1 Introduction to general relativity

1.1 Admin

Recommended reading is Foster & Nightingale 'A Short Course in General Relativity', Springer. I also use Schutz 'A first course in General Relativity', CUP, but I also put all my lecture notes on the web. They are available in DUO or directly at http://duss0.dur.ac.uk/~done/foundations.html - and the second section of the course is at http://duss0.dur.ac.uk/~done/tests.html. For general reading (just for fun) try the Uncle Albert series by Russell Stannard, and Brian Greene 'The Elegant Universe'. These give you some nice ways to think about GR.

1.2 Special Relativity

Einsteins special relativity was a towering intellectual achievement, disposing of the 'intutively obvious' ideas of space and time as separate entities. Time does not flow evenly, the same for everyone, and spatial distances are not identical for all observers. The absolute, invariant thing is not space or time but the spacetime interval

$$ds^2 = c^2 d\tau^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = c^2 dt^2 - dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2 = g_{\mu\nu} dx^\mu dx^\nu$$

1.3 Acceleration

Special relativity is fabuolous, but has a pretty severe limitation. It only describes behaviour in INERTIAL FRAMES. So how do we treat acceleration, and describe behaviour in accelerated frames ?

Lets think about some accelerated frames. Circular motion is a nice way to get an accelerated frame since the speed stays constant. So think of a spinning roundabout. One person crawls around ON THE GROUND with a ruler, and lays it end to end to measure the circumference. Another person crawls across to measure the diameter. The ratio circumference/diameter is then simply π.

But what does someone ON the roundabout measure ? If the ruler is small enough then the ruler lies along the direction of motion, so its length is shortened. so the person on the roundabout needs more rulers to get around the circumference than the person who crawled around the outside circumference who was at aproximately rest with respect to us. (yes, there are some sublties, but see the notes at the back of The Elegant Universe). But along the radius the ruler is perpendicular to the direction of motion so there is no length contraction so the diameter comes out to be the same as before. So the ratio circumference/diameter is BIGGER than π. There is NO WAY to do this in flat space! If we go onto the surface of a sphere (the surface falls away from us in both directions - positively curved space) then the ratio circumference/diameter is SMALLER than π (think of the limit of a circle of radius equal to that of the sphere - the circumference is $2\pi r$ but the diameter is πr so the ratio is 2!). If we go to the inside of a sphere we are still in positively curved space - it curves towards us in two directions. But we CAN do it in negatively curved space - like a saddle shape where it curves away from us in one direction and towards us the other.

So, if we are to handle acclerating frames we MUST use curved space. And since we know from special relativity that time is also a dimension we have to handle curved spacetime. Nasty nasty nasty. So we have to do some fairly advanced maths in order to figure out path lengths in curved space - its called differential geometry and its a bit grim.

1.4 Acceleration=gravity: Principle of equivalence

So do we REALLY want to handle accelerated frames - can't we just stick to inertial frames ? Well, no, because we want to understand gravity. And Newtons gravity is in big trouble as it has a subtle but direct conflict with SR. Newtons theory of gravity is amazingly accurate in calculating the effects of gravity in the world around us, but completely silent on what causes gravity. In Newtons formulae, gravity is solely determined by the mass and separation of objects. Changing the mass or separation then, according to Newton, INSTANTLY changes the gravity. But special relativity says NOTHING can travel faster than light. Einstens 'happiest thought' was the realisation that there was no difference between gravity and an accelerating frame - there is no difference between an accelerated frame without a gravitational field and a stationary frame within a gravitational field. Or, say the same thing a different way around, there is no difference between a free fall frame in a gravitational field and an inertial frame.

This is the equivalence principle. *In a freely falling (nonrotating) frame then the laws of physics are the laws of special relativity*

So there is a deep connection between gravity and accelerated motion. This explains why inertial and gravitational masses are the same, which was a big deep problem. Inertial mass, m_i, is the one which gives the constant of proportionality between force and acceleration - Newtons famous $F = m_i a$. Gravitational mass, m_g, determines how the object is affected by the force of gravity $F = GMm_g/r^2 = m_g g$. So when an object falls under a gravitational field we have $a = m_g/m_i g$. The only way for all objects to fall the same way in a gravitational field (feather and hammer) is if $m_g/m_i =$ constant. If acceleration and gravity are two different things then this is deeply strange. Compare it with electrostatics - a test particle of charge q and inertial mass m_i in a field E experiences a force qE and acceleration $m_i a$. So $a = q/m_i E$ and different particles have difference values of q/m_i. But they don't have different values of m_g/m_i! But if gravity is the same as being accelerated then its not at all strange. But acceleration REQUIRED us to think about curved spacetime, so we need to do differential geometry in order to understand gravity!!!

1.5 Curved space

Think of ants with ink on their feet, walking normally i.e. taking steps of equal length with their left and right feet. On a flat piece of their inky feet leave a path which is a straight line. On a curved surface, still taking steps of equal length, they trace out a curved path. e.g. two ants setting off from different points along the equator of a sphere, heading due south. Their paths get closer and closer together AS IF there were an attractive force between them. But there is not, its just that they are travelling normally (steps of equal length with left and right feet) over a surface that is curved. Now put the ants on the surface of an apple. they still move 'normally' but now their paths show a variety of curves depending on the curvature of the surface of the apple - if the ants walk close to the stalk where the surface is highly curved then their path curves a lot.

These 'walking normally' paths in physical language are constant velocity frames - i.e. inertial frames. But in geometric language they are geodesic paths - shortest distance between two points.

1) locally, geodesics (shortest distance between two points) appear straight

2) over more extended regions of spacetime then geodesics originally receeding from each other begin to approach at a rate governed by the curvature of spacetime, and this effect of geometry on matter is what we mean today by the old word gravitation
3) matter in turn warps geometry

Space tells matter how to move, matter tells space how to curve

1.6 The way ahead

1) figure out how mass(energy) curves spacetime(!!!)
2) find the geodesic paths in our curved spacetime - ie the inertial frames. we know how to do physics here (Special Relativity)
3) write the physical laws in such a way that we can easily transform the equations to any arbitrary frame (not just inertial ones). In these frames the paths will CURVE, AS IF THERE WAS SOME ADDITIONAL FORCE ACTING. But its because we have transformed to a 'bad' coordinate system, where the curvature of spacetime is apparent. This is where tensors come in.

1.7 Gravity and light: The Pounds-Rebka-Snider experiment

This was a Gedanken (thought) experiment first done by Einstein. In the old way to think of gravity, then only particles with mass are acted on by gravity, so photons, being massless, are not. Which leads to some horrors like infinite energy!

Think of a particle dropped from a tower of height h. Its energy when it reaches the ground is $m_0c^2 + m_0gh$. Convert it into a photon then it has $h\nu_{bottom} = m_0c^2 + m_0gh$. If gravity doesn't affect light then it gets to the top with this energy so $h\nu_{top} = m_0c^2 + m_0gh$, to be converted back into a particle. But now the particle mass is $m = m_0 + m_0gh/c^2$. So we can drop it again..... Since we believe in conservation of energy then we'd better have gravity affecting light, giving a gravitational redshift so that $h\nu_{top} = m_0c^2$. Hence

$$\frac{h\nu_{top}}{h\nu_{bottom}} = \frac{m_0c^2}{m_0c^2 + m_0gh} = 1 - gh/c^2 + ...$$

And in fact we can see this from the roundabout again - fast clocks run slow in SR so a clock in an accelerating frame run slow, ie time dilation=gravitational redshift

And we can tie in the equivalence principle as well. Suppose we had our energy to mass converter in a lab which was freely falling such that it arrived at the top of the tower from above just as the photon arrived at the top of the tower from below. The lab is then moving with velocity gh downwards, so the gravitationally redshifted photon is doppler blueshifted with respect to the lab. And the effects cancel (to first order as appropriate for these approximations) so *there is no redshift measured in the free fall frame.* So a free fall frame really is an inertial frame!

2 Mathematical Background

2.1 Introduction

So, we have to do curved spacetime. There are some very powerful mathematical techniques for handling arbitrary curved spaces. They are nasty, but we want to understand gravity, so we have to persevere. This gets us into differential geometry, of which we will do AS MUCH AS WE NEED (no more, no less) to do gravity. We want to be able to characterise the curvature of spacetime which is associated with the presence of mass-energy.

Another separate thing we need is to be able to write equations so that we can transform easily between frames (geodesic paths in curved spacetime are inertial frames. We can write down the equations of physics ie the equations of special relativity in these inertial frames, then transform coordinates out of these into an arbitrary frame and we automatically account for gravity!). Turns out we can do BOTH curvature and coordinate transformations of equations by using TENSORS.

2.2 Notation

First lets start with standard vectors $\vec{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ in standard cartesian coordinates. This notation is incredibly clumsy. So we're going to use instead a much more powerful and very compact notation. Axes are labeled by indices so $\{x, y, z\} = \{x^1, x^2, x^3\}$ where the superscripts DO NOT MEAN 'raised to the power' but instead pick out the components along the x, y and

z axes i.e. along the **i,j,k** directions which we are now going to call $\vec{e_1}, \vec{e_2}, \vec{e_3}$. If we need to raise to the power we will make it obvious e.g $(x^1)^3$. Then we can write this as

$$\vec{r} = x^1\vec{e_1} + x^2\vec{e_2} + x^3\vec{e_3}$$

It doesn't look much better yet. But we are going to get very sick of writing all the coordinates so we're going to use shorthand notation that $\{x^i\}$ means all of the $\{x^1, x^2, x^3\}$. We're also going to use the Einstein summation convection: whereever an expression contains one index as a superscript and the SAME index as a subscript then the summation is implied. So our vector

$$\vec{r} = x^1\vec{e_1} + x^2\vec{e_2} + x^3\vec{e_3} = x^i\vec{e_i} = x^j\vec{e_j}$$

It really doesn't matter what we call the index we sum over - if it occurs in the sum, we sum over it, whatever it is! But we'll have a convention that roman letters from the middle of the alphabet go from 1-3 (3D space, i,j...), while roman letters from the start of the alphabet (a,b,c...) mean $\{x^1, x^2..x^N\}$ - any N dimensional space and capital roman letters from the start of the alphabet (A,B....) mean 2D space. Greek letters are for spacetime, and go from 0-3, so $\{x^0, x^1, x^2, x^3\}$ where $x^0 = ct$, and $\{x^i\} = \{x, y, z\}$.

