Lesson 9 : Einstein field equations

Notes from Prof. Susskind video lectures publicly available on YouTube

Introduction

In this lesson, we finally arrive at Einstein field equations. We shall introduce the continuity equation which says that certains things cannot disappear from here to reappear there without passing in between, and the energy-momentum tensor which extends the idea of distribution of masses in space and their motions. Then we will be able to derive Einstein equations which are the analog in general relativity to Newton equations of motion in classical non-relativistic physics, in their version using the concept of field.

We are not going to get deep into solving the Einstein field equations. They are mathematically rather complicated. Even writing them down explicitely is intricate.

That is a feature of general relativity we have already mentioned : the principles are pretty simple, but the equations are computationally nasty. Almost everything we try to calculate gets complicated fast. There are lots of independent Christoffel symbols, elements of the curvature tensor, derivatives, etc. Each Christoffel symbol has a bunch of derivatives. The curvature tensor has more derivatives. The equations become hard to write on a single piece of paper.

The best way to solve them, or even write them, is just to feed them into your computer. And Mathematica [1] will spit out answers whenever it can. As said, the basic principles are simple, but going anywhere past the basics principles tends to be computationally intensive.

1. Symbolic mathematical computation program, conceived by Stephen Wolfram (born in 1959), English computer scientist, mathematician and physicist.

So we won't do much computation. We will concentrate on the meaning of the symbols. And then we will see what happens when we try to solve them in various circumstances.

In the next and final chapter of the book, we will do a little bit of solving when we talk about gravitational waves. But in the present chapter, the topic is not gravitational waves, it is the fundamental equations of general relativity : Einstein field equations.

Newtonian gravitational field

Before we talk about Einstein field equations, we should talk first about the corresponding Newtonian concepts.

Newton [2] didn't think in terms of fields. He didn't have a concept of field equations. Nevertheless, in classical non-relativistic physics, there are field equations which are equivalent to Newton's equations of motion.

Field equations of motion are always a sort of two-way street. Masses affect the gravitational field and the gravitational field affects the way masses move. So let's talk about the two way street in the context of Newton.

First of all, a field affects particles. That is just a statement that a gravitational force F can be written as minus a mass m times the gradient of the gravitational potential. The

2. Isaac Newton (1642 - 1727), English physicist and mathematician.

gravitational potential is usually denoted as ϕ which is a function of position x. So we have

$$F = -m \ \nabla\phi(x) \tag{1}$$

This equation means that everywhere in space, due to whatever reason, there is a gravitational potential $\phi(x)$ which varies from place to place. You take the gradient of $\phi(x)$. You multiply it by m, the mass of the object whose motion you want to figure out, and put a minus sign. And that tells you the force on the object.

Equation (1) can be one dimensional, in which case the gradient is just the ordinary derivative of the function ϕ with respect to the independent variable x. And equation (1) equates two scalars.

Or equation (1) can be multi-dimensional. It is the case when x, is a *vector* position, which we could then denote $X = (x,\ y,\ z)$. The sign ∇, which reads "gradient" or "del", is then the vector of partial derivatives

$$\nabla\phi(X) = \left(\frac{\partial\phi}{\partial x},\ \frac{\partial\phi}{\partial y},\ \frac{\partial\phi}{\partial z}\right)$$

In that case, of course, the force F is also a vector. For simplicity, however, we will keep with the notation x, be it one- or multi-dimensional. Same for F and for the acceleration introduced below.

So equation (1) is one aspect of field $\phi(x)$. It tells particles how to move. In the case which occupies us, it does that by telling them what their acceleration should be. This comes from Newton's equation linking the force F exerted on the particle to its acceleration a :

$$F = ma \tag{2}$$

Combining equations (1) and (2), the mass m cancels. And we obtain

$$a = -\nabla\phi(x) \tag{3}$$

That is one direction of the two way street. It is how a field tells a particle how to move.

On the other hand – that is the other direction – masses in space tell the gravitational field what to be. The equation that tells the gravitational field what to be is Poisson equation [3]. It says that the second derivative of ϕ with respect to space (or the sum of second partial derivatives) is related to the distribution of masses in space as follows

$$\nabla^2\phi = 4\pi G\ \rho(x,\ t) \tag{4}$$

Here is how to read equation (4) :

a) As said, if we are in several dimensions [4],

$$\nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}$$

b) The multiplicative factor 4π is essentially a convention. It originates from the geometry of the sphere, because we often deal with a field that is invariant under rotation, therefore spherical.

---

3. Named after Siméon Denis Poisson (1781 - 1840), French mathematician and physicist.

4. ∇^2, which is a notation for the formal dot product $\nabla.\nabla$, is called the *laplacian*, and has its own notation : it is sometimes denoted $\triangle$.

c) G is Newton gravitational constant. It is equal to

$$6.674 \text{ x } 10^{-11} Nm^2/kg^2$$

d) Finally, ρ is the *density of mass* at each position x, or $(x,\ y,\ z)$. And it can depend also on time because masses can move around.

That is how the distribution of masses in space determines the gravitational field.

So we have these two aspects : the field tells particles how to move, and masses, particles in other words, tell the field how to be – how to curve, we shall see.

We can easily solve equation (4) in the special case where the distribution of masses (or the density [5]) ρ is simply all the mass concentrated at one point.

In general, the density ρ is by definition the amount of mass per unit volume.

In the case where $\rho(x)$ is "concentrated", it is actually a Dirac distribution – we met these "extended" mathematical functions in volume 2 of the collection *The Theoretical Minimum*. We manipulate them essentially through their integral over some region.

At first, we may be considering a real star, or a planet or a bowling ball of mass M, see figure 1. The only constraint is that it must be a spherically symmetric object of total

5. The mass in any small volume around a point x being then the local density of mass $\rho(x)$ multiplied by the small volume of space dx or $dx\ dy\ dz$.

mass M. It does not matter whether the mass is uniformly distributed inside the volume. It just has to be symmetric with respect to rotation. There could be layers with more density near the center than near the outer surface.

Anyhow, once we are outside the region where there is mass, according to Newton's theorem which we saw in the previous chapter we can treat the body as a point mass, that is, a Dirac density with its peak at the center. And we can easily find the unique solution to equation (4).

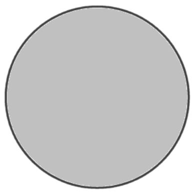

Figure 1 : Spherically symmetric object of mass M, equivalent to a point mass when we are outside its surface.

The solution to equation (4) then is

$$\phi(r) = -\frac{MG}{r} \tag{5}$$

where we replaced the spatial coordinate x by a radius r.

