration and taking into account gravitation in a novel way, Einstein published his work in November 1915. Unlike special relativity which topped off the work of many, general relativity is essentially the work of one man.

We shall start our study of general relativity pretty much where Einstein started.

It was a pattern in Einstein's thinking to start with a really simple elementary fact, that almost a child could understand, and deduce these incredibly far-reaching consequences. We think that it is also the best way to teach it, to start with the simplest things and deduce the consequences.

So we shall begin with the *equivalence principle.* What is the equivalence principle? It is the principle that says that *gravity is in some sense the same thing as acceleration.*

We shall explain precisely what is meant by that, and give examples of how Einstein used it. From there, we shall ask ourselves what kind of mathematical structure does a theory ought to have in order that the equivalence principle be true? What are the kinds of mathematics we must use to describe it?

Most readers have probably heard that general relativity is a theory not only about gravity, but also about geometry. So it is interesting to start at the beginning and ask what is it that led Einstein to say that gravity has something to do with geometry.

What does that mean to say that "gravity equals acceleration"? We will go through a very elementary derivation of what that means. It is not that this derivation is important, but it is worth formalizing. Formalizing means making equations that describe the world.

You all know that if you are in an accelerated frame of reference, an elevator accelerating upward or downward, you feel an effective gravitational field. Children know that because they feel it.

What follows may be overkill, but making some mathematics out of the motion of an elevator is useful to see, in a very simple example, how physicists transform a natural phenomenon into mathematics, and then to see what these mathematics really are and what in turn they can say and predict about the natural phenomenon.

Let's imagine the Einstein thought-experiment: somebody is in an elevator (figure 1). In later textbooks, the elevator got promoted to a rocket ship. But I have never been in a rocket ship, whereas I have been in an elevator. So I know what it feels like when it accelerates or decelerates. Let's say the elevator is moving upward with a velocity v.

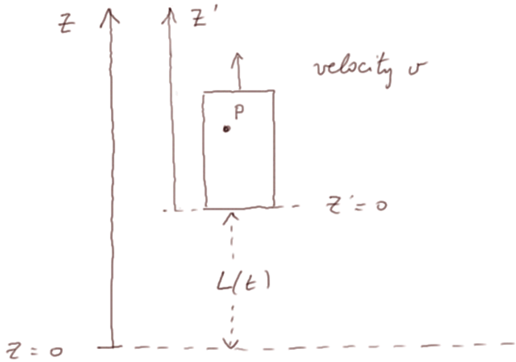

Figure 1: Elevator and two reference frames.

So far the problem is one-dimensional. We are only interested in the vertical direction. There are two reference frames: one is stationary with respect to the Earth. It uses the coordinate z. And the other is fixed with respect to the elevator. It uses the coordinate z'.

A point P anywhere along the vertical axis has two coordinates: coordinate z in the stationary frame, and coordinate z' in the elevator frame. For instance the floor of the elevator has coordinate $z' = 0$. And its z coordinate is the distance L, which is obviously a function of time. So we can write for any point P

$$z' = z - L(t) \tag{1}$$

We are going to be interested in the following question: if we know the laws of physics in the frame z, what are they in the frame z'.

One warning about this lesson: at least in the start we are going to ignore special relativity. This is tantamount to saying that we are pretending that the speed of light is infinite, or that we are talking about motions which are so slow that the speed of light can be regarded as infinitely fast

You might wonder: if general relativity is the generalization of special relativity, how did Einstein get away with starting thinking without special relativity? The answer is that special relativity has to do with very high velocities, while gravity has to do with heavy masses. There is a range of situations where gravity is important but high velocities are not. So Einstein started out thinking about gravity for slow velocities, and then combined it with special relativity to think about the combination of fast velocities and gravity. And that became the general theory.

Let's see what we know for slow velocities. And let's begin with inertial reference frames. Suppose that z' and z are both inertial reference frames. That means, among other things, they are related by uniform velocity. In other words

$$L(t) = vt \tag{2}$$

We have chosen the coordinates such that when $t = 0$, they lineup. At $t = 0$, for any point, z and z' are equal. For instance at $t = 0$ the elevator's floor has coordinate 0 in both frames. Then the floor is rising and its height z is equal to vt. So for any point we can write equation (1). Combined with equation (2), it becomes

$$z' = z - vt \tag{3}$$

Notice that this is a coordinate transformation. For readers who are familiar with Volume 3 of the collection *The Theoretical Minimum*, on special relativity, this naturally raises the question: what about time in the reference frame of the elevator? If we are going to forget special relativity, then we can just say t' and t are the same thing. We don't have to think about Lorentz transformations and their consequences. So the other half of the coordinate transformation would be $t' = t$.

We could also add to the stationary frame a coordinate x going horizontally, and a coordinate y jutting out of the page. Correspondingly coordinates x' and y' could be attached to the elevator, see figure 2. The x coordinate will play a role in a moment with a light beam. As long as the elevator is not sliding horizontally then x' and x can be taken to be equal. Same thing for y' and y.

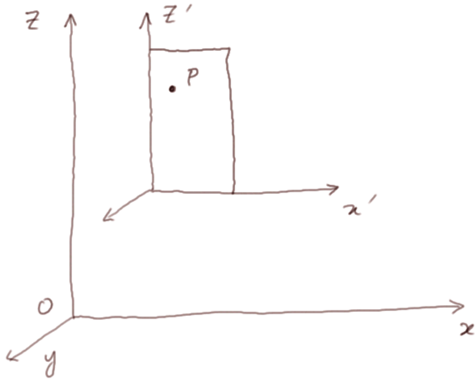

Figure 2: Elevator and two reference frames, three axes in each case.

For the sake of clarity of the drawing, we offset a bit the elevator to the right of the z axis. But think of the two vertical axes as actually sliding on each other, and at $t = 0$ the two origins O and O' as being the same. Once again, the elevator moves only vertically.

Finally our complete coordinate transformation is

$$
\begin{aligned}
z' &= z - vt \\
t' &= t \\
x' &= x \\
y' &= y
\end{aligned}
\tag{4}
$$

It is a coordinate transformation of space-time coordinates. For any point P in space-time it expresses its coordinates in the moving reference frame of the elevator as functions of its coordinates in the stationary frame. It is rather trivial. Only one coordinate, namely z, is involved in an interesting way.

Now let us look at a law of physics expressed in the stationary frame. Let's take Newton's law of motion $F = ma$ applied to an object or a particle in the elevator. The acceleration a is $\ddot{z}$, where z is the vertical coordinate of the particle. So we can write

$$F = m\ddot{z} \tag{5}$$

As we know, $\ddot{z}$ is the second time derivative of z with respect to time. It is called the vertical acceleration. Of course F is the vertical component of force. We forget about any x-component or y-component of force. They are not interesting in this context. In fact we take them equal to zero. Whatever force is exerted, it is exerted vertically.

What could this force be due to? There could be some charge in the elevator exerting a force on the charged particle. Or it could just be a force due to a rope attached to the ceiling and to the particle, that pulls on it. Any number of different kinds of forces could be acting on the particle. And we know that the equation of motion of the particle, expressed in the original frame of reference, is given by formula (5).

What is the equation of motion expressed in the prime frame? Well this is very easy. All we have to do is figure out what the original acceleration is in terms of the primed acceleration.

What is the primed acceleration? It is the second derivative with respect to time of z'. Using the first equation in equations (4)

$$z' = z - vt$$

we get first of all

$$\dot{z'} = \dot{z} - v$$

and then

$$\ddot{z'} = \ddot{z}$$

The accelerations in the two frames of reference are the same.

All of this should be familiar. But I want to formalize it to bring out some points. In particular, I want to stress that we are doing a coordinate transformation. We are asking how laws of physics change in going from one frame to another.

So now what can we say about the Newton's law in the prime frame of reference? We substitute $\ddot{z'}$ for $\ddot{z}$ in equation

(5), since they are equal, and we get

$$F = m\ddot{z}' \tag{6}$$

What do we find? We find that Newton's law in the prime frame is exactly the same as Newton's law in the unprime frame. That is not surprising. The two frames of reference are moving with uniform velocity relative to each other. If one of them is an inertial frame, the other is an inertial frame. Newton taught us that the laws of physics are the same in all inertial frames. So that is a sort of formalization of the argument.