Once we have picked a given space(time) then we might do coordinate transformations within this. Up till now you've often seen this denoted as the primed coordinate frame, with the prime on the kernal letter, e.g. $x' = \gamma(x - vct/c)$ as the Lorentz transformation for the x coordinate I'm going to put a bar on the *index* to denote a transformation to a different frame (a prime can all to easily get mistaken for a 1, and I'm going to do it on the individual indices rather than on the kernal letter as we're going to generalise to things with more than 1 index...!!) so it looks like $x^{\bar{1}} = \gamma(x^1 - vx^0/c)$.

2.3 Generalised coordinates in N dimensions

Lets now think about more general spaces which have arbitrary curvature. Define a point P in some space, and another point a little further on called R. These points have NOTHING to do with the coordinate system. They exist irrespective of the labels we give them. Coordinates just make them easier to handle. So lets take two points, P and R, and look at the displacement vector which points from one event to another whose components are the coordinate differences. A key property of tensors is that their representations in different coordinate systems depend only on the relative orientations and scales on the coordinate axes at that point and NOT on the absolute values of coordinates. So thats why we used the differential position vector not the absolute position vector from the origin to a point. This VECTOR, connecting them, $\vec{r}$, will be the same in all coordinate systems, though its components will be different depending on our choice of coordinates.

$$\vec{r} = x^1\vec{e_1} + x^2\vec{e_2} + ... = x^a\vec{e_a}$$

is an arrow pointing from P to R, with components x^a along whatever basis vectors $\vec{e_a} = \partial\vec{r}/\partial x^a$ we are using (they don't have to be rectilinear or orthoganal).

Now transform to a different coordinate system - the vector is the SAME vector. We can write the new (primed) coordinates, as function of the old (unprimed) ones so $x^{\overline{a}} = x^{\overline{a}}(x^b)$. (our conventions mean that this expands to N equations):

$$x^{\overline{1}} = x^{\overline{1}}(x^1, x^2....x^N)$$
$$x^{\overline{2}} = x^{\overline{2}}(x^1, x^2...x^N)$$

etc... But we know how to do total differentials:

$$\Delta f = \frac{\partial f}{\partial x^1}\Delta x^1 + \frac{\partial f}{\partial x^2}\Delta x^2 + ...\frac{\partial f}{\partial x^N}\Delta x^N = \frac{\partial f}{\partial x^a}\Delta x^a$$

so we can write our coordinate transformations as (the N equations!)

$$\Delta x^{\overline{b}} = \frac{\partial x^{\overline{b}}}{\partial x^a}\Delta x^a$$

where there are N^2 separate $\partial x^{\overline{b}}/\partial x^a$. In the limit of infinitesimals this goes to

$$dx^{\overline{b}} = \frac{\partial x^{\overline{b}}}{\partial x^a}dx^a$$

A superscript index on the denominator counts the same in the summation convention as a subscript on the numerator.

2.4 Example: Spherical polar coordinates

we KNOW how to do coordinate transformations. eg for any point in a 3D space we can write

$$\vec{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$$

where **i,j,k** are simply the basis vectors along the x,y,z axes. we can transform to spherical polar coordinates where $x = r\sin\theta\cos\phi$ and $y = r\sin\theta\sin\phi$ and $z = r\cos\theta$. This is the old coordinates x,y,z i.e. x^i written as functions of the new i.e. r, θ, ϕ i.e. $x^{\bar{j}}$. then the new basis vectors along the new coordinate directions are

$$\vec{e}_{x^{\bar{1}}} = \vec{e}_r = \frac{\partial\vec{r}}{\partial r} = \sin\theta\cos\phi\ \mathbf{i} + \sin\theta\sin\phi\ \mathbf{j} + \cos\theta\ \mathbf{k}$$

$$\vec{e}_{x^{\bar{2}}} = \vec{e}_\theta = \frac{\partial\vec{r}}{\partial\theta} = r\cos\theta\cos\phi\ \mathbf{i} + r\cos\theta\sin\phi\ \mathbf{j} - r\sin\theta\ \mathbf{k}$$

$$\vec{e}_{x^{\bar{3}}} = \vec{e}_\phi = \frac{\partial\vec{r}}{\partial\phi} = -r\sin\theta\sin\phi\ \mathbf{i} + r\sin\theta\cos\phi\ \mathbf{j}$$

2.5 Contravariant tensors of 1st order (4 vectors)

Entities which transform like the coordinate differences are called contravariant tensors of first order, or (1,0) tensors. They are defined by their transformation properties. If $\vec{A} = A^a\vec{e_a}$ has components which transform as

$$A^{\bar{b}} = \frac{\partial x^{\bar{b}}}{\partial x^a}A^a$$

then it is a contravariant tensor - things like 4-velocity, momentum, 4-force etc all transform like this.

2.6 What about the basis vectors?

Back to the position vector

$$\vec{r} = x^a\vec{e}_a = x^{\bar{b}}\vec{e_{\bar{b}}}$$

Old basis vectors are $\vec{e}_a = \partial\vec{r}/\partial x^a$. New ones are again the tangents to the coordinate curves so

$$e_{\bar{b}} = \frac{\partial\vec{r}}{\partial x^{\bar{b}}} = \frac{\partial\vec{r}}{\partial x^a}\frac{\partial x^a}{\partial x^{\bar{b}}} = \frac{\partial x^a}{\partial x^{\bar{b}}}\vec{e_a}$$

where the second step is simply the chain rule. This is a bit different. Our basis vectors don't transform like the coordinate differences i.e. like contravarient tensors which have $A^{\overline{b}} = \partial x^{\overline{b}} / \partial x^a A^a$ - they go the other way (derivative of the old with respect to the new, whereas the components of the vectors - contravarient tensors of first order - transformed as the derivatives of the new with respect to the old). So this looks more like a INVERSE transformation if we were talking about coordinates. Lets look at this in some more detail by looking at the inverse transformations (from primed to unprimed frame). They are done in much the same way as the transformations we wrote before from the unprimed to primed frames.

$$\Delta x^a = \Delta x^a(\Delta x^{\overline{b}})$$
$$\Delta x^a = \frac{\partial x^a}{\partial x^{\overline{b}}} \Delta x^{\overline{b}}$$

or in the infinitesimal limit

$$dx^a = \frac{\partial x^a}{\partial x^{\overline{b}}} dx^{\overline{b}}$$

and yes, we can see that the change in basis vectors from the unprimed to primed frame goes like the coordinate differences from the primed to unprimed frame! So they don't transform like the coordinate differences, they transfrom like the INVERSE of the coordinate differences.

Lets see this from another viewpoint. We can write any contravarient 1st order tensor $\vec{A}$ as

$$\vec{A} = A^a \vec{e_a} = A^{\overline{b}} \vec{e_{\overline{b}}}$$

But we know how contravarient tensor components transform - its $A^{\overline{b}} = \partial x^{\overline{b}} / \partial x^a A^a$. But we can just rename indices and equally well write this as $A^a = \partial x^a / \partial x^{\overline{c}} A^{\overline{c}}$ so

$$A^{\overline{c}} \frac{\partial x^a}{\partial x^{\overline{c}}} \vec{e_a} = A^{\overline{b}} \vec{e_{\overline{b}}}$$

change dummy indices to whatever you like

$$A^{\overline{b}} \frac{\partial x^a}{\partial x^{\overline{b}}} \vec{e_a} = A^{\overline{b}} \vec{e_{\overline{b}}}$$

$$A^{\overline{b}} (\frac{\partial x^a}{\partial x^{\overline{b}}} \vec{e_a} - \vec{e_{\overline{b}}}) = 0$$

$$\vec{e_{\overline{b}}} = \frac{\partial x^a}{\partial x^{\overline{b}}}\vec{e_a}$$

and we get the same answer (phew!). So this is the way the basis vectors themselves change - its is NOT a component transformation. It gives the new *basis vectors* (not components) $\vec{e_{\overline{b}}}$ in terms of the partial differentials of the old coordinates with respect to the new, whereas the contravariant tensors we looked at before transformed as the partial differentials of the new coordinates wrt to the old.

2.7 More on contravariant tensors

$\vec{A}$ and $\vec{B}$ are both first order contravariant (1,0) tensors. $\vec{C} = a\vec{A}+b\vec{B}$ (where a and b are constants). Is $\vec{C}$ also first order contravariant tensor? To answer this we must find out how its components transform. ie how $C^{\overline{d}}$ relate to C^e.

$$C^{\overline{d}} = aA^{\overline{d}} + bB^{\overline{d}} = a\frac{\partial x^{\overline{d}}}{\partial x^e}A^e + b\frac{\partial x^{\overline{d}}}{\partial x^e}B^e = \frac{\partial x^{\overline{d}}}{\partial x^e}(aA^e + bB^e) = \frac{\partial x^{\overline{d}}}{\partial x^e}C^e$$

So the components $C^{\overline{e}}$ transform like a (1,0) tensor, so it IS a (1,0) tensor! And we see that tensors are linear.