If we consider that the body is a point mass, then equation (5) is the valid solution to equation (4) everywhere except at the central point (where ϕ is not defined because $-MG/r$ tends to infinity).

Let's pursue this example where ϕ is given by equation (5). Taking its gradient just puts another r downstairs, while changing the sign. Indeed, differentiating $1/r$ with respect to the radius, which is what ∇ does, produces $-1/r^2$. We wind up with

$$F = \frac{mMG}{r^2} \tag{6}$$

This is a kind of primitive field theoretic way to think about gravitation. Instead of action at a distance we have a gravitational field.

In truth it still implies action at a distance, because in Newton's theory if you move around a mass, ϕ instantly reacts to it and changes, from equation (4). But it is a way of writing the theory, making it look like a field theory.

Equations (3) and (4), corresponding to "field $\rightarrow$ motion of mass" and "mass$\rightarrow$ structure of field", is what we want to replace by something making sense in general relativity.

Just to get a little bit of a handle on what we are going to do in this chapter, let's remember something about the Schwarzschild geometry. In chapter 5, we pulled the Schwarzschild geometry out of a hat. Of course the point was that it is the solution to Einstein's equations in the case of a central point mass field. So we did not have the tools to derive it.

Equations (3) and (4) are not Einstein's equations, they are Newton's equations. And their solution, in the simple case of a central spherically symmetric field, is $\phi(r) = -MG/r$.

So we wrote down the Schwarzschild metric, see equation (32) of chapter 5. We are not going to write the whole shebang again. But let's examine what g_{00} was.

$$g_{00} = 1 - \frac{2MG}{r} \tag{7}$$

As usual, we have set everywhere the speed of light c equal to 1.

Our aim presently is to guess some correspondence between g_{00} given by equation (7), and anything in the Newton framework.

Using equation (5), the above equation (7) can be rewritten as $g_{00} = 1 + 2\phi$. Therefore $\nabla^2 g_{00}$ is just twice $\nabla^2 \phi$. This leads to

$$\nabla^2 g_{00} = 8\pi G\rho \tag{8}$$

This formula should be taken with a grain of salt. It is just a mnemonic device to remember the relationship between some aspects of general relativity and matter.

Nevertheless, interestingly enough, it already suggests that matter or mass is affecting geometry. When we make this correspondence between Newton ϕ and Einstein or Schwarzschild metric, we see roughly that matter is telling geometry how to curve so to speak. We use the word roughly because we are going to be more precise.

Of course equation (8) is not really the part "mass $\rightarrow$ field" of Einstein's equations. In Einstein's equations, it is good deal more complicated than that.

And what about the other part "field $\rightarrow$ motion of mass"? In Newton's framework we saw equations (1) or (2) which specify how the field dictates the motion of mass. In general relativity, the equation $\ddot{x} = -\nabla\phi(x)$ is replaced by the statement that once we know the geometry, that is, once we know g_{00}, the rule is :

particles move on space-time geodesics.

And for the other direction, we saw that the Newtonian field equation, $\nabla^2\phi = 4\pi G\rho$, is replaced by something which we naively wrote as $\nabla^2 g_{00} = 8\pi G\rho$. We are going to do better. We have to figure out exactly how the mass distribution affects the field[6]. Well, a little more than a century ago, Einstein figured out exactly what goes there.

Before we write down the field equations, we need to understand more about the right hand side $4\pi G\rho$ of Newton field equations, or $8\pi G\rho$ of Einstein field equations. We need to know more about the density of mass.

Mass really means energy. Indeed, we saw in volume 3 on special relativity that $E = mc^2$. If we forget about c, that is, if we set it equal to one, then energy and mass are the same thing. Therefore what goes on the right hand side of the field equations is really energy density.

6. Sometimes Einstein field equations refer only to the direction "mass $\rightarrow$ field", and are the analog only of $\nabla^2\phi = 4\pi G\rho$. And sometimes by them we mean the equations corresponding to both directions, that is the two way relationship between matter and field.

We need to understand better what kind of quantity in relativity energy density is.

It is part of a complex of things which includes more than just the energy density. It is part of a complex. In other words it is part of some kind of tensor whose other components have other meanings.

So let's go back and review quickly a little bit about the notion of conservation. In the case that we are concerned with, it will be conservation of energy and momentum.

But we shall begin with a simpler case : conservation of charge.

Continuity equation

Conservation, densities, flows of things like charge and mass, let's review briefly these concepts.

Let's start with electric charge. It is simpler than energy for reasons we will come to.

The total electric charge of a system is called Q. It is the standard notation for the electric charge. In many situations the electric charge *density* is called ρ, but in order not to create confusion with the mass density or the energy density, which we have already called ρ, we shall denote the charge density with the letter σ.

What is electric charge density ? Consider a small volume of space, a differential volume of space at a given point, take the electric charge in that volume, divide it by the volume. That gives you the electric charge density at that point. The density can be schematically written

$$\sigma = \frac{Q}{\text{Volume}} \tag{9}$$

It has units of charge divided by volume.

Let's look at an infinitesimal volume, figure 2.

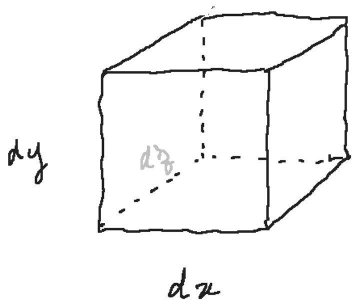

Figure 2 : Infinitesimal volume.

If we look at electric charges moving around in space, between t_1 and t_2, we see some entering the volume $dx\ dy\ dz$, some leaving, some passing by, some passing through, etc., figure 3.

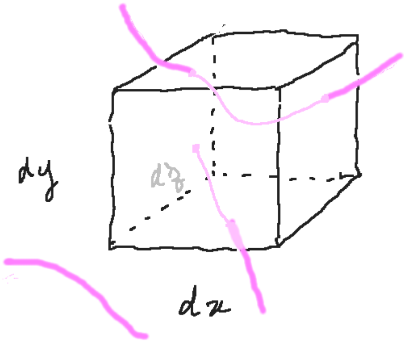

Figure 3 : Trajectory in space, between t_1 and t_2, of various moving electric charges.

The current in direction x is defined as the quantity of charge passing through the side $dy\ dz$ per unit time. It is denoted J^x. There are analogous definitions and notations for y and z.

Then, similarly to equation (9), notice that each current along one spatial dimension can be schematically written

$$J^m = \frac{Q}{\text{Area Time}} \tag{10}$$

But in relativity we have four dimensions more or less on an equal footing. Time and space are nice and symmetric to each other. So dividing by area times time is again dividing by three dimensions. One happens to be time-like, two happen to be space-like.