Now let's turn to an accelerated reference frame.

Accelerated reference frames

Now $L(t)$ is increasing in an accelerated way (see fig 1). The height of the elevator's floor is then given by

$$L(t) = \frac{1}{2}gt^2 \tag{7}$$

We use the letter g for the acceleration, for obvious reasons. If you don't know what the obvious reasons are, you will find out in a minute. We know from Volume 1 of *The Theoretical Minimum*, on classical mechanics, or from high school, that this is a uniform acceleration. Indeed, if we differentiate $L(t)$ with respect to time, after one differentiation we get

$$\dot{L} = gt$$

which means that the velocity of the elevator increases linearly with time. And after a second differentiation with respect to time, we get

$$\ddot{L} = g$$

which means that the acceleration of the elevator is constant. The elevator is uniformly accelerated upward. And the equations connecting the primed and unprimed coordinates are different. For the vertical coordinates it is now

$$z' = z - \frac{1}{2}gt^2 \tag{8}$$

The other equations in (4) don't change.

$$t' = t$$

$$x' = x$$

$$y' = y$$

These four equations are our new coordinate transformation to represent the relationship between coordinates which are accelerated relative to each other.

We will continue to assume that in the z, or unprimed, coordinate system, the laws of physics are exactly what Newton taught us. In other words, the stationary reference frame is inertial. And we have $F = m\ddot{z}$. But the prime frame is no longer inertial. It is in uniform acceleration relative to the unprimed frame. Let's ask what the laws of physics are now in the prime frame of reference.

We have to do the operation of differentiating twice over again on equation (8). We know the answer:

$$\ddot{z}' = \ddot{z} - g \tag{9}$$

The primed acceleration and the unprimed acceleration differ by an amount g. Now we can write Newton's equations in the prime frame of reference. We multiply both sides of (9) by m, the particle mass, and we replace $m\ddot{z}$ by F. We get

$$m\ddot{z'} = F - mg \tag{10}$$

We arrived at what we wanted. Equation (10) looks like a Newton equation, that is, mass times acceleration is equal to some term. That term, $F - mg$, we call the force in the prime frame of reference. You notice, as expected, that the force in the prime frame of reference has an extra term, a sort of fictitious term which is the mass of the particle times the acceleration of the elevator, with a minus sign.

What is interesting about the fictitious term $-mg$, in equation (10), is that it looks exactly like the force exerted on the particle by gravity at the surface of the Earth or the surface of any kind of large massive body. It looks like a uniform gravitational field. But let me spell out in what sense it looks like gravity.

The special feature of gravity is that the gravitational forces are proportional to mass – the same mass which appears in Newton's equation of motion. We sometimes say that the gravitational mass is the same as the inertial mass. That has a deep implication. If the equation of motion is

$$F = ma \tag{11}$$

and the force itself is proportional to mass, then the mass cancels in equation (11). That is a characteristic of gravitational forces: for a small object moving in a gravitational force, its motion doesn't depend on its mass.

An example would just be the motion of the Earth about the Sun. It is independent of the mass of the Earth. If you know where the Earth is at an instant, and you know how fast it is moving, then you can predict its trajectory, without regard for what the mass of Earth is.

So equation (10) is an example of a fictitious force – if you want to call it that way – mimicking the effect of gravity. Most people before Einstein considered this largely an accident. They certainly knew that the effect of acceleration mimics the effect of gravity, but they didn't pay much attention to it. It was Einstein who said: look, this is a deep principle of nature that gravitational forces cannot be distinguished from the effect of an accelerated reference frame.

If you are in an elevator without windows and you feel that your body has some weight, you cannot say whether the elevator, with you inside, is resting on the surface of a planet, or, far away from any massive body in the universe, some impish devil is accelerating your elevator. That is the equivalence principle, that extends the relativity principle which said you can juggle in the same way at rest or in a railway car in uniform motion. In a simple example, we have equated accelerated motion and gravity. We have begun to explain what is meant by "gravity is in some sense the same thing as acceleration".

We will have to discuss this result a bit – whether we really believe it totally or it has to be qualified. But before we do that, let's draw some pictures of what these various coordinate transformations look like.

Curvilinear coordinate transformations

Let's first take the case where $L(t)$ is proportional to t. That is when we have

$$z' = z - vt$$

In figure 3, every point in space-time has a pair of coordinates z and t in the stationary frame, and also a pair of coordinates z' and t' in the elevator frame. Of course $t' = t$ and we left over the two other spatial coordinates x and y which don't change between the stationary frame and the elevator. We represented the time trajectories of fixed z with dotted lines, and of fixed z' with solid lines.

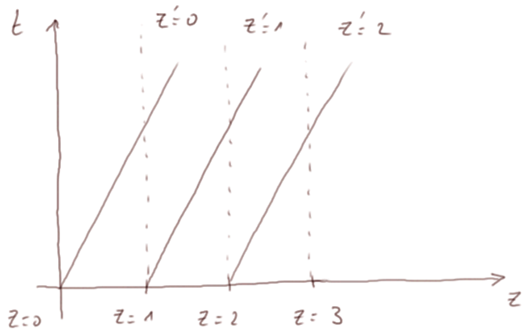

Figure 3: Linear coordinates transformation.

That is called a *linear coordinates transformation* between the two frames of reference. Straight lines go to straight lines, not surprisingly since Newton tells us that free particles move in straight lines in an inertial frame of reference. And what is a straight line in one frame had therefore better be a straight line in the other frame. Not only do free particles move in straight lines in space, but their trajectories are straight line in space-time – they don't change speed on their spatial straight lines.

Now let's do the same thing for the accelerated coordinate system. The transformation equations are now equation (8) linking z' and z, which we repeat here

$$z' = z - \frac{1}{2}gt^2$$

and the other coordinates which don't change.

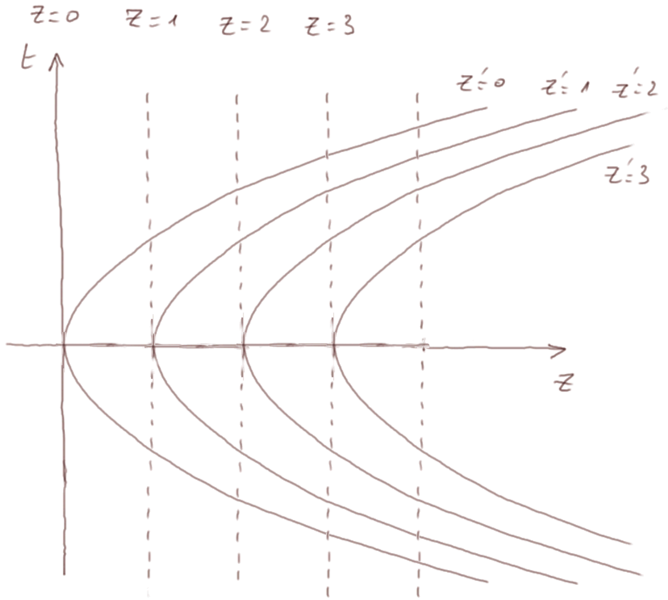

Figure 4: Curvilinear coordinates transformation.

Again, in figure 4, every point in space-time has two pairs of coordinates $(z,\ t)$ and $(z',\ t')$. The time trajectories of fixed z, represented with dotted lines, don't change. But now the time trajectories of fixed z' are parabolas lying on their side.

We can even represent negative times in the past. Think of the elevator that was initially going down with acceleration g. This creates a decelaration up to, say, $t = 0$ and $z = 0$. And then the elevator bounces back upward with the same acceleration g. Each parabola is just shifted relative to the previous one by one unit to the right.

The point of this figure, of course, is, not surprisingly, that straight lines in one frame are not straight lines in the other frame. They become curved lines. The lines of fixed t or fixed t' of course are the same horizontal straight lines in both frames. We haven't represented them.

We should view figure 4 as just two sets of coordinates to locate each point in space-time. On set of coordinates has straight axes, while the second – represented in the first frame – is curvilinear. Its $z' = constant$ lines are actually curves, while its $t' = constant$ lines are horizontal straight lines. So it is a *curvilinear coordinates transformation.*

Something Einstein understood very early is this:

There is a connection between gravity and curvilinear coordinates transformations of space-time.