Now lets think about other things we can do. What is $T^{ab} = A^aB^b$?

$$T^{\overline{cd}} = A^{\overline{c}}B^{\overline{d}} = \frac{\partial x^{\overline{c}}}{\partial x^e}A^e\frac{\partial x^{\overline{d}}}{\partial x^f}B^f = \frac{\partial x^{\overline{c}}}{\partial x^e}\frac{\partial x^{\overline{d}}}{\partial x^f}A^eB^f = \frac{\partial x^{\overline{c}}}{\partial x^e}\frac{\partial x^{\overline{d}}}{\partial x^f}T^{ef}$$

This transforms like 2 powers of the coordinate differentials so we call it a second order covariant tensor or (2,0). Getting higher order tensors from lower order ones in this way is called an outer product.

2.8 Covariant tensors of first order

Quantities which transform as the inverse of the coordinate differences are not just limited to the basis vectors. If we go back into 3D and look at how the components of a scalar field $\phi(x^i)$ transform. Scalar fields are just numbers at a given point. whatever coordinates you give the point makes no change in the number. But its gradient will change depending on coordinate system (because we take gradients as change in number over some coordiante distance). Its gradient is simply $\nabla\phi(x^i)$ which has components

$$\frac{\partial \phi}{\partial x^i}$$

so go to another frame with primed coordindates and look at $\nabla\phi(x^{\bar{j}})$ which has components

$$\frac{\partial \phi}{\partial x^{\bar{j}}} = \frac{\partial \phi}{\partial x^i}\frac{\partial x^i}{\partial x^{\bar{j}}}$$

so the gradient is NOT a contravariant tensor. So there are physical quantities we will need to describe which are NOT contravarient. Gradients have components which transform the other way to the way a contravariant tensor transforms. Goes as old wrt new not new wrt old i.e. like the normal basis vectors. We call things that transform like this COVARIANT vectors (also called (0,1) tensors or covariant tensors of the first kind/order, or one forms) and we denote these by a lower index on the components (like the normal basis vectors are denoted by a lower index), and transformation laws for the components (now in generalised N dimensions) are:

$$A_{\bar{b}} = \frac{\partial x^a}{\partial x^{\bar{b}}} A_a$$

2.9 Kronekar delta

We saw in the last lecture that the difference in position vector between two points (something which MUST be the same in any frame) $\vec{r} = \Delta x^a \vec{e}_a = \Delta x^{\bar{b}} \vec{e}_{\bar{b}}$ The coordinate differences transformed as $\Delta x^{\bar{b}} = \partial x^{\bar{b}} / \partial x^a \Delta x^a$. Substitute this into the equation of $\vec{r}$ and get that the basis vectors transform as $\vec{e}_{\bar{b}} = \partial x^a / \partial x^{\bar{b}} \vec{e}_a$. This means that any quantity $\vec{A} = A^a \vec{e}_a$ in another frame,

$$A^{\bar{b}} \vec{e}_{\bar{b}} = \frac{\partial x^{\bar{b}}}{\partial x^a} A^a \frac{\partial x^d}{\partial x^{\bar{b}}} \vec{e}_d$$

where we changed a to d in the implied sum for the basis vectors as we've already used a - we can call them anything we like in a sum, but they are BOTH sums. So finally we get

$$\frac{\partial x^d}{\partial x^a} A^a \vec{e}_d = \delta^d_a A^a \vec{e}_d = A^a \vec{e}_a$$

where the $\delta^a_b = 1$ for $a = b$ and 0 otherwise is termed the Kronekar delta function. So if we have something that transforms as the coordinate differences, then this means its tensor equation looks the same in ANY frame!!!

A quick way to see if you have dropped any indices is to count them - cancel out the indices which have an impled sum, and then the indices we have on the left hand side MUST be equal to the indices we have on the right.

2.10 Metric tensor

Lets go back to our generalised differential position vector again

$$\vec{\Delta x} = \Delta x^a \vec{e_a}$$

This is telling us the distance between two points. Suppose in fact we're looking at a curve in this N dimensional space, where position along the curve can be parameterized as some parameter t

$$\vec{r} = x^a(t)\vec{e_a}$$

$$\vec{\Delta r} = \frac{\partial \vec{r}}{\partial x^b}\frac{dx^b}{dt}\Delta t$$

$$\frac{\vec{dr}}{dt} = \frac{dx^b}{dt}\vec{e_b}$$

as $\partial\vec{r}/\partial x^b = \vec{e_b}$

So the length of the curve is

$$L = \int_{t1}^{t2} |\vec{dr}/dt.\vec{dr}/dt|^{1/2} dt$$

or in infinitesimals

$$ds^2 = \vec{dr}.\vec{dr}$$

$$= dx^a\vec{e_a}.dx^b\vec{e_b}$$

$$= \vec{e_a}.\vec{e_b}dx^a dx^b$$

$$= g_{ab}dx^a dx^b$$

where $g_{ab} = \vec{e_a}.\vec{e_b}$ is called the metric. Its transformation under coordinate change can be seen as we derived the basis vector transformations

$$\vec{e_{\overline{a}}}.\vec{e_{\overline{b}}} = \frac{\partial x^c}{\partial x^{\overline{a}}}\vec{e_c}.\frac{\partial x^d}{\partial x^{\overline{b}}}\vec{e_d}$$

$$= \frac{\partial x^c}{\partial x^{\overline{a}}}\frac{\partial x^d}{\partial x^{\overline{b}}}\vec{e_c}.\vec{e_d}$$

$$= \Lambda^c_{\overline{a}}\Lambda^d_{\overline{b}}g_{cd}$$

So the components transform like the basis vectors twice - called covariant tensor of second order - this is the METRIC tensor and its wildly important in everything that we are going to do next, because it includes what happens to distances (spacetime intervals!!) if the space is curved.

From its definition in terms of the dot product of basis vectors then we can see that $g_{ab} = g_{ba}$, ie that it is symmetric.

The metric tensor is basically a rule for taking two vectors and getting a single number out.

2.11 Example metric for the surface of a sphere

I've said that the machinary we've developed is very general - we can use it on spaces of arbitrary curvature as well as flat Euclidean spaces. So lets see it at work on a CURVED surface - the surface of a sphere. Here we are embedding a 2D curved surface into a 3D Euclidean geometry. So we have $\vec{r} = x\mathbf{i}+y\mathbf{j}+z\mathbf{k}=x^i\vec{e_i}$. Now transform coordinates to the surface of a sphere of radius a where a point can be parameterised in terms of θ and ϕ on the surface where $x = a\sin\theta\cos\phi$, $y = a\sin\theta\sin\phi$, $z = a\cos\theta$. so now we can write $\vec{r} = x^{\overline{A}}e_{\overline{A}}$ where I'm using the convention that capital roman letters mean a 2D space. So lets write the coordinate trasnformation as primed coordinates to make it all look like the stuff we've been dealing with so

$$x^1 = a\sin x^{\overline{1}}\cos x^{\overline{2}} \quad x^2 = a\sin x^{\overline{1}}\sin x^{\overline{2}} \quad x^3 = a\cos x^{\overline{1}}$$

the new basis vectors (the ones along the surface of the sphere) are

$$e_{\overline{A}} = \frac{\partial\vec{r}}{\partial x^{\overline{A}}} = \frac{\partial x^i}{\partial x^{\overline{A}}}\vec{e_i}$$

so for the $x^{\overline{1}}$ i.e. θ coordinate we have

$$e_{\overline{1}} = \frac{\partial x^1}{\partial x^{\overline{1}}}\vec{e_1} + \frac{\partial x^2}{\partial x^{\overline{1}}}\vec{e_2} + \frac{\partial x^3}{\partial x^{\overline{1}}}\vec{e_3} = a\cos x^{\overline{1}}\cos x^{\overline{2}}\vec{e_1} + a\cos x^{\overline{1}}\sin x^{\overline{2}}\vec{e_2} - a\sin x^{\overline{1}}\vec{e_3}$$

or more simply $\vec{e_\theta} = a\cos\theta\cos\phi\mathbf{i}+a\cos\theta\sin\phi\mathbf{j}-a\sin\theta\mathbf{k}$. Similarly

$$e_{\overline{2}} = \frac{\partial x^1}{\partial x^{\overline{2}}}\vec{e_1} + \frac{\partial x^2}{\partial x^{\overline{2}}}\vec{e_2} + \frac{\partial x^3}{\partial x^{\overline{2}}}\vec{e_3} = -a\sin x^{\overline{1}}\sin x^{\overline{2}}\vec{e_1} + a\sin x^{\overline{1}}\cos x^{\overline{2}}\vec{e_2}$$

or more simply $\vec{e_\phi} = -a\sin\theta\sin\phi\mathbf{i}+a\sin\theta\cos\phi\mathbf{j}$.