Both equations (9) and (10) can be thought of as a charge divided by a three-dimensional volume. In the case of equation (9) is it a pure spatial volume. In the case of equation (10) it is a mixed space and time three-dimensional thing.

The combined density σ plus the three currents J^x, J^y, and J^z happen to form a 4-vector in the sense of relativity [7].

Following the standard notation we became familiar with in volume 3, instead of $(\sigma,\ J^x,\ J^y,\ J^z)$ we simply write

$$J^\mu \tag{11}$$

where the Greek index μ runs from 0 to 3. J^0 (read "J naught") is the electric charge density σ. And the other three elements are simply the components of the current indexed with 1, 2 and 3 instead of x, y and z.

Now we arrive at an important physical fact :

the conservation of charge is a local property.

We haven't yet talked about the familiar relationship between the evolution of the charge Q inside the volume $dx\ dy\ dz$ and the current going through its sides, because we want to stress that in theory conservation could hold in a system without being local.

The electric charge carried by an object in a system, could at time t disappear from the object and instantaneously reappear elsewhere in another distant part of the system. If the system we are considering is the entire universe, the

7. That is, in a change of coordinates they change according to the tensor equations of a change of frame of reference, see chapter 2.

charge could disappear from my pen and reappear in Alpha Centauri. I always use Alpha Centauri as someplace which is so far away that it doesn't matter.

If that were possible, conservation would still hold. You would say rightly : well charge is conserved. But I would retort : who cares if charge is conserved, if it can just disappear to some arbitrarily distant place. It is just as good as saying it wasn't conserved.

In the laboratory, however, charge doesn't disappear that way. If it leaves the laboratory, it goes through the walls, or the windows, or whatever, or simply through the door. In other words, the electric charge cannot change in a given volume without some current flowing through the boundary.

That idea is called *continuity*. And there is an equation that goes with it. The equation is the *continuity equation*.

If we look at a box of volume one in some units, the charge inside the box is σ, that is the charge density times one And the charge leaving the box over a unit time is $-\dot{\sigma}$. In other words, it is minus the time derivative of σ. Why minus ? Because it is leaving the box.

That has to be equal to the sum of the currents passing through the box. A basic theorem of multivariate calculus, called the *divergence theorem*, also known as Gauss's theorem or Ostrogradsky's theorem [8], states that the dimi-

8. It was actually first stated by Lagrange in 1762, then independently by Gauss in 1813, Ostrogradsky in 1826, Green in 1828, etc. It is the multivariate generalization of $\int_a^b f(x)dx = F(b) - F(a)$.

nution of charge inside the box is equal to the divergence of the current.

$$-\dot{\sigma} = \nabla . J \tag{12}$$

The formal dot product $\nabla . J$ on the right hand side means $\frac{\partial J^x}{\partial x} + \frac{\partial J^y}{\partial y} + \frac{\partial J^z}{\partial z}$.

Remembering that t is one of the four components of space-time in relativity, this can be rewritten more nicely, first of all as

$$\frac{\partial \sigma}{\partial t} + \frac{\partial J^x}{\partial x} + \frac{\partial J^y}{\partial y} + \frac{\partial J^z}{\partial z} = 0 \tag{13}$$

Then, since J^0 is by definition σ, see expression (11), using the summation convention we can write this even more nicely

$$\frac{\partial J^\mu}{\partial X^\mu} = 0 \tag{14}$$

So the divergence theorem has lead us to a nice tensor-type equation satisfied by J^μ.

J^μ is 4-vector. The X^μ's are the four components of space-time. Therefore equation (14) has the earnest look of a good equation involving the derivative of a tensor with respect to position, and contracting it. Since equation (14) is true in any frame (because it is a tensor equation), it expresses a conservation law.

In curved coordinates, in general, an expression like equation (14) would be correct only in flat space. In curved coordinates, we might have to replace the ordinary derivative by the *covariant derivative* of the tensor.

$$\frac{DJ^\mu}{DX^\mu} = 0 \tag{15}$$

Remember covariant derivative of tensors which we studied in chapter 3. It turns out in the case of charge current it doesn't matter. But in general it would matter.

When you go to curved coordinates, you should replace all derivatives by covariant derivatives, otherwise the equations are not good tensor equations.

Why do we want tensor equations ? We want tensor equations because we want them to be true in any set of coordinates.

That is the theory of electric charge flow of current and the continuity equation. Equation (15) is called the continuity equation. And the physics of it is that when charge either appears in or disappears from a volume, it is always traceable to current flowing into or out through the boundaries of that volume.

Now let's come to energy and momentum.

Energy-momentum tensor

Energy and momentum are also conserved quantities. They can be described in terms of density of energy and density of momentum, that is, density of each component of momentum because it is a vector quantity.

We can ask how much energy there is in the form of particles or whatever it happens to be, including the mc^2 part of energy. We can ask how much energy is in a volume.

We can ask how much momentum is in a volume. Just look at all the particles within a volume and count their momentum.

Photons, or electromagnetic radiations, have both energy and momentum. And that energy and momentum can be regarded as the integral of a density.

So, in that sense, each one of them, each energy and each component of the momentum, is like the electric charge. They are conserved. They can flow. If an object is moving, the energy and momentum may be flowing.

The question is : how do we represent the same set of ideas for energy and momentum as we did for charge ?

Now there is a difference between charge and energy and momentum. Electric charge is an invariant. No matter how the charge is moving, the charge of an electron is always the same. It does not depend on its state of motion. Therefore charge itself is an invariant.

The density of charge and the current of charge, however, are not invariants. For example, suppose we have a given charge, taking a certain volume of space, and I decide to look at it in a different frame of reference than you. I walk by it with a certain velocity. Because of Lorentz contraction I say that the volume of that charge is one thing. You, sitting still relative to the charge, assign a different volume to it. Since the charge itself does not change, we will not agree about the value of the charge density.

That's okay : charge density is a component of a 4-vector. Remember 4-vectors. Depending on the reference frame, their components change according to tensor equations of change of frame. But if the 4-vector is zero in one frame, it is zero in every frame.

Similarly, I see charges moving, and you see them at rest. You say there is no current. I'm moving and I see a wind of charges passing me by. I say there is current.

We are both right of course. Charge density and charge current are not invariant. They form together a 4-vector.

Energy and momentum are more complicated. The total energy and momentum, not the density of them but the total energy and momentum, are not invariant.

I see a particle standing still – I'm talking about the whole particle not the density. I say there is some energy of a certain magnitude. You are walking past it and you see not just the $E = mc^2$ part of the energy, but you also see kinetic energy of motion.