Special relativity was only about linear transformations, transformations which take uniform velocity to uniform velocity. Lorentz transformations are of that nature. They take straight lines in space-time to straight lines in space-time.

But if we want to mock-up gravitational fields with the effect of acceleration, we are really talking about transformations of the coordinates of space-time which are curvilinear.

That sounds extremely trivial. When Einstein said it, probably every physicist knew it and thought: oh yeah, no big deal. But Einstein was very clever and very persistent. He realized that if he dug very deep into the consequences of this, he could then answer questions that nobody knew how to answer.

Let's look at a simple example of a question that Einstein answered using the curvature of space-time created by acceleration and therefore – if the two are the same – by gravity. The question is: what is the influence of gravity on light?

Effect of gravity on light

After his work of 1905 on special relativity, Einstein began to think about gravity. When he first asked himself the question "what is the influence of gravity on light?", around 1907, most physicists would have answered: there is no effect of gravity on light. Light is light. Gravity is gravity. A light wave moving near a massive object creating a gravitational field moves in a straight line. It is a law of light that it moves in straight lines. And there is no reason to think that gravity has any effect on it.

But Einstein said: no, if this equivalence principle between acceleration and gravity is true, then acceleration and gravity will affect light in the same way. And the argument was very simple. It was again one of these arguments which you could almost explain to a child.

Let's start out with a light beam that is moving horizontally with respect to the stationary reference frame, starting out at $t = 0$ (figure 5). And let's see what is the motion of the light beam expressed in the elevator reference frame.

At $t = 0$ the two frames of reference match: $z' = z$ for every point in space-time, see equation (8). And they are both at rest relative to each other, that is, at $t = 0$ the velocity of the elevator is also zero because for any point of the elevator we have $\dot{z} = qt$. But it begins to move up.

A flashlight sends a beam of light in the x direction, from point $(x = 0,\ z = 0,\ t = 0)$ in the stationary frame. This point in space-time has also the coordinates $(x' = 0,\ z' = 0,\ t' = 0)$ in the elevator frame.

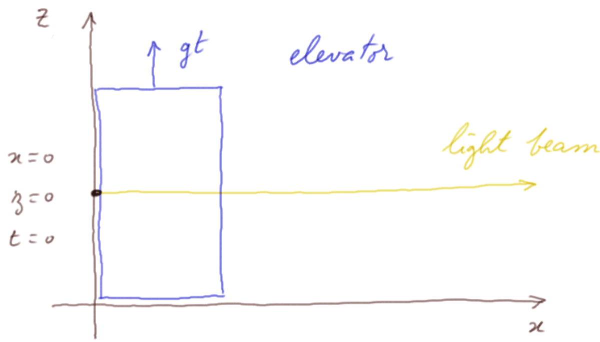

Figure 5: Trajectory of a light beam in the *stationary* reference frame.

In the stationary frame of reference we know how the light beam moves. Einstein said: in the frame of reference without any gravity the light beam goes in a straight line. Of course there is gravity on the surface of the Earth, but let's neglect it – or transport all our experiment in outer space in an inertial frame away from any gravitational field. The only apparent gravity, in the elevator, will be due to its acceleration.

The equations of motion for the light beam – think of it, for instance, as one photon – are, in the stationary frame of reference,

$$\begin{aligned} x &= ct \\ z &= 0 \end{aligned} \tag{12}$$

Now, let's replace the unprimed coordinates of the photon (that is, the coordinates in the stationary frame) by its primed coordinates (that is, the coordinates in the elevator frame). Equations (12) become

$$\begin{aligned} x' &= ct \\ z' + \frac{qt^2}{2} &= 0 \end{aligned} \tag{13}$$

In the elevator, the photon trajectory is shown in figure 6. It is a parabola, whose equation linking z' and x' we can obtain by eliminating t in equations (13).

If I am in the elevator, I say: gravity is pulling the light beam down. You say: no, it's just that the elevator is accelerating upward. That makes it look like the light beam moves on a curved trajectory. And Einstein said they are the same thing.

This proved to him that a gravitational field must bend a light ray. And that was something that no other physicist knew at the time.

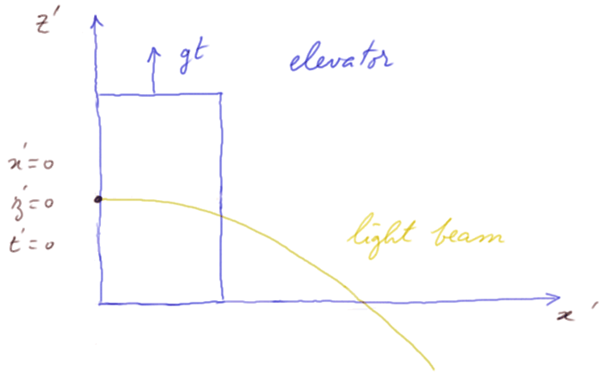

Figure 6: Trajectory of a light beam in the *elevator* reference frame.

Till now what we have learned – the thing we want to abstract – is that it is interesting to think about curvilinear coordinates transformations.

When we do think about curvilinear coordinates transformations, the form of Newton's laws do change. One of the things that happen is that apparent gravitational fields materialize, that are physically indistinguishable from ordinary gravitational fields.

Well are they really physically indistinguishable? Not really. So let's talk about real gravitational fields, namely gravitational field of gravitating objects like the Sun or the Earth.

Tidal forces

In figure 7 is represented the Earth or the Sun. And the gravitational acceleration doesn't point vertically on the page. It points toward the center.

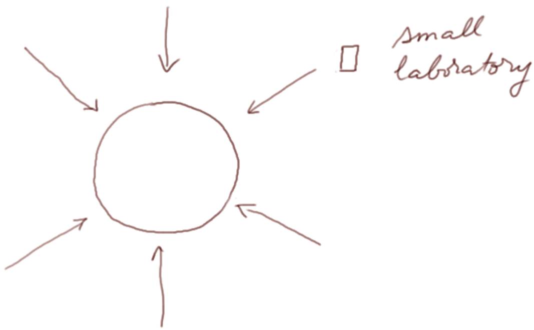

Figure 7: Gravitational field of a massive object, for instance the Earth

It is pretty obvious that there is no way that you could do a coordinate transformation like we did in the preceding section which will remove the effect of the gravitational field.

OK, if you are in a small laboratory in space and that laboratory is allowed to simply fall toward the Earth, or whatever massive object you are considering, then you will think that in that laboratory there is no gravitational field.

Exercise 1 : If we are falling freely in a uniform gravitational field, prove that we feel no gravity and that things float around us like in the International Space Station.

But there is no way globally to introduce a coordinate transformation which is going to get rid of the fact that there is a gravitational field pointing inward toward the center. For instance a very simple transformation similar to equations (12) might get rid of the gravity in a small portion on one side of the Earth, but the same transformation will increase the gravitational field on the other side. And even more complex transformations will not solve the problem.

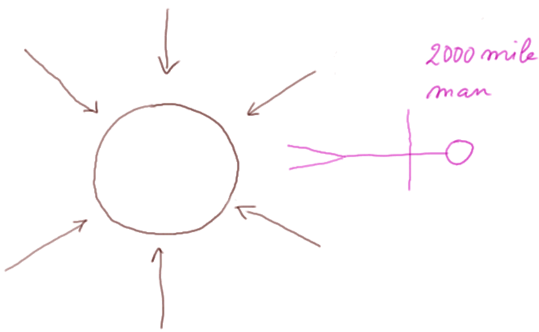

Figure 8: 2000-mile-man falling toward the Earth

One way to understand why we can't get rid of gravity is to think of an object which is not very very small compared to the gravitational field. My favorite example is a 2000-mile-man who is falling in the Earth gravitational field, figure 8.

Because he is large, not a point mass, the different parts of him feel different gravitational fields. The further away you are the weaker the gravitational field is. So his head feels a weaker gravitational field than his feet. His feet are being pulled harder than his head. He feels like he is being stretched. If he is freely falling he can't distinguish this from a stretching feeling.