The metric tensor components are then the dot products of the basis vectors so

$$g_{\overline{11}} = \vec{e_{\overline{1}}}.\vec{e_{\overline{1}}} = a^2\cos^2 x^{\overline{1}}\cos^2 x^{\overline{2}} + a^2\cos^2 x^{\overline{1}}\sin^2 x^{\overline{2}} + a^2\sin^2 x^{\overline{1}} = a^2$$

$$g_{\overline{12}} = \vec{e_{\overline{1}}}.\vec{e_{\overline{2}}} = -a^2\cos x^{\overline{1}}\cos x^{\overline{2}}\sin x^{\overline{1}}\sin x^{\overline{2}} + a^2\cos x^{\overline{1}}\sin x^{\overline{2}}\sin x^{\overline{1}}\cos x^{\overline{2}} = 0$$

$$g_{\overline{22}} = \vec{e_{\overline{2}}}.\vec{e_{\overline{2}}} = a^2\sin^2 x^{\overline{1}}\sin^2 x^{\overline{2}} + a^2\sin^2 x^{\overline{1}}\cos^2 x^{\overline{2}} = a^2\sin^2 x^{\overline{1}}$$

So the line element showing us how distance relates to changing coordinates on the surface of a sphere is the familiar one

$$ds^2 = g_{\overline{AB}}dx^{\overline{A}}dx^{\overline{B}} = a^2(dx^{\overline{1}})^2 + a^2\sin^2 x^{\overline{2}}(dx^{\overline{2}})^2 = a^2d\theta^2 + r^2\sin^2\theta d\phi^2$$

Well, its familar if you can remember anything about spherical polar coordinates!

The surface of a sphere is a 2D space - position/distances can be completely defined by 2 coordinates. Other 2D metrics are the obviuos flat space metric $ds^2 = dx^2 + dy^2$. We could instead coose to write this in polar coordindates where $x = r\cos\theta$ $y = r\sin\theta$ which gives $ds^2 = dr^2 + r^2d\theta^2$. At first sight you might look at this and think that it represents a curved surface as distances are no longer doing a Pythagoros law - the metric tensor components are not given by a diagonal, unit matrix. BUT - there exists a coordinate transformation that gets us back to a metric in the form of $ds^2 = dx^2 + dy^2$ - the underlying space is FLAT. The difference for our surface of the sphere 2d surface is that there is no transformation we can make to get this metric back to $ds^2 = dx^2 + dy^2$. This surface of a sphere metric represents space which is intrinscally CURVED.

2.12 Basis vectors for covariant components

We got the basis vectors for the contravariant components by getting the vectors tangent to the coordinate curves. We've been working with these basis vectors which define a tangent space T_P at P. A contravariant vector $\vec{\lambda} = \lambda^a \vec{e_a}$ is an arrow emmanating from P - its not in the manifold like the curve is in the manifold, but is attached to it like the tangent space is attached to a surface.

But we could instead have defined a set of basis vectors by the vectors which were normal to the coordinate surfaces rather than tangent to the coordinate curves. We get a coordinate surface by holding that coordinate constant. and then we want to get the normal to this surface. We get normals by taking the GRADIENT. We can regard each coordinate as a scalar field Then the vector perpendicular to this scalar field is $\vec{e^a} = \nabla x^a$. And if we do a transformation of coordinates $x^{\overline{b}} = x^{\overline{b}}(x^a)$ then these basis components transform as

$$\vec{e^{\overline{b}}} = \nabla x^{\overline{b}} = \frac{\partial x^{\overline{b}}}{\partial x^a} \nabla x^a = \frac{\partial x^{\overline{b}}}{\partial x^a} \vec{e^a}$$

i.e. these transform in the opposite way to our tangent basis vectors.

This gives us another vector space of the manifold - call it T_P^* the dual or cotangent space of the manifold at P. The basis vectors themselves look identical if we have an orthonormal set of coordinates (they don't have to be reclilinear, just be 90 degrees where they meet). BUT THEY ARE NOT IDENTICAL IF THE COORDINATES ARE NOT ORTHOGANAL.

2.13 Covariant and contravariant: more on the metric

But if we have another set of basis vectors then we can write any arbitrary vector either on the old basis in the tangent space OR the new basis in the cotangent space i.e. $\vec{\lambda} = \lambda^a \vec{e_a} = \lambda_b \vec{e^b}$. If the basis vectors are the same i.e. we had orthonormal bases then the contravariant and covariant components are IDENTICAL. But of course in general they are not.

For the pure mathematically inclined, setting the contravariant and covariant forms of writing this as equal is a bit of a fudge as $\lambda^a \vec{e_a}$ and $\lambda_b \vec{e^b}$ are actually are in two different vector spaces T_P and $T*_P$, respectively. But we'll do a bit more on this later

Take another vector $\vec{\mu} = \mu^a \vec{e_a} = \mu_a \vec{e^a}$. The dot product of these is

$$\vec{\lambda}.\vec{\mu} = \lambda^a \vec{e_a}.\mu^b \vec{e_b} = \lambda^a \mu^b \vec{e_a}.\vec{e_b} = \lambda^a \mu^b g_{ab} = |\vec{A}||\vec{B}| \cos \chi$$

where *chi* is the angle between the two vectors. But we could have done this with the covariant components alone

$$\vec{\lambda}.\vec{\mu} = \lambda_a \vec{e^a}.\mu_b \vec{e^b} = \lambda_a \mu_b \vec{e^a}.\vec{e^b} = \lambda_a \mu_b g^{ab}$$

where we have defined $g^{ab} = \vec{e^a}.\vec{e^b}$. Again this is symmetric by definition of the dot product so $g^{ab} = g^{ba}$. But we could also have done this with the mixed components

$$\vec{\lambda}.\vec{\mu} = \lambda^a \vec{e_a}.\mu_b \vec{e^b} = \lambda^a \mu_b \vec{e_a}.\vec{e^b}$$

What is this dot product ? lets do it explicitally in 3D

$$\vec{e_i}.\vec{e^j} = \frac{\partial \vec{r}}{\partial x^i}.\nabla x^j$$

where $\vec{r} = x^i \vec{e_i}$. Then we have

$$\begin{aligned} e_i.e^j &= (\frac{\partial x^1}{\partial x^i} i + \frac{\partial x^2}{\partial x^i} j + \frac{\partial x^3}{\partial x^i} k).(\frac{\partial x^j}{\partial x^1} i + \frac{\partial x^j}{\partial x^2} j + \frac{\partial x^j}{\partial x^3} k) \\ &= \frac{\partial x^1}{\partial x^i}\frac{\partial x^j}{\partial x^1} + \frac{\partial x^2}{\partial x^i}\frac{\partial x^j}{\partial x^2} + \frac{\partial x^3}{\partial x^i}\frac{\partial x^j}{\partial x^3} \\ &= \frac{\partial x^j}{\partial x^i} = \delta_i^j \end{aligned}$$

where the kronekar delta function is $\delta_i^j = 1$ for $i = j$, 0 otherwise. NB because of the summation convention then $\delta_i^i = \delta_1^1 + \delta_2^2 + \delta_3^3 = 3$ So now we have

$$\lambda^a \vec{e_a}.\mu_b \vec{e^b} = \lambda^a \mu_b \vec{e_a}.\vec{e^b} = \lambda^a \mu_b \delta_a^b = \lambda^a \mu_a$$

By symmetry of the dot product and metric then $\lambda^a \mu_a = \lambda_a \mu^a$. So collecting all these together we have

$$\lambda^a \mu_a = \lambda_a \mu^a = g_{ab} \lambda^a \mu^b = g^{ab} \lambda_a \mu_b$$

So this gives us an easy way of swapping between contravariant and covariant components - we use the METRIC $g_{ab} \lambda^a \mu^b = \lambda_a \mu^a$ so

$$\lambda_a = g_{ab} \lambda^b$$

and as $g^{ab}\lambda_a\mu_b = \lambda^b\mu_b$ so

$$\lambda^a = g^{ab}\lambda_b$$

Tensors derived from other tensors by raising or lowering the indices via the metric are called **associated tensors**.

We also get another useful metric relation in that $\lambda_a = g_{ab}\lambda^b = g_{ab}g^{bc}\lambda_c$ so

$$g_{ab}g^{bc} = \delta^a_c$$

For the pure mathematically inclined, heres how its done. A contravariant vector $\vec{\lambda} = \lambda^a\vec{e_a}$ in T_p is associated with a covariant vector $\tilde{\lambda} = \lambda_a\vec{e^a}$ in $T*_P$ via the metric (or dot or inner product) such that for all vectors $\vec{\mu}$ in T_P then $\tilde{\lambda}(\vec{\mu})$ is a real number $= \mu^a\lambda_a$. i.e. it define covariant vectors as linear functions which map covariant vectors (one–forms) into real numbers via the metric. Schutz takes this approach (eventualy), as do Misner, Thorne and Wheeler. Foster and Nightingale take the approach I've used here.