You see more energy in the particle or object not because of any Lorentz contraction of the volume it is in, but because the same object when you look at it has more energy than when I look at it.

The same is true of the momentum, not the flow of it, not the density of it, but the momentum itself. It is also frame dependent. You see an object in motion, you say there is momentum there. I see the same object at rest, I say there is no momentum.

So energy and momentum, unlike charge, are not invariant. They together form the components of a 4-vector.

$$P^{\mu} = (E,\ P^{m}) \tag{16}$$

E is the energy. It is also P^0. And the P^m are the components of momentum, where the Roman index m labels the directions of space. So zero is the index for energy, and the others for momentum.

Each of the four components of P^{μ} is like a charge Q. In other words, it is a conserved quantity in the sense of invariance of 4-vectors : if two 4-vectors are equal in one frame, they are equal in every frame.

In equations (14) or (15), density was the zeroth component of J^{μ}. So zero is also the index for density, and the others for flow.

Let's return to the energy in P^{μ}. It is P^0. And consider the *density of energy*. In other words, how much energy is in a small volume. We are going to denote this

$$T^{00} \tag{17}$$

The first index 0 is refering to energy as opposed to momentum, and the second 0 to the fact that we are considerting a density.

It is a function of position, so we will sometimes write

$$T^{00}(X) \tag{18}$$

where X stands for $(X^0,\ X^1,\ X^2,\ X^3)$.

Energy can also move. Like charge it can disappear from a region, and like charge too it does so by crossing the boundary of the region. In other words, the continuity equation applies to energy density and energy flow, as it applied to charge density and electric current.

The amount of energy flowing through a unit surface, along the X^1 direction, per unit time – the "current of energy" you might call it – is denoted

$$T^{01}(X) \tag{19}$$

the first index 0 refering to an energy, and the second index 1 to a flow along the first spatial axis. Likewise there is a T^{02} and a T^{03}.

In summary, the three components ($T^{02},\ T^{02},\ T^{03}$) form the flow of energy. And T^{00} is the density of energy.

The continuity equation for energy is derived in exactly the same way as we did for charge. It involves now the components $T^{0\nu}$. And what does it say ? For the moment the first index 0 is just passive It signals that are talking about energy. It is the second index ν running over 0, 1, 2 and 3 which is like the index μ of J^μ in equation (15). Since ν is a dummy index, we may as well use μ. The *continuity equation for energy* then is

$$\frac{DT^{0\mu}}{DX^\mu} = 0 \tag{20}$$

Next step : everything that we said about energy we could now say about any one of the components of the momentum.

So let's go to the momentum components of the 4-vector P^μ. It is the last three components, customarily indexed with the Roman letter m running from 1 to 3.

The component P^m also has a density, denoted T^{m0}. The naught indicates that it is a density. The m indicates that we are talking about the m-th component of momentum. It is also a function of X.

Likewise we can consider the flow of the m-th component of momentum, along n-th direction of X. It is naturally denoted T^{mn}.

Momentum is a conserved quantity, in the sense that, with the energy, the three momenta form a 4-vector. Momentum can flow. Each of its components can flow along some direction.

So we can go a little further. We can say the same equation as equation (20) is true even if we replace energy by a component of momentum. In other words we could replace naught by n. But then we have the same equation for all four possibilities of the first index. So, switching μ and ν again for cosmetic reasons, we can finally write

$$\frac{DT^{\mu\nu}}{DX^{\nu}} = 0 \tag{21}$$

We can now step back and see where we arrived. The densities and flows of energy and momentum form a tensor with two indices. The first index tells us what were talking about, energy or momentum. The other index tells us whether we are talking about density or flow. That is called the *energy-momentum tensor*.

The energy-momentum tensor has an interesting property which we have not proved. Take for example T^{m0}, that is, the density of the m-th component of momentum. Compare it with T^{0m}, that is, the flow of energy in the m-th direction.

It is a general property of relativistic systems, which we are not going to prove now, that the matrix formed by the components of the energy-momentum tensor is symmetric. In other words, $T^{m0} = T^{m0}$, and more generally $T^{\mu\nu} = T^{\nu\mu}$.

It takes a little bit of work to prove it. Relativistic invariance allows you to connect T^{m0} with T^{m0}. It is a theorem of all relativistic field theories, that the energy-momentum tensor is symmetric. So let's take this into account and write the energy-momentum tensor in matrix form

$$T^{\mu\nu} = \begin{pmatrix} T^{00} & T^{01} & T^{02} & T^{03} \\ T^{01} & T^{11} & T^{12} & T^{13} \\ T^{02} & T^{12} & T^{22} & T^{23} \\ T^{03} & T^{13} & T^{23} & T^{33} \end{pmatrix} \tag{22}$$

We will come back to the meaning of these elements in a little while. T^{00} is clear : it is energy density. The components $(T^{01},\ T^{02},\ T^{03})$ are fairly clear : they are flow of energy. T^{10} (which is the same as T^{01}) is momentum density. Then we have flow of momentum. So their meaning is pretty clear. But we will find out that some of the elements of $T^{\mu\nu}$ have other meanings, connected with pressure, things of that nature.

The important idea is that the flow and density of energy and momentum are combined into an energy-momentum tensor. And this energy-momentum tensor satisfies four continuity equations, one for each type of the stuff that we are talking about (energy, that is P^0, or momentum along one of the three spatial axes, that is P^m).

What we have learned it that the notion of energy density, ρ, which we tried to relate to the metric in equation (8), reproduced below,

$$\nabla^2 g_{00} = 8\pi G\rho \tag{23}$$

is incomplete. It is part of a complex of things. It is part of a tensor of course.

This is a fundamental observation : the right hand side of equation (23) is part of a tensor, because ρ is T^{00}. So the left hand side must also be part of a tensor.

In physics an equation which says that some particular component of a tensor is equal to the same component of some other tensor is usually meaningless or wrong – unless its full version say that the two tensors are equal. We will understand this with vectors, which are a simple type of tensor.

Suppose we have two 3D spatial vectors A and B. Suppose we assert that there is a law of physics which says that A_3 is equal to B_3, the index 3 being the z direction. Does this make sense as a law of physics ?

Well, it only makes sense as a law of physics if it is also true that $A_2 = B_2$ and $A_1 = B_1$. Why is that ? Because if it is a law of physics, it means that it is true in every reference frame. But we can always rotate the coordinates so that the first axis in the old frame becomes the third axis in the new frame, so A_1 must also be equat to B_1, etc.

This is an illustration of the general idea that vectors and tensors are geometric entities which exists independently of any consideration of coordinates or components. "Go five steps over there", if I show you a direction while telling you that, is meaningful without refering to coordinates. Of course, with appropriate coordinates, it may become "go three steps to the east, and four steps to the south". But the coordinates are frame dependent. My injonction may have first coordinate 3 in one frame and first coordinate 2 in another.