But he knows that there is a gravitating object there. The sense of discomfort that he feels, due to the non-uniform gravitational field, cannot be removed by going to a falling reference frame. And indeed no change of mathematical description whatsoever has ever changed a physical phenomenon.

The forces he feels are called tidal forces. They cannot be removed by a coordinate transformation.

Let's also see what happens if he is falling not vertically but sideways? In that case his head and his feet will be at the same distance of the Earth. Both will be subjected the same force in magnitude pointing to the Earth. But since they are radial, they are not parallel. A part of component of the force pulling his head will be along his body, and a part of the force pulling his feet too. So they will also tend to compress or shrink him.

That sense of compression is again not something that you can remove by any coordinate transformation. Being stretched or shrunk or both by the Earth field – if you are big enough – is an invariant fact. It is not something you can get rid of by doing a coordinate transformation.

So it is not true that gravity is equivalent to going to an accelerated reference frame. Einstein of course knew this. What he really meant is:

Small objects for a small length of time cannot tell the difference between a gravitational field and an accelerated frame of reference.

That raises the question: if I present you with with a force field, does there exist a coordinate transformation which will make it disappear?

For example the force field, inside the elevator, associated with its uniform acceleration with respect to an inertial reference frame, was just a vertical force field pointing downward and uniform everywhere. There was a transformation making it disappear: simply to use the z coordinates instead of the z'. It is a nonlinear coordinate transformation, but nevertheless it gets rid of the force field.

With other kinds of coordinate transformations – from inertial to moving –, you can make the gravitational field look more complicated, for example transformations which affect also the x-coordinate. They can make the gravitational field bend toward the x-axis.

You might simultaneously accelerate along the z-axis while oscillating back and forth in the x-axis. What kind of gravitational field do you see? A very complicated one: it's got a vertical component and it has got a time dependent oscillating component along the x-axis.

If instead of the elevator you use a merry-go-round, and instead of the $(x',\ z',\ t)$ coordinates of the elevator, you use polar coordinates $(r,\ \theta,\ t)$, an objet that in the initial frame was fixed, or had a simple motion like the light beam, may have a weird motion. You may think that you have discovered some repulsive gravitational field phenomenon. But no matter what, the reverse coordinate change will reveal that your apparently messy field is only the consequence of a coordinate change.

By appropriate coordinate transformations you can make some pretty complicated apparent gravitational fields. They noneless are not genuine, in the sense that they don't result from the presence of a massive object.

I tell you what the gravitational field is everywhere, how do you determine whether it is just the sort of fake gravitational field coming from a coordinate transformation to a frame with various kinds of accelerations with respect to an inertial one, or it is a real genuine gravitational field?

Well, if we are talking about Newtonian gravity – not a field for instance coming from electromagnetism – there is an easy way. You just calculate the tidal forces. You determine whether that gravitational field will have an effect on an object which will cause it to squeeze and stretch.

If calculations are not pratical, you take an object, a mass, a crystal. You let it fall and you see whether there was stresses and strains on it. These will be detectable things. If you detect stresses and strains then it is a real gravitational field. If you discover on the other hand that the gravitational field has no such effect, that any object, wherever it is located and let free to move, experiences no tidal force, in other words that the field has no tendency to distort a freely falling system, then it is not a real gravitational field.

The test object must be let free. It should not be held in place by ropes and pulleys or anything else, be it in the elevator or on the merry-go-round.

In summary: how do we know when it is possible to make a coordinate transformation that removes the gravitational field? How do we know when the field is just an artifact of the coordinate system, and when it is a real physical thing, due to some gravitating mass for example?

If the gravitational field is small enough that it doesn't vary between one part of the system and another, that is when the equivalence principle holds with accuracy. Einstein knew all that. So he asked himself the question: what kind of mathematics goes into trying to answer the question of whether a field is a genuine gravitational one or not.

Non-Euclidean geometry

Atfer his work on special relativity, and after learning of the mathematical structure in which Minkowski[7] had recast it, Einstein knew that special relativity had a geometry associated with it.

So let's take a brief rest from gravity to remind ourselves of this important idea in special relativity. The only thing we are going to use about special relativity is that space-time has a geometry.

In the Minkowski geometry of special relativity, there exists a kind of length between two points, that is, between two events in space-time.

[7] Hermann Minkowski (1864-1909), Polish-German mathematician and theoretical physicist.

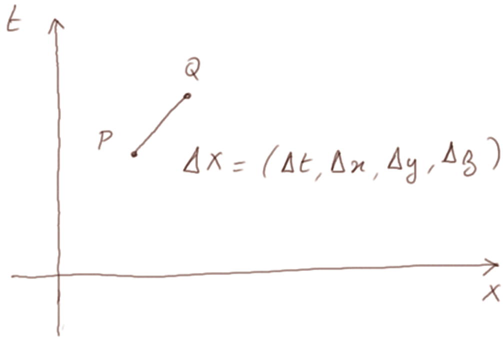

Figure 9: Minkowski geometry.

The distance between P and Q is not the usual Euclidean distance that we could be tempted to think of.

The distance, or length, between any pair of points P and Q, close to each other or not, is defined as follows. Let's call ΔX the 4-vector going from P to Q. To the pair $P\,Q$ is associated the quantity denoted $\Delta\tau$, defined by

$$\Delta\tau^2 = \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$$

$\Delta\tau$ is called the *proper time* between P and Q. It is an invariant under Lorentz transformations. That is why it qualifies as a kind of distance, just like in Euclidean 3D space the distance of a point from the origin, $x^2 + y^2 + z^2$, is invariant under rotations.

Equivalently we may define Δs by

$$\Delta s^2 = -\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$$

Δs is called the *proper distance* between P and Q.

$\Delta\tau$ and Δs are not two different concepts. They are the same – just differing by an imaginary factor i. They are just two ways to talk about the Minkowski distance between P and Q. Depending on which physicist is writing the equations, he or she will rather use $\Delta\tau$ or Δs as the distance between P and Q.

Einstein knew about this non-Euclidean geometry of special relativity. In his work to include now gravity, and to investigate the consequences of the equivalence principle, he also realized that the question we asked at the end of the previous section – are there coordinate transformations which can remove the effect of forces? – was very similar to a certain mathematics problem that had been studied at great length by Riemann. It is the question of deciding whether a geometry is flat or not.

Riemannian geometry

What is a flat geometry?

Intuitively, it is the following idea: the geometry of a page is flat, the geometry of the surface of a sphere or a section of a sphere is not flat.

And the *intrinsic geometry* of the page remains flat even if we furl the page like in figure 10. We shall expound mathematically on the idea in a moment. For now, let's just say that the intrinsic geometry of a surface is the geometry that a bug roaming on it, equipped with tiny surveying tools, would see if it were trying to establish an ordnance survey map of the surface. If it worked carefully it might see hills and valleys, bumps and troughs, if there were any, but it would not determine, as we see it, that the page is furled. We see it because *for us* the page is embedded in the 3D Euclidean space we live in. And by unfurling the page, we can make its flatness obvious again.

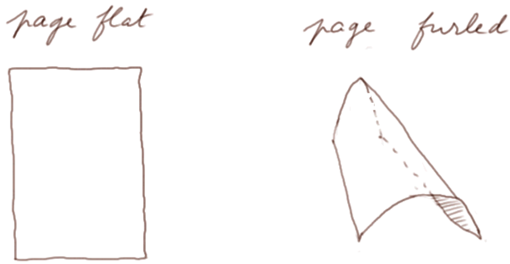

Figure 10: The intrinsic geometry of a page remains flat.

Einstein realized that there was a great deal of similarity in the two questions of whether a geometry is non-flat and whether a space-time has a real gravitational field in it.

Riemann had studied the first question. But Riemann had never dreamt about geometries which have a minus sign in the definition of the square of the distance. He was thinking about geometries which were non-Euclidean but were similar to Euclidean geometry.

Minkowski geometry is non-Euclidean in two respects. The square of the distance between two points is not just $\Delta x^2 + \Delta y^2 + \Delta z^2$. But furthermore the square of the distance can be negative!

So we want to start with the mathematics of Riemannian geometry, that is of spaces where the distance between two points is not the Euclidean distance, but its square is always positive[8].