2.14 Summary: Tensor definitions and algebra

Suppose that at a point P on a manifold M there are N^{r+s} quantities $\tau^{a_1...a_r}{}_{b_1...b_s}$ which under a change of coordinates transform as

$$\tau^{\overline{a_1}...\overline{a_r}}{}_{\overline{b_1}...\overline{b_s}} = \frac{\partial x^{\overline{a_1}}}{\partial x^{c_1}}...\frac{\partial x^{\overline{a_r}}}{\partial x^{c_r}}\frac{\partial x^{d_1}}{\partial x^{\overline{b_1}}}...\frac{\partial x^{d_s}}{\partial x^{\overline{b_s}}}\tau^{c_1...c_r}{}_{d_1...d_s}$$

where the partial derivatives of the coordinates are also evaluated at P. Then the quantities $\tau^{a_1...a_r}{}_{b_1...b_s}$ are called components of as type (r,s) tensor at P. The sum (r+s) is called the rank or order of the tensor, (r,0) might be called a contravariant tensor of rank r, while a (0,s) might be called a covariant tensor of rank s. If both r and s are nonzero then its called a mixed tensor. scalars are type (0,0) tensors. The whole point is that tensor equations maintain physical meaning under coordinate transformation.

linear combination

$$T^c_d = aA^c_d + bB^c_d$$

where a and b are scalars and A^c_d, B^c_d are (1,1) mixed tensors. Then T^c_d is also a (1,1) mixed tensor.

$$T^{\overline{e}}_{\overline{f}} = aA^{\overline{e}}_{\overline{f}} + bB^{\overline{e}}_{\overline{f}} = a\frac{\partial x^{\overline{e}}}{\partial x^c}\frac{\partial x^d}{\partial x^{\overline{f}}}A^c_d + b\frac{\partial x^{\overline{e}}}{\partial x^c}\frac{\partial x^d}{\partial x^{\overline{f}}}B^c_d$$

$$= \frac{\partial x^{\overline{e}}}{\partial x^c}\frac{\partial x^d}{\partial x^{\overline{f}}}T^c_d$$

Direct (outer) Products We can form higher order tensors out of lower order ones by multiplying the components.

$$T^a_b = A^a B_b$$

Again, its easy to show that this transforms as required.

$$T^{\overline{a}}_{\overline{b}} = A^{\overline{a}} B_{\overline{b}} = \frac{\partial x^{\overline{a}}}{\partial x^c}A^c \frac{\partial x^d}{\partial x^{\overline{b}}} B_d = \frac{\partial x^{\overline{a}}}{\partial x^c}\frac{\partial x^d}{\partial x^{\overline{b}}}T^c_d$$

General Inner product (contraction)

Set an upper and lower index as equal i.e. summing it, gives a new tensor of rank 2 less than the original one. Generally we keep the same kernal letter e.g. if we have a tensor $T^a{}_b{}^{cd}$ then $T^{ac} = T^a{}_d{}^{cd}$. Though obviously this is not unique as we could have also had the DIFFERENT contractions $T^{ad} = T^a{}_b{}^{bd}$ or $T^{cd} = T^a{}_a{}^{cd}$. Lets check that this is a valid tensor operation, and see how it transforms:

$$T^{\overline{a}}{}_{\overline{a}}{}^{\overline{cd}} = \frac{\partial x^{\overline{a}}}{\partial x^e}\frac{\partial x^f}{\partial x^{\overline{a}}}\frac{\partial x^{\overline{c}}}{\partial x^g}\frac{\partial x^{\overline{d}}}{\partial x^h}T^e{}_f{}^{gh} = \delta_e{}^f\frac{\partial x^{\overline{c}}}{\partial x^g}\frac{\partial x^{\overline{d}}}{\partial x^h}T^e{}_f{}^{gh} = \frac{\partial x^g}{\partial x^{\overline{c}}}\frac{\partial x^h}{\partial x^{\overline{d}}}T^e{}_e{}^{gh} = \frac{\partial x^g}{\partial x^{\overline{c}}}\frac{\partial x^h}{\partial x^{\overline{d}}}T^{gh}$$

Raising and lowering indexes

This is simply a direct product of a tensor with the metric, followed by a contraction. We can use it to swap between covariant and contravariant components AND basis vectors so $A_a = g_{ab}A^b$, $\vec{e^a} = g^{ab}\vec{e_b}$ where $g^{ab}g_{bc} = \delta^a_c$

Inner Product of 2 vectors (dot or scalar product)

$$\vec{A}.\vec{B} = g_{ab}A^a B^b = A^a B_a = A_b B^b = A^1 B_1 + A^2 B_2 + ...A^N B_N = |\vec{A}||\vec{B}|\cos\chi$$

where χ is the angle between the two vectors. Because its a scalar it must be invarient. We can see this explicitally on a change of coordinates

$$A^{\overline{a}} B_{\overline{a}} = \frac{\partial x^{\overline{a}}}{\partial x^c}A^c \frac{\partial x^d}{\partial x^{\overline{a}}} B_d$$

$$= \delta_c^d A^c B_d = A^c B_c$$

An obvious and VERY important invarient scalar product is $ds^2 = c^2 d\tau^2 = dx^\alpha dx_\alpha = g_{\alpha\beta} dx^\alpha dx^\beta$

ANY EQUATION WILL BE INVARIANT UNDER GENERAL COORDINATE TRANSFORMATIONS IF IT STATES THE EQUALITY OF TWO TENSORS WITH THE SAME UPPER AND LOWER INDICES. ${A^a}_{bc} = {B^a}_{bc}$ can be written in another coordinate frame as ${A^{\overline{a}}}_{\overline{bc}} = {B^{\overline{a}}}_{\overline{bc}}$ since both sides transform between coordinate frames in the same way

3 Curved Space

We're going to assume that we'll always be dealing with a space which is continuous and differentiable - a differentiable manifold. We're also going to assume that it has a metric. If the metric is positive definite (inner products +ve) then this is called a Riemannian manifold. Our metric won't be - so called pseudo–Riemannian. We've seen that the metric can carry (unimportant) information about coordinates - if space is flat then there is always some coordinate choice which gets it back into the unit matrix, and that the metric can also carry important non–transformable away information about the space curvature e.g. the metric of SR spacetime and the metric's we'll be using to describe the curved spacetime of gravity.

So the whole idea is that there is no 'force-at-a-distance' gravity. Paths curve only because they are following locally curved spacetime (which is curved because of mass). Any paths which follow the local curvature of spacetime (free-fall) are inertial frames, and we know how to do physics in inertial frame so we need to find out what these paths are.

3.1 Parallel transport

Some strange things happen when space is curved rather than flat. In flat space, we can take a vector from point and it keeps its direction - this is called parallel transport. This is important as when we do differentials we're comparing a vector at some point with a vector at another point. In curved space e.g. the surface of a sphere it just doesn't happen and the direction the vector points at the end of the path depends on the PATH as well as the start and end points. This is a really big problem when we start to look at differentiation.

Lets think some more about this, and have our vector $\vec{\lambda} = \lambda^a(u)\vec{e_a}$ defined along some curve given by coordinates $x^a(u)$. Suppose we tried to define derivitive in the 'obvious' way as

$$\frac{d\lambda^a}{du} = \frac{\lim}{\delta u \to 0}\frac{[\lambda^a(u+\delta u) - \lambda^a(u)]}{\delta u}$$

For a start this isn't a tensor as it doesn't transform.

$$\frac{d\lambda^{\overline{a}}}{du} = \frac{d(\partial x^{\overline{a}}/\partial x^b \lambda^b)}{du} = \frac{\partial x^{\overline{a}}}{\partial x^b}\frac{d\lambda^b}{du} + \lambda^b \frac{d}{du}\frac{\partial x^{\overline{a}}}{\partial x^b} = \frac{\partial x^{\overline{a}}}{\partial x^b}\frac{d\lambda^b}{du} + \frac{\partial^2 x^{\overline{a}}}{\partial x^b \partial x^c}\frac{dx^c}{du}\lambda^b$$

Theres a funny extra term, which comes from the fact that $\lambda^a(u+\delta u)$ and $\lambda^a(u)$ are located at different points. In general the coordiate transforms depend on position so $\partial x^{\overline{a}}/\partial x^b$ evaluated at $u+\delta u$ is not the same as $\partial x^{\overline{a}}/\partial x^b$ at u. For differentiation to give a vector (tensor) we must take component differences at the same point. In flat space we slide one of the vectors to the other one by moving one of the vectors parallel to itself (called parallel transport).

So how do we parallelly transport a vector ? We don't change its length or direction so

$$\frac{d\vec{\lambda}}{du} = 0 = \frac{d\lambda^a \vec{e_a}}{du} = \frac{d\lambda^a}{du}\vec{e_a} + \lambda^a \frac{d\vec{e_a}}{du} = \frac{d\lambda^a}{du}\vec{e_a} + \lambda^a \frac{\partial \vec{e_a}}{\partial x^b}\frac{dx^b}{du}$$

There is this nasty component $\partial \vec{e_a}/\partial x^b$ but since this is the derivitive of a vector then it is itself a vector so we can write it as a linear combination of basis vectors

$$\frac{\partial \vec{e_a}}{\partial x^b} = \Gamma^c{}_{ab}\vec{e_c}$$

these are called Christoffel symbols or connection coefficients. We have written them in tensor notation but THEY DO NOT TRANSFORM AS TENSORS IN GENERAL.

$$= \frac{d\lambda^a}{du}\vec{e_a} + \lambda^a \Gamma^c{}_{ab}\frac{dx^b}{du}\vec{e_c} = \frac{d\lambda^a}{du}\vec{e_a} + \lambda^b \Gamma^a{}_{bc}\frac{dx^c}{du}\vec{e_a}$$

so since this is all the same vector component we can see that

$$\frac{d\lambda^a}{du} + \lambda^b \Gamma^a{}_{bc}\frac{dx^c}{du} = 0$$

So if we parallelly transport a vector λ at u to $u + \delta u$ we have

$$\frac{\delta\lambda^a}{\delta u} + \lambda^b\Gamma^a{}_{bc}\frac{\delta x^c}{\delta u} = 0$$
$$\delta\lambda^a = -\lambda^b\Gamma^a{}_{bc}\frac{\delta x^c}{\delta u}\delta u$$
$$\delta\lambda^a = -\lambda^b\Gamma^a{}_{bc}\delta x^c$$

so the parallely transported vector has components $\lambda^a_{||}(u + \delta u) = \lambda^a(u) - \lambda^b\Gamma^a{}_{bc}\delta x^c$, and we can compare this with the vector $\vec{\lambda}(u+\delta u)$ and use this to define absolute derivitive. The point is that $\vec{\lambda}(u+\delta u)$ contains two differences with respect to $\vec{\lambda}(u)$ - firstly there can be real physical differences - maybe this contravariant tensor is telling us about velocity and maybe there has been some real physics going on - accelerations and decelerations, that mean that $\vec{\lambda}(u + \delta u)$ is different from $\vec{\lambda}(u)$. BUT THERE CAN ALSO BE CHANGES BETWEEN THESE TWO VECTORS JUST BECAUSE THE SPACE IS CURVED. And we want to take out these 'space changes' which are what the parallelly transported vector tells us about, so that we can see the real physical changes. So these are

$$\frac{\lambda^a(u + \delta u) - \lambda^a_{||}(u + \delta u)}{\delta u} = \frac{\lambda^a(u) + \frac{d\lambda^a}{du}\delta u - \lambda^a(u) + \lambda^b\Gamma^a{}_{bc}\delta x^c}{\delta u} = \frac{d\lambda^a}{du} + \lambda^b\Gamma^a{}_{bc}\frac{\delta x^c}{\delta u}$$

so

$$\frac{D\lambda^a}{du} = \frac{d\lambda^a}{du} + \lambda^b\Gamma^a{}_{bc}\frac{dx^c}{du}$$

We saw that $d\lambda^a/du$ is not a tensor, and I've told you that $\Gamma^a{}_{bc}$ isn't one either. But the two non-tensor bits concel each other out so $D\lambda^a/du$ DOES transform as a tensor.