When we go to relativity, the same considerations are true for 4-vectors, including the time component, because it can be transformed into a space component through a Lorentz transformation. And they are also true for any two tensors of the same type.

So tensor equations, to be good laws of physics, must equate all the components of two tensors, not just two components of them. And if we have an equation, which we believe should reflect a law of physics, which equates two components of tensors, then it should be true for all the components of the two tensors.

Equation (23) is a formula which equates the naught naught component of the energy-momentum tensor with something else. If it is to be a law of physics, we should figure out which tensors it actually equates.

Let's not worry too much about whether the left hand side is correct or not. We just made a guess of what the left hand side might look like, and found something related to the metric, that is, to the geometry.

But the right hand side is the energy density, and it is what you would expect to be on the right hand side of *Newton's equations*.

Therefore the right hand side of *Einstein's equations* must involve not a particular component of a tensor, but it must generalize to something that involves all components of a tensor. That means Einstein's generalization of Newton must be as follows. The right hand side ought to be

$$... = 8\pi G T^{\mu\nu}$$

A special case of it is when both μ and ν are time. Then it becomes the energy density. But if the equation that we are constructing is to be true in every frame, it has to be a tensor equation involving all the μ ν components.

What has to be on the left hand side ? It must also be a rank 2 tensor. Otherwise the equation would not make sense. It must be symmetric because the right hand side is symmetric. It must have whatever other properties the right hand side has. But the left hand side will not be something made up out of matter. It will be made up out of the metric. It has to do with geometry and not masses and sources.

So we shall call the left hand side $G^{\mu\nu}$, and we shall pursue our guess work on the equation

$$G^{\mu\nu} = 8\pi G T^{\mu\nu} \tag{24}$$

The only thing we know about $G^{\mu\nu}$ is that it is made up out of the metric. It probably has second derivatives in it to be the analog of the laplacian of Newton equation (4). In other words, it will involve the metric in some form or other and very likely second derivatives of the metric.

Ricci tensor and curvature scalar

We have begun to see what kind of object we would like to find and put on the left hand side of equation (24), which will be the generalization of Newton's field equation in the general relativity setting of Einstein.

When we find a good candidate for it, in order to evaluate it we can ask the following question : when we are in a situation where non-relativistic physics should be a good approximation, does this $G^{\mu\nu}$ reduce to just $\nabla^2 g_{00}$? Perhaps it does. Perhaps it doesn't. If it doesn't, then we throw it away and try to find a different candidate.

Let's explore the possibilities for $G^{\mu\nu}$. It is a tensor made up out of the metric. It has second derivatives, or at least it must have some terms which have second derivatives. So it is not the metric by itself. What kind of tensor can we make out of a metric and second derivatives ? We have already met one : it is the curvature tensor.

Let's recall what is the curvature tensor. It is made up out of the Christoffel symbols, which themselves are functions of the metric.

Let's start with the Christoffel symbols

$$\Gamma^{\sigma}_{\nu\tau} = \frac{1}{2}\, g^{\sigma\delta} \;\left[\, \partial_\tau g_{\delta\nu} + \partial_\nu g_{\delta\tau} - \partial_\delta g_{\nu\tau} \,\right] \tag{25}$$

The only important thing in this equation is that the right hand side involves the first derivatives of g terms.

Next, let's write the curvature tensor in all its glory [9].

$$\mathcal{R}_{\mu\nu\tau}^{\ \ \ \ \sigma} = \partial_\nu \Gamma^{\sigma}_{\mu\tau} - \partial_\mu \Gamma^{\sigma}_{\nu\tau} + \Gamma^{\lambda}_{\mu\tau}\Gamma^{\sigma}_{\lambda\nu} - \Gamma^{\lambda}_{\nu\tau}\Gamma^{\sigma}_{\lambda\mu} \tag{26}$$

This rank 4 tensor is the object – the "mathematical probe" – which tells us whether there is real curvature in spacetime, not just curvature due to curvilinear coordinates that could be flattened (like a uniform acceleration can be flattened, see chapter 1 fig 4, and chapter 4 fig 17). If any component of the curvature tensor is non-zero at any point or in any region of space, then the space is curved.

The right hand side of equation (26) uses the summation convention for the index λ in the last two terms. Again, the only important point in the expression of $\mathcal{R}_{\mu\nu\tau}^{\ \ \ \ \sigma}$ is that it involves another differentiation.

The Christoffel symbol, which is not a tensor, involves first derivatives of the metric. And the first two terms of the

9. It is exactly the same as expression (25) in chapter 3, with different dummy indices.

curvature tensor are first derivatives of the Christoffel symbols. Therefore the curvature tensor has second derivatives of the metric in its first two terms, and squares of first derivatives in its last two terms.

Therefore $\mathcal{R}$ is a candidate, or various functions of it are candidates, to appear on the left hand side of equation (24). But wait ! The Riemann curvature tensor $\mathcal{R}_{\mu\nu\tau}^{\sigma}$ has four indices, whereas the energy-momentum tensor has only two.

What can we do with the curvature tensor to transform it into a thing with only two indices ? We can contract it. Remember the rule : if you set the upper index σ equal to one of the lower indices and apply the summation convention, we eliminate two indices and get a rank 2 tensor. If we used τ for the lower index in the contraction, we would discover that we get zero. The various symmetries and minus signs in the expression of $\mathcal{R}_{\mu\nu\tau}^{\sigma}$ would wind up yielding 0.

However if we set σ and ν equal, we won't get zero. We will get something that we can call the tensor $\mathcal{R}_{\mu\tau}$. So from a tensor with four indices, by contraction we can build a tensor with two indices. But we have to be careful not to get zero.

We won't get the zero tensor when we contract σ with ν or σ with μ. And in fact the two tensors we get by contracting σ with ν and σ with μ happen to be the same tensor except for the sign.

So there is only one tensor that we can build out of two derivatives acting on the metric and which has only two indices. It is actually a well-known theorem.

This tensor is called the *Ricci tensor*. And it is a contraction of the Riemann tensor. The Riemann tensor has a lot more components. The Ricci tensor has less information. As a consequence, it can be zero while the Riemann tensor is not zero.

As always, if you have a tensor you can raise and lower its indices. That means there are also things denoted $\mathcal{R}_{\tau}^{\mu}$, $\mathcal{R}_{\mu}^{\tau}$ and $\mathcal{R}^{\mu\tau}$. You can raise and lower indices using the metric tensor.