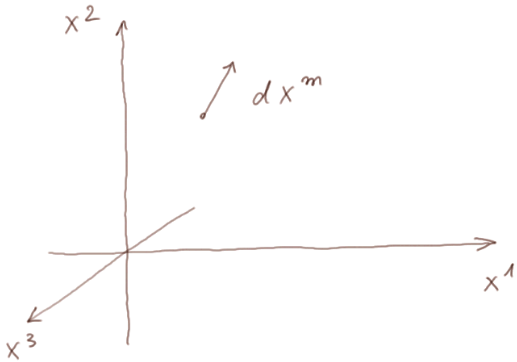

Figure 11: Distance between two points in a space.

We are not going to spend a huge amount of time on it, but we shall nevertheless study a little bit the geometry of Riemannian surfaces.

We look at two points in a space (figure 11). In our picture there are three dimensions, three axes, X^1, X^2 and X^3. There could be more. So a point has three components, that we can write X^m, where m is understood to run from one to three or to whatever number of axes there is. And a little shift between two points can be denoted ΔX^m or, if it is to become an infinitesimal, dX^m.

[8]In mathematics, they are called positive definite distances.

If this space has the usual Euclidean geometry, the square of the length of dX^m is given by Pythagoras theorem

$$dS^2 = (dX^1)^2 + (dX^2)^2 + (dX^3)^2 + ... \tag{14}$$

If we are in three-dimensions then there are three terms in the sum. If we are in two dimensions, there are two of them. If the space is 26 dimensional, there are 26 of them and so forth. That is the formula for Euclidean distance between two points in the Euclidean space.

To pursue our study of Riemannian geometry, it is now best to think of a two-dimensional space. It can be the ordinary plane, or it can be a two-dimensional surface that we may visualize embedded in 3D Euclidean space with clearly defined unit distance, because it enables us to conveniently view its curvature. See figure 12.

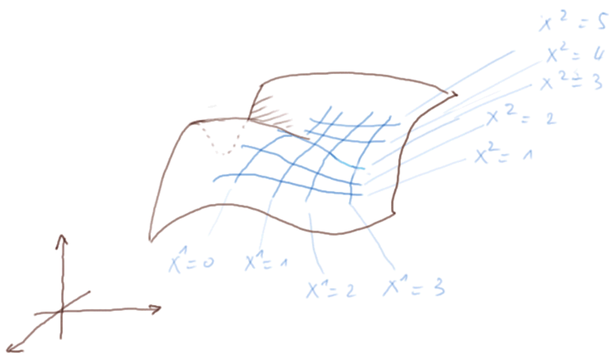

Figure 12: Two dimensional variety.

There is nothing special about two dimensions for such a surface, except that it is easy to visualize. Mathematicians think of "surfaces" even when those have more dimensions. And usually they don't call them surfaces but *varieties*.

Gauss had already understood that on curved surfaces the formula for the distance between two points was more complicated in general than equation (14).

Indeed we must not be confused by the fact that in figure 12, the surface is shown embedded in the usual three-dimensional Euclidean space. This is just for convenience of representation.

We ought to think of the surface, or variety, as a space in itself, equipped with a coordinate system to locate any point, with curvy lines corresponding to one coordinate constant, etc. and where a distance has been defined. We must forget about the embedding 3D Euclidean space.

The distance between two points on the surface is certainly not their distance in the embedding 3D Euclidean space, and is not even necessarily defined on the surface with the equivalent of equation (14).

Riemann generalized these surfaces and their metric (the way to compute distances) to any dimensions. But let's continue to use our picture with two dimensions in order to sustain intuition. And let's go slowly, so as not to miss any important detail.

So the first thing we do with a surface is to put some coordinates on it. Indeed we want to be able to quantify various statements involving its points, therefore we need to introduce some coordinates onto it.

We just lay out some kind of coordinates as if drawing them with a chalk. We don't worry at all about whether the coordinate axes are straight lines or not, because for all we know when the surface is a really curved surface there probably won't even be things that we can call straight lines. So we just lay out some coordinates and we still call them X's.

The values of the X's are not related directly to distances. They are just numerical labels. The points ($X^1 = 0,\ X^2 = 0$) and ($X^1 = 1,\ X^2 = 0$) are not necessarily separated by a distance of one. Now we take two neighboring points (figure 13).

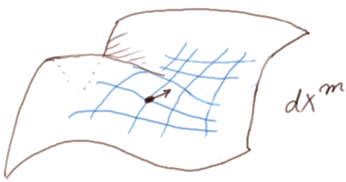

Figure 13: Two neighboring points,
and the shift dX^m they form.

The two neighboring points are again related by a shift of coordinates. But, unlike in figure 11 which was still a Euclidean space, now we are on an arbitrary curved surface with arbitrary coordinates.

Now we *define* a distance on the surface for points separated by a small shift like dX^m. It won't be as simple as equation (14). It will have similarities with it. Here is the new definition of dS^2

$$dS^2 = \sum_{m,\ n} g_{mn}(X)\ dX^m dX^n \tag{15}$$

This is the formula in any geometry curved or otherwise.

Incidentally equation (15) applies even to flat geometries equipped with curvilinear coordinates. Suppose you take a flat geometry, like the surface of the page, but you use some curvilinear coordinates to locate points. And you ask what is the distance between two points close to each other. Then in general the square of the distance between two points close to each other will be a quadratic form in the coordinate shifts dX^m's. A quadratic form means a sum of terms, each of which is either the product of two little shifts, or the square of one little shift, times a coefficient like g_{mn} which depends on X.

Let's consider a simple example of distance on a curved surface. It is the distance on the surface of the Earth between two points nearby characterized by longitude and latitude. Let's denote R the radius of the Earth.

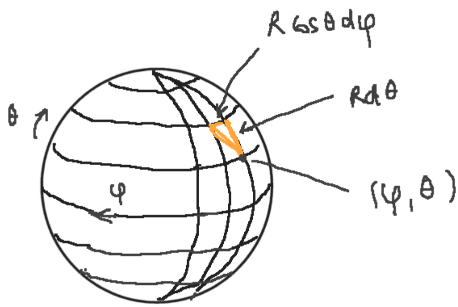

Figure 14: Distance on the surface of the Earth.

We take two points $(\phi,\ \theta)$, and $(\phi + d\phi,\ \theta + d\theta)$, where θ is the latitude and ϕ the longitude.

We apply Pythagoras theorem in a small approximately flat rectangular region to compute the square of the length of its diagonal. One side along a meridian has length $Rd\theta$. The other side along a parallel has a length $Rd\phi$ but corrected by the cosine of the latitude. At the equator it is the full $Rd\phi$, but at the pole it is zero. So the formula is

$$dS^2 = R^2 \left[d\theta^2 + (cos\theta)^2 d\phi^2 \right] \tag{16}$$

Here is an example of squared distance not just being $d\theta^2 + d\phi^2$ but having some coefficient functions in front. In this case the interesting coefficient function is $(cos\theta)^2$ in front of one of the terms $(dX^m)^2$. We also note that there are no terms of the form $dX^m dX^n$ because the natural curvilinear coordinates we chose on the sphere are still orthogonal at every point.

In other examples – on the sphere with more involved coordinates, or on a more general curved surface like in figure 13 – where the coordinates are not necessarily locally perpendicular the formula for dS^2 would be more complicated and have terms in $dX^m dX^n$. But it will still be a quadratic form. There will never be $d\theta^3$ terms. There will never be things linear, everything will be quadratic.

You may wonder why we define the distance dS only for small – actually infinitesimal – displacements. The reason is that to talk about distance between two points A and B far away from each other, we must first of all define what we mean. There may be bumps and troughs in between. Well, we could mean the shortest distance as follows: we would put a peg at A, a peg at B and pull a string as tight as we can between the two points. That would define one notion of distance. Of course there might be several paths with the same value. One might go around the hill one way.

And then the other would go around the hill the other way. Simply think on the Earth of going from the North pole to the South pole.

Furthermore even if there is only one answer, we have to know the geometry on the surface everywhere in the whole region where A and B are, not only to calculate the distance but to know actually where to place the string. So the notion of distance between any two points is a complicated one.

But the notion of distance between two neighboring points is not so complicated. That is because locally a smooth surface can be approximated by the tangent plane and the curvilinear coordinates lines by straight lines – not necessarily perpendicular but straight.