So this is the way to define physical or absolute derivative - use parallel transport to take out ths SPACE changes. But what we are left with is PATH DEPENDANT in general.

3.2 Covariant derivative

We can do this in terms of coordinates as opposed to a single parameter, and remove that path dependance. We had

$$\frac{D\lambda^a}{du} = \frac{d\lambda^a}{du} + \lambda^b\Gamma^a{}_{bc}\frac{dx^c}{du} = \frac{\partial\lambda^a}{\partial x^c}\frac{dx^c}{du} + \lambda^b\Gamma^a{}_{bc}\frac{dx^c}{du} = \Big(\frac{\partial\lambda^a}{\partial x^c} + \lambda^b\Gamma^a{}_{bc}\Big)\frac{dx^c}{du}$$

So the components of the covariant derivative, are

$$\lambda^a{}_{;c} = \frac{\partial \lambda^a}{\partial x^c} + \Gamma^a{}_{bc}\lambda^b$$

This is obviously closely related to the absolute derivative, which is a (0,1) tensor, so this is a (1,1) tensor

3.3 Christoffel symbols and the metric

What are these horrors ? well they come from the derivitives of the tangent basis vectors, but these themselves are defined as $\vec{e_a} = \partial\vec{r}/\partial x^a$ so

$$\Gamma^c{}_{ab}\vec{e_c} = \frac{\partial \vec{e_a}}{\partial x^b} = \frac{\partial^2 \vec{r}}{\partial x^a \partial x^b} = \frac{\partial^2 \vec{r}}{\partial x^b \partial x^a} = \frac{\partial \vec{e_b}}{\partial x^a} = \Gamma^c{}_{ba}\vec{e_c}$$

so they are symmetric (changing the order of integration assumes that the second partial derivatives are continuous). To get some more about them the next step is to look at how the metric changes as a function of the coordinates.

$$\frac{\partial g_{ab}}{\partial x^c} = \frac{\partial \vec{e_a}.\vec{e_b}}{\partial x^c} = \frac{\partial \vec{e_a}}{\partial x^c}.\vec{e_b} + \vec{e_a}.\frac{\partial \vec{e_b}}{\partial x^c} = \Gamma^d{}_{ac}\vec{e_d}.\vec{e_b} + \Gamma^f{}_{bc}\vec{e_a}.\vec{e_f} = \Gamma^d{}_{ac}g_{db} + \Gamma^f{}_{bc}g_{af}$$

relabel to get:

$$\frac{\partial g_{ab}}{\partial x^c} = \Gamma^d{}_{ac}g_{db} + \Gamma^d{}_{bc}g_{ad}$$

then move all suffixes around to get

$$\frac{\partial g_{bc}}{\partial x^a} = \Gamma^d_{ba}g_{dc} + \Gamma^d_{ca}g_{bd}$$

$$\frac{\partial g_{ca}}{\partial x^b} = \Gamma^d_{cb}g_{da} + \Gamma^d_{ab}g_{cd}$$

add the first two and subtract the third to get an expression for the Christoffel symbols in terms of the derivities of the metric - SO IF THE METRIC IS CONSTANT THEN THE CHRISTOFFEL SYMBOLS ARE ZERO. So

$$2\Gamma^d_{ca}g_{db} = \frac{\partial g_{ab}}{\partial x^c} + \frac{\partial g_{bc}}{\partial x^a} - \frac{\partial g_{ca}}{\partial x^b}$$

we can make this *slightly* simpler by multipling by $\frac{1}{2}g^{fb}$ to get

$$\Gamma^f_{ca} = \frac{1}{2}g^{fb}(\frac{\partial g_{ab}}{\partial x^c} + \frac{\partial g_{bc}}{\partial x^a} - \frac{\partial g_{ca}}{\partial x^b})$$

So we can calculate these horrors!!

3.4 notes on notation!

partial derivative $\partial A^a/\partial x^c = \partial_c A^a = A^a{}_{,c}$ (I won't use the latter as its too easy to lose a comma!)
covariant derivative $\nabla_c A^a = A^a{}_{;c}$ (I will use the semicolon notation because physicists tend to freeze when they see del!!)
absolute derivative $\nabla_{\vec{u}} A^a = DA^a/du$
Christoffel symbols (or connection coefficients) $\Gamma^a{}_{bc}$ or $\left\{ {a \atop bc} \right\}$ or $\{a, bc\}$

3.5 Example: flat space

The metric for flat space DOES NOT DEPEND ON POSITION. So whatever coordinates we use (cartesian or spherical polar) gives that the partial derivatives of the metric are ZERO. so $\Gamma^a_{bc} = 0$.

3.6 Example: surface of a sphere

We had previously worked out the basis vectors for the surface of a sphere of radius a and the metric (section 2.10) so I'll just repeat them here, working in coordinates where $x^1 = \theta$, $x^2 = \phi$. Then

$$\vec{e}_1 = \vec{e}_\theta = a\cos\theta\cos\phi\mathbf{i} + a\cos\theta\sin\phi\mathbf{j} - a\sin\theta\mathbf{k}$$

and

$$\vec{e}_2 = \vec{e}_\phi = -a\sin\theta\sin\phi\mathbf{i} + a\sin\theta\cos\phi\mathbf{j}$$

giving $ds^2 = a^2 d\theta^2 + a^2\sin^2\theta d\phi^2$ i.e. $g_{AB} = \mathrm{diag}(a^2, a^2\sin^2\theta)$. so the metric IS a function of coordinates - its got an $x^1 = \theta$ term but no $x^2 = \phi$ term. So all the $\partial_C g_{AB} = 0$ for $C = 2$ ie $\partial_\phi g_{AB} = 0$, but all the terms involving $\partial_1 g_{22} = \partial_\theta g_{\phi\phi} = 2a^2\sin\theta\cos\theta$ need to be included.

Christoffel symbols are defined as

$$\Gamma^f_{ca} = \frac{1}{2} g^{fb}(\partial_c g_{ab} + \partial_a g_{bc} - \partial_b g_{ca})$$

so we need to also find the covariant metric components g^{AB} from $g^{AB}g_{BC} = \delta^A_C$. In general this leads to a set of simultaneous equations to solve

$$g^{AB}g_{BC} = g^{A1}g_{1C} + g^{A2}g_{2C} = \delta^A_C$$

BUT for diagonal metrics we don't have to do a matrix inverse to solve these! in a diagonal metric all cross terms are zero so we get $g^{AB}g_{BA} = g^{AA}g_{AA} = 1$ or $g^{AA} = 1/g_{AA}$ ONLY FOR DIAGONAL METRICS!

$$g^{AB} = \mathrm{diag}[1/a^2, 1/(a^2 \sin^2\theta)]$$

So in our 2D space we have

$$\Gamma^\theta_{\theta\theta} = 1/2g^{\theta B}(\partial_\theta g_{\theta B} + \partial_\theta g_{B\theta} - \partial_B g_{\theta\theta})$$

as $g^{\theta B} = 0$ except for $B = \theta$ then this collapses to

$$\Gamma^\theta_{\theta\theta} = 1/2g^{\theta\theta}(\partial_\theta g_{\theta\theta} + \partial_\theta g_{\theta\theta} - \partial_\theta g_{\theta\theta}) = 1/2\ .\ 1/a^2\ .\ 0 = 0$$

$$\Gamma^\theta_{\theta\phi} = \Gamma^\theta_{\phi\theta} = 1/2g^{\theta B}(\partial_\phi g_{\theta B} + \partial_\theta g_{B\phi} - \partial_B g_{\theta\phi})$$

again $g^{\theta B}$ is only nonzero for $B = \theta$ since the metric is diagonal so

$$\Gamma^\theta_{\theta\phi} = \Gamma^\theta_{\theta\phi} = 1/2g^{\theta\theta}(\partial_\phi g_{\theta\theta} + \partial_\theta g_{\theta\theta} - \partial_\theta g_{\theta\phi}) = 1/2\ .\ 1/a^2\ .\ 0 = 0$$

$$\Gamma^\theta_{\phi\phi} = 1/2g^{\theta B}(\partial_\phi g_{\phi B} + \partial_\phi g_{B\phi} - \partial_B g_{\phi\phi}) = 1/2\ .\ 1/a^2(0 + 0 - \partial_\theta g_{\phi\phi})$$

$$= 1/2\ .\ 1/a^2(-2a^2 \sin\theta\cos\theta) = -\sin\theta\cos\theta$$

$$\Gamma^\phi_{\theta\theta} = 1/2g^{\phi B}(\partial_\theta g_{\theta B} + \partial_\theta g_{B\theta} - \partial_B g_{\theta\theta}) = 1/2g^{\phi\phi}(\partial_\theta g_{\theta\phi} + \partial_\theta g_{\theta\phi} - \partial_\phi g_{\theta\theta}) = 0$$

$$\Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = 1/2g^{\phi B}(\partial_\phi g_{\theta B} + \partial_\theta g_{B\phi} - \partial_B g_{\theta\phi}) = 1/2g^{\phi\phi}(\partial_\phi g_{\theta\phi} + \partial_\theta g_{\phi\phi} - \partial_\phi g_{\theta\phi})$$

$$= 1/2\ 1/(a^2\sin^2\theta)\partial_\theta g_{\phi\phi} = 1/2\ 1/(a^2\sin^2\theta)2a^2\sin\theta\cos\theta = \cot\theta$$

$$\Gamma^\phi_{\phi\phi} = 1/2g^{\phi B}(\partial_\phi g_{\phi\phi} + \partial_\phi g_{\phi\phi} - \partial_\phi g_{\phi\phi}) = 0$$