Another fact about the Ricci tensor, is that it happens to be symmetric. And this is true of its version with upper indices as well

$$\mathcal{R}^{\mu\tau} = \mathcal{R}^{\tau\mu} \tag{27}$$

Now, this you just check by using its definition. I don't know any simple quick argument. All these tensors have fairly complicated expressions, but checking the facts we stated is most of the time pretty mechanical.

The Ricci tensor $\mathcal{R}^{\mu\tau}$ is symmetric. The right hand side of equation (24) is a rank 2 tensor, with two contravariant indices, and is symmetric. So the Ricci tensor $\mathcal{R}^{\mu\tau}$ is a candidate for the left hand side :

$$\mathcal{R}^{\mu\nu} \ \ =? \ \ 8\pi G T^{\mu\nu} \tag{28}$$

Now there is another tensor that we can make out of the Ricci tensor. It is by contracting its two indices (after having raised or lowered one of them). We define

$$\mathcal{R} = \mathcal{R}_{\mu}^{\mu} = \mathcal{R}^{\mu\tau} g_{\mu\tau} \tag{29}$$

The $\mathcal{R}$ defined by equation (28) is called the *curvature scalar*. It is a scalar. It has no indices left at all. So it's not something we want on the left hand side of equation (24). But we can multiply it.

For instance we can multiply it like this $g^{\mu\nu}\mathcal{R}$. This does give us again a tensor of the type we are looking for. That is another possibility.

$$g^{\mu\nu}\mathcal{R} \ =? \ \ 8\pi G T^{\mu\nu} \tag{30}$$

I'm not recommending either of these at the moment. I'm just saying from what we have said up till now, either of these could be possible laws of gravitation. The left hand sides of equations (28) or (30) both involve second derivatives of the metric tensor, equated to something on the right hand side which looks like a density and flow of energy and momentum.

Which one should we pick ? Well, we know one more thing. That is the conservation of energy and momentum or better yet the continuity equation for energy-momentum. If we believe that energy and momentum have the property that they can only disappears by flowing through walls of systems, then we are forced to the conclusion that

$$D_\mu T^{\mu\nu} = 0$$

That is the continuity equation.

From that, it follows that we must also have

$$D_\mu G^{\mu\nu} = 0 \tag{31}$$

So the first thing we could do is check whether either of the candidate solutions to equation (24), that is, those in equations (28) or (30) satisfy condition (31). If not then the left hand side simply cannot equal the right hand side, unless we give up local continuity of the energy and momentum.

Let's check the left hand side of equation (30). We want to calculate

$$D_\mu \ (g^{\mu\nu}\mathcal{R})$$

Covariant derivatives satisfy the usual product rule of differentiation. So we have

$$D_\mu \ (g^{\mu\nu}\mathcal{R}) = (D_\mu \ g^{\mu\nu})\mathcal{R} + g^{\mu\nu}(D_\mu \ \mathcal{R}) \tag{32}$$

First of all, the covariant derivative of the metric tensor is zero. It is a consequence of the definition of covariant derivative, see chapter 3. Covariant derivatives are by definition tensors which in the special good frame of reference are equal to ordinary derivatives. But the good frame of reference is by definition the frame of reference in which the derivative of g is 0. So the first term on the right hand side of equation (32) is eliminated.

Secondly, $\mathcal{R}$ is a scalar. The covariant derivative of a scalar is just the ordinaray derivative. So we can rewrite

$$D_\mu \ (g^{\mu\nu}\mathcal{R}) = g^{\mu\nu} \ \partial_\mu \mathcal{R} \tag{33}$$

Certainly, in general, the derivative of the curvature is not zero. We know that there are geometries which are more curved in some places and less curved in others. So it cannot be the case that $\partial_\mu \mathcal{R}$ is identically equal to zero. And the factor $g^{\mu\nu}$ doesn't change anything. So we are lead to eliminate the candidate $g^{\mu\nu}\mathcal{R}$.

What about the other candidate, the Ricci tensor with upper indices ? We do the same thing. We calculate

$$D_\mu \ \mathcal{R}^{\mu\nu}$$

The calculation is a little harder, but not much. The answer is

$$D_\mu \ \mathcal{R}^{\mu\nu} = \frac{1}{2} \ g^{\mu\nu} \ \partial_\mu \mathcal{R} \tag{34}$$

Again, it cannot be zero, for the same reason that the right hand side of equation (33) can't be zero. The covariant derivative of the Ricci tensor happens to be exactly one half that of the other candidate, $g^{\mu\nu}\mathcal{R}$.

But now we know the answer !

Einstein tensor and Einstein field equations

We now know which tensor on the left hand side of equation (24) will satisfy all the conditions, including having covariant derivative zero. The left hand side of the gravity equation (24) that will work is

$$G^{\mu\nu} = \mathcal{R}^{\mu\nu} - \frac{1}{2} \ g^{\mu\nu} \ \mathcal{R} \tag{35}$$

And a theorem can be proved, which says that there is no other tensor (up to a multiplicative factor) made up out of two derivatives acting on the metric that is covariantly conserved.

Of course, we could have twice $G^{\mu\nu}$ or half of it, or seventeen times it. But now it just becomes a question of matching equation (24), with the $G^{\mu\nu}$ we found, to Newton's equations in the appropriate approximation, that is, where everything is moving non-relativistically. Either there is some correct numerical multiple that will ensure the match or there isn't. If there isn't, then we're in trouble.

When we say that we want to match formulas, we mean that we take equation (24), we look at the time-time component of $G^{\mu\nu}$ given by equation (35), and we want, with some appropriate multiplicative factor in front of $G^{\mu\nu}$, in the non-relativistic limit – that is, everybody moving slowly, not too strong a gravitational field – to find Newton's equation (4).

The answer is that it works. And the multiplicative factor is simply 1. It is just a piece of luck that it is not some other number. There is nothing deep about this factor.

So, we finally arrived at the equation we were looking for, generalizing Newton's equation.

$$\boxed{\mathcal{R}^{\mu\nu} - \frac{1}{2}\, g^{\mu\nu}\, \mathcal{R} = 8\pi G\, T^{\mu\nu} \qquad (36)}$$

The path we followed in this lesson is essentially Einstein's calculations. He knew pretty much what was going on. But he didn't quite know what the right equation linking the energy-momentum tensor and the geometry was. I believe in the beginning he actually did try $\mathcal{R}^{\mu\nu} = T^{\mu\nu}$. But he eventually realized that it didn't work. And he searched hard for a suitable left hand side. I don't know how many weeks of work it took him to do all of this [10]. But in the end he discovered $G^{\mu\nu}$.