Metric tensor

Now we shall go deeper into the geometry of a curved surface and its links with equation (15) which defines the distance between two neighboring points on it. Recall the equation

$$dS^2 = \sum_{m,\ n} g_{mn}(X)\ dX^m dX^n$$

In order to get a feel about the geometry of the surface and its behavior, let's imagine that we arrange elements from a Tinkertoy Construction set along the curved surface. For instance they could approximately follow the coordinate lines on the surface. And we would also add more rigid elements diagonally. This would create a lattice as shown in figure 15. But any reasonably dense lattice, sort of triangulating the surface, would do as well.

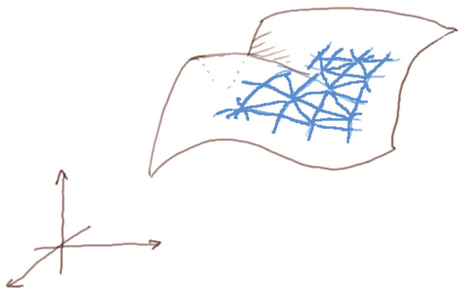

Figure 15: Lattice of rigid Tinkertoy elements arranged on the surface.

Suppose furthermore that the Tinkertoy elements are hinged together in a way that let them freely move in any direction from each other.

Now let's lift our lattice from the surface. Sometimes it will keep its shape rigidly, sometimes it won't. It will not keep its shape if going from the initial shape to a new shape doesn't force any Tinkertoy element to be stretched or compressed or bend.

And in some cases it will even be possible *to lay it out flat*. It is the case, for instance, in figure 10 going from the shape on the right to the shape on the left – which is just a flat page.

Exercise 2 : Is it possible to find a curved surface and a lattice of rods arranged on it, which cannot be flattened out, but which can change shape?

We shall see that the initial surface being able to take other shapes or not corresponds to the g_{mn}'s of equation (15) having certain mathematical properties.

The collection of g_{mn}'s has a name. It is called the *metric tensor*. It is the mathematical object that enables us to compute the local distance between two neighboring points on our Riemannian surface.

When the lattice of Tinkertoy elements can be laid out flat, the geometry of the surface is said to be flat. We will define it more carefully later.

Sometimes on the other hand the lattice of little rods cannot be laid out flat. For example on the sphere, if we initially lay out a lattice triangulating a large chunk of sphere, we won't be able to remove it to lay it out on a flat plane[9].

The question now we have to address is this: if I had made a lattice of little rods covering a surface, and I gave you the length of each rod, without yourself building the lattice how could you tell me whether it is a flat space or it is an intrinsically curved space which cannot be flattened out

[9]This is a well-known problem of cartographers, which lead to the invention of various kinds of maps of the world, the most famous being the Mercator projection map invented by Flemish cartographer Gerardus Mercator (1512-1594).

and laid out on a flat plane?

Let's formulate it more precisely and mathematically: we start from the metric tensor $g_{mn}(X)$ which is a function of position, in some set of coordinates. Keep in mind that there are many different possible sets of curvilinear coordinates on the surface, and in every set of coordinates the metric tensor may look different. It will have different components, just like the same 3-vector in ordinary 3D Euclidean space has different components depending on the basis used to represent it.

So I select one set of coordinates and I give you the metric tensor of my surface. In effect I tell you the distance between every pair of neighboring points. Now the question is: is my surface flat or is it not flat?

Of course we have to explain what we mean by flat.

You may think of "checking Pi". Here is the way it would go – think of a 2D surface embedded in the usual 3D Euclidean space as shown in figure 12. You select a point and mark out a disk around it. Then you measure its radius r as well as its circonference l. And you divide l by $2r$. If you get 3.14159... you would say that it is flat. Otherwise you would say that it is not flat, it is curved.

That procedure is good for a two-dimensional surface under certain conditions. But anyway it is not so great for higher dimensional surfaces.

So what is the mathematics of taking a metric tensor and asking if its space is flat. What does it mean for it to be flat? By definition

The space is flat if you can find a coordinate transformation, that is, a different set of coordinates, in which the distance formula for dS^2 is just $(dX^1)^2 + (dX^2)^2 + ...(dX^n)^2$, as it would be in Euclidean geometry.

That the initial g_{mn}'s form the unit matrix, with ones on the diagonal and zeroes elsewhere – as if equation (15) were just Pythagoras theorem – is not necessary. But we must find a coordinate transformation which can make it look like that.

In that sense, it has a vague similarity – it turned out not to be a vague similarity at all but a close parallel – with the question of whether you can find a coordinate transformation which removes gravitational field. The question is: can you find a coordinate transformation which removes the curvy character of the metric tensor g_{mn}?

In order to answer that geometric question, we have to do some mathematics essential to relativity. It is not possible to understand general relativity without it. The mathematics is tensor analysis and a little bit of differential geometry.

At first it looks annoying because we shall have to deal with all these indices floating around, and different coordinate systems, and partial derivatives of components, etc. But once you get used to it, it is really very simple. It was invented by Ricci-Curbastro and Levi-Civita at the end of the XIXth century to build on works of Gauss and Riemann. And it was actually made even simpler by the famous Einstein summation convention which astutely gets rid of most summation symbols.

To find a set of coordinates that make equation (15) become equation (14) is a more involved procedure than just diagonalizing the matrix g_{mn}. The reason is that there is not one matrix. g_{mn} depends on X. It is the same tensor field, but it has a different matrix at each point[10]. You cannot diagonalize all of them at the same time. At a given point, you can indeed diagonalize $g_{mn}(X)$ even if the surface is not flat. It is equivalent to working locally in the tangent plane of the surface at X, and orthogonalizing out the coordinate axes there. But you cannot say that a surface is flat because it can be made at any given point locally to look like the Euclidean plane.

So we are looking for a coordinate transformation, $X \to Y$, which turns g_{mn} into the Kronecker delta function δ_{mn}. Remember, X and Y represent the *same point* on the surface. They are its coordinates in different coordinate systems. And the Kronecker delta function is

$$\begin{aligned} \delta_{mn} &= 1, \text{ if } m = n \\ \delta_{mn} &= 0, \text{ if } m \neq n \end{aligned}$$

We want the square of the interval distance to become

$$dS^2 = \sum_{m,\ n} \delta_{mn}(Y)\ dY^m dY^n \tag{17}$$

You can also think of δ_{mn} as the unit matrix, with ones everywhere on the diagonal, and zeroes everywhere off the

[10]For a given set of coordinates, it has a collection of matrices – one for each point. And for another set of coordinates, it will have another collection of matrices. And we still talk of the same tensor field. Its components depend on the coordinates. But the tensor itself is an abstract object which doesn't. We already met the distinction with 3D vectors.

diagonal.

In summary: I give you the metric tensor of my surface, that is, the g_{mn} of equation (15), which is

$$dS^2 = \sum_{m,\ n} g_{mn}(X)\ dX^m dX^n$$

and I ask you if by a coordinate transformation $X \to Y$ you can reduce it to equation (17).

If yes, the space is called flat. If no, the space is called curved. Of course the space could have some portions which are flat. There could exist a set of coordinates such that in a region the metric tensor be the Kronecker delta. But it is called flat if it is everywhere flat.

This is now a mathematics problem: given a tensor field $g_{mn}(X)$ on a multi-dimensional variety, how do we figure out if there is a coordinate transformation which would change it into the Kronecker delta function?

To answer that question, we have to understand better how things transform when you make coordinate transformations. And that is the subject of tensor analysis.

The analogy between tidal forces and curvature actually is not an analogy, it is a very precise equivalence. In the general theory of relativity, the way you diagnose tidal forces is by calculating the curvature tensor. A flat space is defined as a space where the curvature tensor is zero everywhere. Therefore it is a very precise correspondence:

Gravity is curvature.

But we will come to it as we get through tensor analysis.

Obviously, in trying to determine whether we can transform away $g_{mn}(X)$ and turn it into the trivial $\delta_{mn}(Y)$, the first question to ask is how does $g_{mn}(X)$ transform when we change coordinates?

We have to introduce notions of tensor analysis which are rather easy.