3.7 Geodesic equations

And now we have everything we need to get the geodesic paths (intertial frames). We could define a geodesic as the shortest distance between two points but actually its easier to start from another property which is that they are the 'natural' paths taken by ants crawling across the surface (left feet, right feet, middle feet). Taking even steps means that the path is always tangent to the surface. In mathematical language this is related to parallel transport - geodesics are curves which parallel transport their tangent vectors. In effect this is saying that the geodesics have tangent vectors all pointing along tangents to the spatial curvature.

Suppose we take some path u and look its tangent vector $\vec{\tau}$ to any point on the curve $\vec{r}(u) = x^a(u)\vec{e_a}$.

$$\vec{\tau} = \frac{d\vec{r}}{du} = \frac{\partial \vec{r}}{\partial x^a}\frac{dx^a}{du} = \frac{dx^a}{du}\vec{e_a}$$

as basis vectors are defined as $\partial\vec{r}/\partial x^a = \vec{e_a}$. So the components $\tau^a = dx^a/du$. From the parallel transport equation above, parallel transport implies

$$\frac{d\tau^a}{du} + \tau^b\Gamma^a_{bc}\frac{dx^c}{du} = 0$$

substitute in$\tau^a = dx^a/du$ from the definition of tangent vector and this becomes

$$\frac{d^2x^a}{du^2} + \Gamma^a_{bc}\frac{dx^b}{du}\frac{dx^c}{du} = 0$$

This works for any parameter linearly related to path length $s = A + Bu$ so the **affinely parameterized geodesic equation** in an N dimensional manifold is

$$\frac{d^2x^a}{ds^2} + \Gamma^a_{bc}\frac{dx^b}{ds}\frac{dx^c}{ds} = 0$$

or in slighly cleaner notation, where dot denotes derivative w.r.t. s

$$\ddot{x}^a + \Gamma^a{}_{bc}\dot{x}^b\dot{x}^c = 0$$

So now we know how to do geodesic paths, as well as derivatives.

3.8 Example: Geodesics in flat space

we saw that in flat space all the $\Gamma^i{}_{jk} = 0$ so the geodesic equations become

$$\ddot{x}^a = 0$$

integrate to get $\dot{x}^a = A$ and $x^a = Ax^a + B$ so geodsics in flat space have constant velocity and direction - Newtonian inertial frame (constant motion, straight line!)

3.9 Example: surface of a sphere

We know that the only non-zero christoffel symbols are ${\Gamma^{\theta}}_{\phi\phi} = -\sin\theta\cos\theta$, and $\Gamma^{\phi}_{\theta\phi} = \Gamma^{\phi}_{\phi\theta} = \cot\theta$. Geodesic paths satisfy the equation

$$\frac{d^2x^A}{du^2} + \Gamma^A_{BC}\frac{dx^B}{du}\frac{dx^C}{du} = 0$$

suppose the path u is just a change in θ then $u = a\theta$ so there is no dependance on ϕ i.e. ϕ =constant so $d\phi/du = 0$. While for θ we have $d\theta/du = 1/a$ and $d^2\theta/du^2 = 0$. For θ then

$$\frac{d^2\theta}{du^2} + \Gamma^{\theta}_{BC}\frac{dx^B}{du}\frac{dx^C}{du} = \frac{d^2\theta}{du^2} + \Gamma^{\theta}_{\phi\phi}\frac{d\phi}{du}\frac{d\phi}{du} = 0$$

as required. Check also that ϕ = constant is consistent with the geodesic equations:

$$\frac{d^2\phi}{du^2} + \Gamma^{\phi}_{BC}\frac{dx^B}{du}\frac{dx^C}{du} = \frac{d^2\phi}{du^2} + \Gamma^{\phi}_{\theta\phi}\frac{d\theta}{du}\frac{d\phi}{du} + \Gamma^{\phi}_{\phi\theta}\frac{d\phi}{du}\frac{d\theta}{du} = 0$$

so all such paths are geodesics. - these are GREAT CIRCLES.

what about paths where θ = constant and $u = a\phi$?

$$\frac{d^2\phi}{du^2} + \Gamma^{\phi}_{BC}\frac{dx^B}{du}\frac{dx^C}{du} = \frac{d^2\phi}{du^2} + \Gamma^{\phi}_{\phi\theta}\frac{d\phi}{du}\frac{d\theta}{du} + \Gamma^{\phi}_{\theta\phi}\frac{d\phi}{du}\frac{d\theta}{du} = 0$$

as required. now check θ:

$$\frac{d^2\theta}{du^2} + \Gamma^{\theta}_{\phi\phi}\frac{d\phi}{du}\frac{d\phi}{du} = -a^2\sin\theta\cos\theta$$

so not 0 unless $\theta = 0, \pi/2, \pi$ ie at the poles or on the equator - at the poles you don't move anywhere! and at the equator its another great circle.

3.10 Easier Christoffel Symbols for geodesic equations

But the Christoffel symbols we defined as derivites of the metric and it all looked horrid to calculate. And geodesics in generality are horrid to calculate as well. An easier way (note I didn't say EASY) is to use the Euler–Lagrange equations. We could have solved for the geodesics (at least for matter particles) by saying that geodesics are paths which give the shortest distance between two points, i.e. we are looking for the extremal path which has

$$\delta \int ds = 0$$

But this is often what we do in classical mechanics, where we look for the minimum energy path by getting the Lagrangian $L = T - V$ as the difference in kinetic T and potential V energies, and then finding the minimum action path by integrating this over time i.e. $\delta \int L dt = 0$. For geodesics then there are no potential energy terms so $V = 0$ so we are looking only at kinetic energy, so $\delta \int T dt = 0$. We'll do it in terms of per unit mass so $T = 1/2[(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2]$, or in metric terms (Euclidean 3D flat space)

$$T = \frac{1}{2} g_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt} = \frac{1}{2} g_{ij} \dot{x}^i \dot{x}^j = \frac{1}{2} \Big(\frac{ds}{dt}\Big)^2$$

So the minimum energy condition (Hamiltons principle in classical mechanics) gives

$$\delta \int T dt = \delta \int \Big(\frac{ds}{dt}\Big)^2 dt = 0$$

i.e. this is basically the same as the minimum path requirement which defines our geodesic which is $\delta \int (ds/dt) dt = 0$.

We know that in classical mechamics that the solution with the minimum energy satisfies the Euler-Lagrange equations i.e.

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{x}^i}\Big) - \frac{\partial L}{\partial x^i} = 0$$

where $L = 1/2 g_{ij} \dot{x}^i \dot{x}^j$. If we wanted we could do long and tedious algebra on these equations to get the equation for a geodesic path, and it would turn out to be of the same form as we got before with far less work by doing parallel transport of a vector i.e.

$$\ddot{x}^i + \Gamma^i_{jk} \dot{x}^j \dot{x}^k = 0$$

or, writing it explicitally

$$\frac{d^2 x^i}{dt^2} + \Gamma^i_{jk} \frac{dx^j}{dt} \frac{dx^k}{dt} = 0$$

While there are obviously some technical problems in carrying this over into our 4D spacetime, where the metric is indefinite (can be +ve, 0 or -ve), if you do it, then it comes out to be the same as above except we are now going to take our derivative wrt some affine parameter u related to the invariant path length rather than to the COORDINATE t. So we could write the geodesic equation as before but where the dot now stands for derivative wrt the affine parameter u. So the geodesic paths are $\ddot{x}^\alpha + \Gamma^\alpha_{\beta\gamma}\dot{x}^\beta\dot{x}^\gamma = 0$.

But then the midstep Euler–Lagrange equations still hold, so we have the condition that

$$\frac{d}{du}\Big(\frac{\partial L}{\partial \dot{x}^\alpha}\Big) - \frac{\partial L}{\partial x^\alpha} = 0$$

And these equations give us directly the Christoffel symbols by comparison with the geodesic equations and equating coefficients, and can give us geodesics themselves.