$G^{\mu\nu}$ is the *Einstein tensor*. $\mathcal{R}^{\mu\nu}$ is the *Ricci tensor*. $\mathcal{R}$ is the *curvature scalar*. Equation (36) is now known as *Einstein field equations*. They generalize Newton equations. And in the appropriate limit, they do reduce to Newton equations.

The continuity equation has played a fundamental role in this derivation. It is what leads us to look only for candidates, on the left hand side, with are covariantly conserved. Notice that the principle is as fundamental in classical non-relativistic physics as in relativity. It says that conservation of energy or of momentum is a local property. Things cannot disappear here and immediately materialize in Alpha Centauri. They have to travel through intermediate locations.

There is another way to derive Einstein field equations, using the action principle. It is much more beautiful and much more condensed. We introduce the principle of least

10. In early summer 1915, at the invitation of Hilbert, Einstein gave a series of lectures in Göttingen, exposing his work of the past few years and his present problem to link sources and geometry. This lead, during the summer and fall of 1915, to an exchange of friendly letters, but also to a race, between Einstein and Hilbert. Einstein published his final correct equations on November 25. Simultaneously Hilbert published the same equations obtained from the lagrangian approach. That is the reason why the name of Hilbert is sometimes also attached next to Einstein's to general relativity. However Hilbert himself never denied that all the credit for general relativity should go to Einstein.

action for the gravitational field. The calculations are harder than what we did. But in the end the field equations just pop out.

Equation (36) reveals something interesting. We see that in general the source of the gravitational field is not just energy density. It can involve energy flow. It can involve momentum density and it can even involve momentum flow.

Now as a rule the momentum components, and even the energy flow, but certainly the momentum flow and the momentum density are much smaller than the energy density. Why can we say that ? It has to do with the impact of the speed of light in the formulas. When we put back the speed of light (that is, when we use ordinary units), the energy density is always huge because it gets a factor c^2, like in $E = mc^2$.

On the other hand momentum is typically not huge because it is just mass times velocity. When you are in a non-relativistic situation, when velocity is slow, energy density is by far the biggest component. The other components of the energy-momentum tensor are much smaller. They are typically smaller by powers of v/c.

In other words, in a frame of reference where the sources of gravitational field are moving slowly, the only important thing in the right hand side of equation (36) is T^{00}, that is, the energy density ρ. It is also true that, in the same limit, the only important thing on the left hand side is the second derivative of G^{00}. So in non-relativistic limit these things match and reduce to Newton's equations.

But if you are outside the non-relativistic limit, in places where sources are moving rapidly, or even places where the sources are made up out of particles which are moving rapidly, even though the whole system may not be moving so much, other components of the energy-momentum tensor do generate gravitation.

Since we now know that, in a sense, gravitation is just geometry, we see that in relativistic situations, all the components of the energy-momentum tensor do participate in the generation of the curvature. It is not just energy – or mass – that causes it.

Let's now consider a special case. Just as in Maxwell's equations there are solutions which involve no sources, the same situation is true here. So let's consider the case with either no sources or when we are in a region of space where there are no sources. Then, on the right hand side of equation (36), we have zero.

It is just a little simplification, but the equation does become simpler in that case. Suppose that $T^{\mu\nu} = 0$. In other words

$$\mathcal{R}^{\mu\nu} = \frac{1}{2} \; g^{\mu\nu} \; \mathcal{R} \tag{37}$$

Let's calculate $\mathcal{R}$ by, on both sides, contracting μ and ν. Contraction is done on one upper and one lower index. So let's first lower the ν index on both sides (by multiplying by the appropriate version of the metric tensor). We get

$$\mathcal{R}^{\mu}_{\nu} = \frac{1}{2} \; g^{\mu}_{\nu} \; \mathcal{R} \tag{38}$$

Then contraction of the left hand side gives precisely $\mathcal{R}$. On the right hand side, g^μ_ν is the Kronecker delta. Its contraction yields 4. We arrive at $\mathcal{R} = 2\mathcal{R}$. Therefore $\mathcal{R} = 0$.

In the case of no sources, the curvature scalar is zero. Therefore in equation (36), where the right hand side is already zero, we can also drop the $\mathcal{R}$ term on the left hand side. We arrive at

$$\mathcal{R}^{\mu\nu} = 0 \tag{39}$$

Einstein equation becomes simpler. It is called the *vacuum case*. In this case the Ricci tensor is zero. Notice, as we have already said, that it doesn't imply that the Riemann curvature tensor be zero. The solutions to equation (39) are non trivial. They contain gravitational waves with no sources, just like there can be electromagnetic waves even in space without electric charges.

Another example is the Schwarzschild metric. It is roughly speaking analogous to a point mass. Outside the point mass there is no matter, nothing. So just like Newton's equations, everywhere outside the mass, are the same as for empty space, the Schwarzschild metric, outside the singularity, is a solution to the vacuum Einstein equation.

It is conceptually a simple check, although the calculations are a real nuisance. You start from the Schwarzschild metric, sit down and spend the rest of the day calculating a ream of Christoffel symbols, get to the curvature tensor, contract it to arrive at the Ricci tensor, $\mathcal{R}^{\mu\nu}$, and observe that it is equal to zero. You can also do it by feeding the Schwarzschild metric into Mathematica on your computer.

Metrics for which the Ricci tensor is zero are called *Ricci* *flat*. So anywhere except at the singularity – where everything is crazy and meaningless anyway, just like for Newton's point masses – the Schwarzschild metric is Ricci flat. It is not the same as flat.

So gravitational waves satisfy equation (36). Schwarzschild metric satisfies it, except at the singularity. And that is the basic fact about Einstein field equations.

Questions / answers session

Q. : What happens if we do the contraction to calculate $\mathcal{R}$ not on the simplified equation (38), but directly on the Einstein equation where we leave $T^{\mu\nu}$?

A. : Okay, let's do that. So we start from

$$\mathcal{R}_{\mu\nu} - \frac{1}{2}\, g_{\mu\nu}\, \mathcal{R} = 8\pi G\, T_{\mu\nu}$$

It doesn't matter if the indices are upstairs or downstairs. Then, after multiplying by the appropriate version of g to move one of the indices upstairs, we contract the two indices. That gives us

$$\mathcal{R} - 2\mathcal{R} = 8\pi G\, T^{\mu}_{\mu}$$

T^{μ}_{μ} is the same thing as $T_{\mu\nu}\, g^{\mu\nu}$. Let's call it T. This scalar T is by definition what you get when you contract the two indices of T^{μ}_{ν}. So we have

$$\mathcal{R} = -8\pi G\, T$$

Then we put this back into Einstein equation. After minor reorganisation we get

$$\mathcal{R}_{\mu\nu} = 8\pi G \ \left(T_{\mu\nu} - \frac{1}{2}\ g_{\mu\nu} T \right) \tag{40}$$

In other words, Einstein equation can be written like equation (40) equating the Ricci tensor to something. But on the right hand side you have to compensate by subtracting $\frac{1}{2}\ g_{\mu\nu} T$.