First tensor rule: contravariant vectors

Sometimes tensor notations are a bit of a nuisance because of all the indices. At first we can get confused by them. But soon you will discover that the manipulations obey strict rules and turn out to be rather simple.

We shall begin with a simpler thing than $g_{mn}(X)$. Suppose there are two sets of coordinates on our surface. There a set of coordinates X^m. And I could call the other coordinates X' as we did earlier. But then we would be running into horrible notation like X^{11} and X'^{11}. So the second set of coordinates, we call Y^m.

To be very explicit, a same point P on the surface has coordinates $\left[\ X^1(P),\ X^2(P),\ X^3(P),\ ...\ , X^N(P)\ \right]$ if we are on variety of dimension N. And it also has coordinates $\left[\ Y^1(P),\ Y^2(P),\ Y^3(P),\ ...\ , Y^N(P)\ \right]$.

But we will almost never recall the variable P. So we have at every point a collection of X^m coordinates and a collection of Y^m coordinates.

The X's and Y's are related because if you know the coordinates of a point P in one set of coordinates then in principle you know where the point is. Therefore you know its coordinates in the other coordinate system. So each coordinate X^m is a function of all the coordinates Y^n. We can use whatever dummy index variable we want if that helps avoid confusion. In fact we write simply

$$X^m(Y)$$

Likewise each Y^m is assumed to be a known function of all the X^n's.

$$Y^m(X)$$

We have two coordinate systems, each one being a function of the other.

Now we ask: how the differential elements dX^m transform? The collection of differential elements dX^m is a small vector as shown in figure 16. Remember, the vector itself is a pair of points (an origin and an end). It is independent of the coordinate system. But in order to work with it, most of the time we must look at it expressed with components. And those dX^m are the components of the little displacement under consideration, in the X coordinate system.

Figure 16: Displacement dX^m.

The notation dX^m is used to represent the small vector

$$dX^m = [\ dX^1,\ dX^2,\ dX^3,\ \dots\ , dX^N\ ]$$

Said another way, when we change a little bit X^1, and change a little bit X^2, and change a little bit X^3 etc. the point X moves to a nearby point, and the displacement is dX^m.

Now we look at how dY^m, for any given m, can be expressed in terms of the dX^p's. It is an elementary result of calculus that

$$dY^m = \sum_p \frac{\partial Y^m}{\partial X^p} \, dX^p \tag{18}$$

Remember elementary calculus: if a function f depends on two variables a and b we write

$$f(a,\ b)$$

Then the total differential of f when a and b each moves a little bit is given by the formula

$$df = \frac{\partial f}{\partial a} da + \frac{\partial f}{\partial b} db$$

Think of a function f as simple as the area of a rectangle of length a and width b: $f(a,\ b) = ab$.

Equation (18), reproduced here

$$dY^m = \sum_p \frac{\partial Y^m}{\partial X^p} \, dX^p$$

says nothing more than figure 17 in a more general setting.

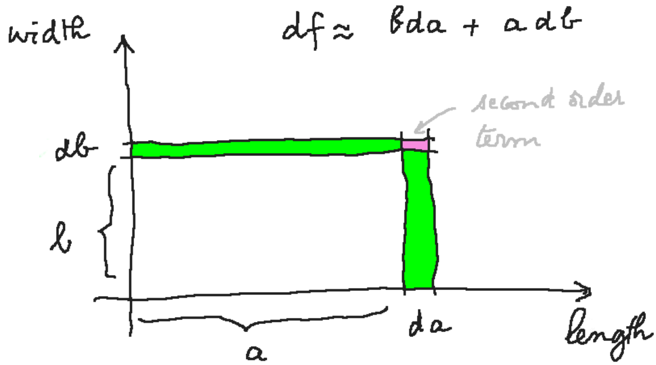

Figure 17: $df = \frac{\partial f}{\partial a} da + \frac{\partial f}{\partial b} db$

it is the green area.

Let's spell out even more explicitly what equation (18) says: the total change of some particular component Y^m is the sum of the rate of change of Y^m when you change only X^1, times the little change in X^1, namely dX^1, plus the rate of change of Y^m when you change only X^2, times the little change in X^2, namely dX^2, and so forth.

Now right here we have the *first example of the transformation of a tensor.*

We now have the expression of the small displacement of a point on the surface (figure 16), expressed in two different coordinate systems, dX^m and dY^m. dX^m and dY^m are two sets of components for the *same* displacement. And we know how to go from one set to the other. Let's represent again the small displacement, and also locally the two sets of coordinates, figure 18:

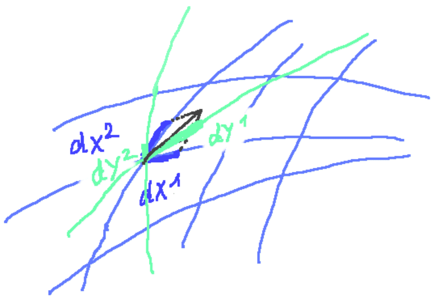

Figure 18: Small displacement, and two sets of coordinates.

Equation (18) is the transformation property of the components of the vector. Now usual vectors, we shall call *contravariant vectors*. So equation (18), which is really a set of equations, is the transformation properties of the components of what is called a contravariant vector.

If we leave for a moment geometric representations or visualizations, and think in more abstract mathematical terms, a contravariant vector simply means a thing which transforms as in equation (18).

Let's give the general definition of a *contravariant vector*. It is an object, a "thing", which has components

$$V^m$$

These components depend on the coordinate system. V^m is the collection of components of the object in the X frame. Let's denote $(V')^m$ the components of the *same* object in the Y frame. For the object to be a contravariant vector its two sets of components must satisfy

$$(V')^m = \sum_p \frac{\partial Y^m}{\partial X^p} V^p \tag{19}$$

Equation (19) is the fundamental equation of a contravariant vector. We will soon see other types of objects called covariant vectors.

Remember, the *ordinary vectors* of $\mathbb{R}^3$, or $\mathbb{R}^n$, or of any vector space *are contravariant vectors.* The position of a point is a contravariant vector. The speed of a particle is a contravariant vector. In tensor analysis, they are called contravariant because they change *contrary* to the change of basis. For instance if your new basis vectors are simply the old basis vectors, all divided by 10, then V' will simply be V multiplied by 10.

There are many things which transform according to equation (19) and therefore are contravariant vectors.

Einstein summation convention

One of Einstein's great contributions to notations was that you can leave out the summation sign in equation (19):

$$(V')^m = \sum_p \frac{\partial Y^m}{\partial X^p} V^p$$

can be rewritten

$$(V')^m = \frac{\partial Y^m}{\partial X^p} V^p \tag{20}$$

People will figure out that there is an implicit summation just by looking at the equation. They will see that the left hand side of equation (20) has no index p, while the right hand side has an index p, so it must be summed over.

Whenever there is a repeated index like p, that seems not to make sense, because it appears on the right side of the equation, but disappears from the left side, it really stands for summation over that index.

Second tensor rule: covariant vectors

Now let's talk about covariant vectors. They are completely different creatures. The most typical covariant vectors are gradients.

Suppose we have a scalar function defined on our space, surface, variety – whatever you call it. That is, if we want to use the terminology introduced in Volume 3, we consider a scalar field. A scalar function is a function that doesn't have to be transformed. It just has some meaningful value at every point in space. Let's call the function S. Its values are $S(X)$, if we use X for the coordinates of a point.

Then the gradient of S is also a kind of vector. What is the gradient of S? It is the set of derivatives of S with respect to each X^p

$$\frac{\partial S(X)}{\partial X^p} \tag{21}$$

Example of scalar fields are the temperature, the atmospheric pressure, the Higgs field, whatever has, at any point in the space, a value that is not multi-dimensional but simply a number, and which doesn't change if we change coordinates.

The wind velocity is not a scalar field because at every point it has a vector value. It is a vector field. And if we tried to consider only the first component of the vector representing the wind, it would not be a scalar field, because it would not be invariant under change of coordinates.

So the gradient of a scalar function is a vector. It is the gradient vector field. But *it doesn't transform in the same way as contravariant vectors.* If we compute the gradient of S with respect to the Y coordinates instead of the X coordinates, we get a different set of values, and we want to know how there are related.

It is easy to figure out. Suppose we know the partial derivatives of S with respect to the X^p's, equation (21). We want to compute

$$\frac{\partial S}{\partial Y^m}$$

Again it is an elementary calculus problem. The rate of change in S when we change Y^m a little bit, keeping all the other Y coordinates fixed, is just derivative of S with respect to X^p times the derivative of X^p with respect to Y^m – summed over all the p's.

$$\frac{\partial S}{\partial Y^m} = \frac{\partial S}{\partial X^p} \; \frac{\partial X^p}{\partial Y^m} \tag{22}$$

This is some version of the chain rule (see Volume 1, Chapter 2 in which the chain rule is explained).

Again the right hand side of equation (22) means sum over p. We shall no longer write the big Σ_p sign meaning that the sum is over p. If there is a repeated index on one side, which doesn't show up on the other side, it must mean that it is summed over. That is Einstein summation convention. It made life a lot easier for publishers and printers who didn't have to put in summation signs all over the place. And it makes mathematics texts on certain topics less cluttered and easier to read.

Now let's see how equation (22) compares with equation (20) for contravariant vectors?

Let's denote W the gradient of S with respect to the X coordinates, and W' the gradient of S with respect to the Y coordinates. Equation (22) can then be rewritten

$$(W')_m = \frac{\partial X^p}{\partial Y^m} \, W_p \tag{23}$$

Notice that we now write the indices m of W' and p of W downstairs. We will see why this makes nice sense in a minute.

Equation (23) is the fundamental equation linking the primed and unprimed version of components of a covariant vector, that is its components in the Y system and in the X system.

Let's rewrite both equations (20) and (23) next to each other, and relabel them (24a) and (24b):

Contravariant vectors

$$(V')^m = \frac{\partial Y^m}{\partial X^p} \, V^p \tag{24a}$$

Covariant vectors

$$(W')_m = \frac{\partial X^p}{\partial Y^m}\, W_p \tag{24b}$$

They look very much alike except that $\partial Y^m / \partial X^p$ appears in the first one, and the inverse, $\partial X^p / \partial Y^m$, in the second.

Let's repeat that ordinary vectors – displacements of positions, or velocities for instance – are contravariant vectors. We saw that they change *contrary* to the basis change.

Gradients, on the other hand, change *like* the basis change. For instance if the new basis vectors are obtained simply by dividing the old ones by 10, then the gradient components will also be divided by 10. That is why such objects are called covariant. They are called covariant vectors, but they are not ordinary vectors. In fact in mathematics, they are dual vectors, like the components of a linear form that can be applied to ordinary vectors.

Equations (24a) and (24b) are fundamental equations for this course. The reader needs to understand them, become familiar and at ease with them, because they are completely central to the entire subject of general relativity.

He or she needs to know where the indices go for different kinds of objects, how these objects transform. That is in some sense what general relativity is all about. It is about the transformation properties of different kinds of objects.

Covariant and contravariant vectors and tensors

We have seen two ways to think about an ordinary vector. First of all we can think of it like we have learned in high school: it is a displacement with a length and a direction, that is, *an arrow* in a space. And this is geometrically well defined even before we consider any basis.

But we can also think of it more abstractly as some object which has components. And its components depend on the basis. If the components transform in a certain way when we change basis, namely according to equation (24a), then the object behaves exactly like our old vectors. Therefore we can also equate the object to an ordinary vector. In tensor analysis we call them contravariant vectors.

Similarly, some other objects have components which transform according to equation (24b). They cannot be equated to our old ordinary vectors, but to other geometric things – dual vectors which don't concern us much here. We call them covariant vectors.

A vector – be it contravariant or covariant – is a special case of a tensor. We are not going to define tensors geometrically. For us, at first, tensors will be things which are defined by the way that they transform. The way they transform means the way they change when we go from one set of coordinates to another.

Later on we will give a geometric interpretation of tensors. We will also go deeper into contravariant and covariant vectors. We will see that an object with one index can have a contravariant version and a covariant version. All this will be studied in the next chapter. For the time being, let's continue to proceed step by step in our construction of the mathematical tools necessary for general relativity.

So the next step, for us now, is to talk about tensors with more than one index.

The best way to approach tensors with several indices is to consider a special very simple case. Let's imagine the product of two contravariant vectors[11] . We consider the two contravariant vectors V and U, and we consider the product

$$V^m U^n$$

V and U don't have to come from the same space. If the dimensionality of the space of V is M, and the dimensionality of the space of U is N, there are MN such products. As usual, we use the notation $V^m U^n$ to denote one product as well as the collection of all of them – just like V^m denotes one component of the contravariant vector V, but is also a notation, showing explicitely the position of the index, of the full vector V itself.

Let's define T^{mn} as

$$T^{mn} = V^m U^n \tag{25}$$

Notice that it matters where we write the indices of T^{mn}, because T^{mn} is not the same as T^{nm}

T^{mn} is a special case of a tensor or rank 2. Rank two means that the collection of component products has two indices. It runs over two ranges: m runs from 1 to M and n runs from 1 to N. For example, if both V and U come from a

[11]It is not the dot product nor the cross product. It is going to be called the outer product or the tensor product. Anyway, it is an operation which, to two things, associates a third thing.

four dimensional space, there will be 16 components $V^m U^n$. In that case T^{mn}, as we saw, represents one component but also the entire collection of 16 components.

How does T^{mn} transform?

For example V^m and U^n could be the components of the vectors V and U in the unprimed frame of reference, the reference frame using the X coordinates. Since we know how the individual components transform, when we go to the Y coordinates, we can figure out how T transforms. Let's call $(T')^{mn}$ the mn-th component of the tensor in the prime frame.

$$(T')^{mn} = (V')^m (U')^n$$

Then using equation (24a) twice, this can be rewritten

$$(T')^{mn} = \frac{\partial Y^m}{\partial X^p}\ V^p\ \frac{\partial Y^n}{\partial X^q}\ U^q$$

The four terms on the right hand side are just four numbers, so we can change their order and rewrite it

$$(T')^{mn} = \frac{\partial Y^m}{\partial X^p}\ \frac{\partial Y^n}{\partial X^q}\ V^p\ U^q$$

Finally $V^p\ U^q$ is just T^{pq}. So the way T transforms is

$$(T')^{mn} = \frac{\partial Y^m}{\partial X^p}\ \frac{\partial Y^n}{\partial X^q}\ T^{pq} \tag{26}$$

We found in the special case of a product of vectors how T transforms. Now this leads us to the following definition:

Anything which transforms according to equation (26) is called a contravariant tensor of rank 2.

If there were more indices upstairs, the rule would be adapted in the obvious manner. A tensor of rank 3, all indices contravariant, would transform like this:

$$(T')^{lmn} = \frac{\partial Y^l}{\partial X^p} \frac{\partial Y^m}{\partial X^q} \frac{\partial Y^n}{\partial X^r} T^{pqr}$$

What kind of things are tensors like that? Many things. Products of vectors are particular examples, but there are other things which are not products and still are tensors according to the above definition.

We are going to see that the metric object g_{mn} is an tensor. But it is a tensor with covariant indices. So to finish this chapter let's see how things with covariant indices transform. Equation (24b) shows how an object with only one covariant index transforms. It is a tensor of rank 1 of covariant type.

Let's begin again with the particular case of the product of two covariant vectors W and Z. Their product transforms as follows

$$(W')_m (Z')_n = \frac{\partial X^p}{\partial Y^m} \frac{\partial X^q}{\partial Y^n} W_p Z_q$$

So here we have discovered a new transformation property of a thing with two covariant indices, that is two downstairs indices.

More generally let's call it T_{mn} – a different object, but we still use T for tensor. It is a tensor with two lower indices and it transforms according to this equation

$$T'_{mn} = \frac{\partial X^p}{\partial Y^m} \frac{\partial X^q}{\partial Y^n} T_{mn} \tag{27}$$

It is left to the reader to figure out how a tensor with one upper and one lower index must transform.

Next lesson, we will also see how the metric object g transforms. And we will discover that it is a tensor with two covariant indices.

Then the question we will ask is: given that equation (27) is the transformation property of g, can we or can we not find a coordinate transformation which will turn g_{mn} into δ_{mn}.

That is the mathematics question. It is a hard mathematics question in general. But we will find the condition.