Q. : Does T have an interpretation ?

A. : Yes. It is called the trace of the energy-momentum tensor. It is 0 for electromagnetic radiations, that is, for massless particles like photons or gravitons. For particles with mass the trace of the energy-momentum tensor is not zero.

Q. : You explained the interpretation of the Riemann tensor. Can you explain in the same way how to interpret the Ricci tensor and the curvature scalar ?

A. : We saw indeed that the Riemann tensor has to do with going around a little bump and seeing how much a vector, which locally doesn't change in each local flat ("best") coordinates, has rotated when we come back. I don't know any particular physical significance or geometric significance to the Ricci tensor or the curvature scalar. Whatever it is, it is not very transparent. They are much simpler objects than the full Riemann tensor. They average over directions. We lose information when we go from Riemann to Ricci. We can have $\mathcal{R}_{\mu\nu} = 0$ while the Riemann curvature tensor is not zero. An example that we will explore a little bit is gravitational waves.

Gravitational waves are just like electromagnetic waves : they don't require any sources to exist. Of course in the real world you expect electromagnetic waves to be made by an antenna or something. But as solutions of Maxwell's equations, you can have electromagnetic waves that just propagate from infinity to infinity.

In the same way you can have gravitational waves which also have no sources. Those waves satisfy $\mathcal{R}_{\mu\nu} = 0$. But they are certainly not flat space. It will be somewhat satisfying to meet a geometry that is Ricci flat but whose curvature tensor itself is not equal to 0.

This is of course also the case of the Schwarzschild metric as long as you stay away from the singularity.

There is something in such space-time. There is real curvature, tidal forces, all sorts of stuff. But $\mathcal{R}_{\mu\nu} = 0$. So $\mathcal{R}_{\mu\nu}$ has less information in it than the curvature tensor.

Notice that it actually depends on the dimensionality. In four dimensions there is less information in the Ricci tensor than in the Riemann tensor. In three dimensions, it turns out that the amount of information is the same. You can write one in terms of the other. And in two dimensions, all of the information is in the scalar. That is all there is. There is the scalar, and from it you can make the other things.

Q. : The information which gets lost when we go from the Riemann tensor to the Ricci tensor does not affect the energy-momentum tensor nor Einstein equations. What is the meaning of this lost information then ?

A. : It means that for a given source configuration, there can be many solutions to Einstein equations. They all have the same right hand side, namely $T_{\mu\nu}$. But they simply have different physical properties. So for example the simplest case is to ask : what if this energy-momentum stuff is 0 ?

If it is zero, does it mean that there is no gravitation, no interesting geometry at all ? No. It allows gravitational waves.

Furthermore, select an energy momentum tensor and construct a solution. Then you can add gravitational waves on top of it. It is not exactly true, but it is roughly true that to any solution you can always add gravitational waves. So the gravitational waves must be something that contains more information than just the Ricci tensor. And they do.

Q. : Shouldn't Einstein equations contain a cosmological constant ?

A. : We haven't said anything about the cosmological constant – whether it exists or not – because it can be thought of as part of $T_{\mu\nu}$.

From this point of view, the "cosmological constant" is an extra tensor term on the right hand side of equation (36). We could denote it

$$T_{\mu\nu}^{cosmological}$$

And that would not change the look of Einstein equation.

If it was indeed a scalar, we could write it

$$T_{\mu\nu}^{cosmological} = \Lambda \ g_{\mu\nu}$$

Then it would make sense to shift it on the left hand side, because it would only be a part of the geometry. This yields the following equation frequently met in the litterature

$$\mathcal{R}^{\mu\nu} - \frac{1}{2} \ g^{\mu\nu} \ \mathcal{R} + \Lambda \ g_{\mu\nu} = 8\pi G \ T^{\mu\nu}$$

Here is a brief history of this added term. When Einstein wrote his equation, the common view among cosmologists was that the universe was the Milky Way, and that it was stable. But Einstein soon realized that his equation implied that the universe could not be stable. So he added an extra term to counter the instability. In 1929, Hubble discovered that the universe was expanding. Then the cosmological became irrelevant. Later in his life Einstein called it "the biggest blunder of my career".

In fact, to this day, the cosmological constant is the subject of a debate because it can frequently be called to the rescue as a "explanatory factor" in various riddles of the cosmos. However most theoretical physicists dislike ad hoc factors or procedures to solve (or fudge) a problem. Yet it is sometimes fruitful if one thinks of the introduction of quantas by Max Planck as an ad hoc procedure to overcome a problem with black-body radiations.

Q. : Did Einstein use the equation you derived in this lesson to explain the precession of the perihelion of the orbit of Mercury ? And do you know of a source where we can see how this calculation was done ?

A. : It is certainly in one of Einstein's papers.

Let's talk about the two great successes that general relativity quickly achieved : the prediction of the bending of light-ray from distant stars by the Sun, and the explanation of the oddities in the orbit of Mercury.

Concerning the bending of light-rays, he did not have the Schwarzschild solution. But he did have the approximation to the Schwarzschild solution at fairly large distances. Of course the Sun is not a black hole. But outside the solar Schwarzschild radius the geometry is exactly the same as that of Schwarzschild.

Far away from the Schwarzschild radius, Einstein knew how to make a good approximation. The fact that the Sun is so big means that the corrections from Newton are small. They can be done using the techniques of perturbation theory : you add some small corrections to something you already know, and fit them so that the equations still hold.

For a light-ray you start from a nice straight line, do a little bit of perturbation theory and work out the modification of the trajectory.

For the orbit of Mercury, most likely what he did was just start from the Keplerian orbits. That is, he started from Newton solution. Then he fitted the small correction on the left hand side to the small correction on the right hand side. And he found exactly the right discrepancy in the perihelion of Mercury.

But then Schwarzschild, within a few weeks after Einstein first publication, calculated his exact solution of the equation. And from them you could do much better than perturbation theory.

Notice one last point concerning the orbits of planets. The ordinary gravitational potential in $1/r$, leading to a centripetal force law in $1/r^2$, is pretty much the only law that leads, in the simplest case, to elliptical closed orbits without precession. It is a curious and somewhat accidental fact [11].

11. Notice, though, that "accidental facts" are precisely the kind of facts that Einstein did not take for granted. It is the rejection of the accidental character of the equality between inertial mass and gravitational mass which ultimately lead him to the theory of general relativity.