3.11 Example on spherical surface

$ds^2 = a^2 d\theta^2 + a^2\sin^2\theta d\phi^2$ so $L = \frac{1}{2}(a^2\dot{\theta}^2 + a^2\sin^2\theta\dot{\phi}^2)$. then the EL equations for θ are

$$\frac{d}{du}\frac{\partial L}{\partial\dot{\theta}} - \frac{\partial L}{\partial\theta} = 0$$

$$\frac{d}{du}(a^2\dot{\theta}) - a^2\sin\theta\cos\theta\dot{\phi}^2 = 0$$

$$a^2\ddot{\theta} - a^2\sin\theta\cos\theta\dot{\phi}^2 = 0$$

$$\ddot{\theta} - \sin\theta\cos\theta\dot{\phi}^2 = 0$$

cf the geodesic equation $\ddot{\theta} + \Gamma^\theta{}_{BC}\dot{x}^B\dot{x}^C = 0$. so only nonzero $\Gamma^\theta{}_{BC} = \Gamma^\theta{}_{\phi\phi} = -\sin\theta\cos\theta$. do same for ϕ and get

$$\frac{d}{du}\frac{\partial L}{\partial\dot{\phi}} - \frac{\partial L}{\partial\phi} = 0$$

$$\frac{d}{du}(a^2\sin^2\theta)\dot{\phi} - 0 = 0$$

$$a^2\sin^2\theta\ddot{\phi} + a^2\dot{\phi}2\sin\theta\cos\theta\dot{\theta} = 0$$

$$\ddot{\phi} + 2\cot\theta\dot{\phi}\dot{\theta} = 0$$

cf geodesic $\ddot{\phi} + \Gamma^\phi{}_{BC}\dot{x}^B\dot{x}^C = 0$. so $\Gamma^\phi{}_{\theta\phi}\dot{\theta}\dot{\phi} + \Gamma^\phi{}_{\phi\theta}\dot{\phi}\dot{\theta} = 2\cot\theta$. so $\Gamma^\phi{}_{\theta\phi} = \Gamma^\phi{}_{\phi\theta} = \cot\theta$.

3.12 Equivalence of Euler-Lagrange and geodesic equations

The justification of the Euler-Lagrange equations was simply that - a justification, not a proof. can we really use them? answer is yes, here we'll expliciatly prove that they are the same thing

$L = \frac{1}{2}g_{ab}\dot{x}^a\dot{x}^b$ and the metric is a function of position NOT velocity so $\partial g_{ab}/\partial\dot{x}^c = 0$

$$\frac{\partial L}{\partial \dot{x}^c} = \frac{1}{2}g_{ab}\Big(\dot{x}^a\frac{\partial \dot{x}^b}{\partial \dot{x}^c} + \frac{1}{2}g_{ab}\dot{x}^b\frac{\partial \dot{x}^a}{\partial \dot{x}^c}\Big)$$

$$= \frac{1}{2}g_{ab}\dot{x}^a\delta^b_c + \frac{1}{2}g_{ab}\dot{x}^b\delta^a_c = \frac{1}{2}g_{ac}\dot{x}^a + \frac{1}{2}g_{cb}\dot{x}^b = g_{cb}\dot{x}^b$$

whereas $\partial_c L = \frac{1}{2}\dot{x}^a\dot{x}^b\partial_c g_{ab}$. put this into the E-L equations and they become

$$\frac{d}{ds}(g_{cb}\dot{x}^b) - \frac{1}{2}\dot{x}^a\dot{x}^b\frac{\partial g_{ab}}{\partial x^c} = 0$$

$$g_{cb}\ddot{x}^b + \dot{x}^b\frac{dg_{cb}}{ds} - \frac{1}{2}\dot{x}^a\dot{x}^b\partial_c g_{ab} = 0$$

$$g_{cb}\ddot{x}^b + \dot{x}^b\dot{x}^a\partial_a g_{cb} - \frac{1}{2}\dot{x}^a\dot{x}^b\partial_c g_{ab} = 0$$

$$g_{cb}\ddot{x}^b + \frac{1}{2}\dot{x}^b\dot{x}^a\partial_a g_{cb} + \frac{1}{2}\dot{x}^b\dot{x}^a\partial_a g_{cb} - \frac{1}{2}\dot{x}^a\dot{x}^b\partial_c g_{ab} = 0$$

but $\dot{x}^b\dot{x}^a\partial_a g_{cb} = \dot{x}^d\dot{x}^a\partial_a g_{cd}$ and $\dot{x}^b\dot{x}^a\partial_a g_{cb} = \dot{x}^a\dot{x}^d\partial_d g_{ca}$ and $\dot{x}^a\dot{x}^b\partial_c g_{ab} = \dot{x}^a\dot{x}^d\partial_c g_{ad}$ as these are dummy indices so we can relabel them to whatever we want and now we can get everything in terms of $\dot{x}^a\dot{x}^d$

$$g_{cb}\ddot{x}^b + \frac{1}{2}\dot{x}^d\dot{x}^a(\partial_a g_{cd} + \partial_d g_{ca} - \partial_c g_{ad}) = 0$$

multiply by the contravariant metric g^{fc}

$$g^{fc}g_{cb}\ddot{x}^b + \frac{1}{2}g^{fc}(\partial_a g_{cd} + \partial_d g_{ca} - \partial_c g_{ad})\dot{x}^d\dot{x}^a = 0$$

$$\delta^f_b\ddot{x}^b + \Gamma^f{}_{ad}\dot{x}^d\dot{x}^a = 0$$

$$\ddot{x}^f + \Gamma^f{}_{ad}\dot{x}^d\dot{x}^a = 0$$

i.e. identical to the geodesic equations!

3.13 Geodesics in weakly curved space = Newtonian gravity!

Now lets go apply some of some of this new found knowledge of Christoffel symbols to REAL PHYSICS! Suppose we are looking at a material particle in free fall. It goes along a geodesic so

$$\frac{d^2x^\alpha}{du^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{du}\frac{dx^\gamma}{du} = 0$$

where u is an affine parameter, linearly related to path length s. well, proper time τ is an affine parameter as $c^2d\tau^2 = ds^2$ so

$$\frac{d^2x^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = 0$$

$$\frac{d^2x^\alpha}{d\tau^2} = -\Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau}$$

Left hand side is acceleration - so the right hand side must be gravity! Go to the Newtonian limit in a stationary gravitational field, and $v \ll c$, then $dx^0/d\tau = cdt/d\tau \sim c$ while $dx^i/d\tau = v^i dt/d\tau \sim v \ll c$ so we can neglect the spatial derivatives in comparison to the time ones.

$$\frac{d^2x^\alpha}{d\tau^2} \approx -\Gamma^\alpha_{00}\frac{dx^0}{d\tau}\frac{dx^0}{d\tau} = -c^2\Gamma^\alpha_{00}\left(\frac{dt}{d\tau}\right)^2 \approx -c^2\Gamma^\alpha_{00}$$

The field is stationary (in time) so $\partial g_{\mu\nu}/\partial x^0 = 0$ i.e. only the spatial derivatives of the metric can matter.

$$\Gamma^f_{ca} = \frac{1}{2}g^{fb}(\partial_c g_{ab} + \partial_a g_{bc} - \partial_b g_{ca})$$

$$\Gamma^\alpha_{00} = \frac{1}{2}g^{\alpha\beta}(\partial_0 g_{0\beta} + \partial_0 g_{\beta 0} - \partial_\beta g_{00}) = -\frac{1}{2}g^{\alpha\beta}\partial_\beta g_{00}$$

If the field is weak then we can write the metric as s sum of the flat spacetime metric which we'll now call $\eta_{\alpha\beta} = \mathrm{diag}(1, -1, -1, -1)$ plus some little perturbation: $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$ so

$$\Gamma^\alpha_{00} = -\frac{1}{2}g^{\alpha\beta}\partial_\beta g_{00} = -\frac{1}{2}(\eta^{\alpha\beta} + h^{\alpha\beta})\partial_\beta(g_{00} + h_{00}) \approx -\frac{1}{2}\eta^{\alpha\beta}\partial_\beta h_{00}$$

But we said there is no time derivative so we can go to 3D rather than 4D and the metric is diagonal so

$$\Gamma^i_{00} = -\frac{1}{2}\eta^{ii}\partial_i h_{00}$$

The geodesic (equation of motion) becomes

$$\frac{d^2x^\alpha}{d\tau^2} \approx \frac{d^2x^\alpha}{dt^2} \approx \frac{c^2}{2}\eta^{ii}\partial_i h_{00} = \nabla(-c^2 h_{00}/2)$$

But we know that if we are a distance r from a star of mass M then we measure an acceleration (per unit mass) of

$$\frac{d^2x^i}{dt} = \nabla - \phi \qquad \phi = -GM/r$$

so

$$h_{00} = \frac{2\phi}{c^2} + \text{const} = \frac{2\phi}{c^2}$$

as for $r \to \infty$ then $\phi \to 0$ and $g_{\alpha\beta} \to \eta_{\alpha\beta}$ so $h_{00} \to 0$. So this gives us

$$g_{00} = \eta_{00} + h_{00} = 1 + \frac{2\phi}{c^2} = 1 - 2GM/(c^2 r)$$

So we started from our geodesic equation which only told us about curved space. But we ended up with something that looks exactly like gravity!

3.14 Summary: Tensor derivatives

Absolute derivative of a contravariant tensor over some path

$$\frac{D\lambda^a}{du} = \frac{d\lambda^a}{du} + \lambda^b\Gamma^a_{bc}\frac{dx^c}{du}$$

gives a tensor field of same type (contravariant first order) in this case. The first bit takes into account both real physical changes in λ^a AND the way the curvature of the space can swing the vector while the second one takes out the swinging the vector from the curvature of space alone. So if there is no physical change then $D\lambda^a/du = 0$ and the vector is simply parallelly transported. Absolute derivative obeys all the normal rules for derivatives so $D(\lambda^a + k\mu^a)/du = D(\lambda^a)/du + kD(\mu^a)/du$ (linear) and $D(\lambda^a\mu^a)/du = \mu^a D(\lambda^a)/du + \lambda^a D(\mu^a)/du$ (Leibniz' rule).

Absolute derivative of a scalar

There is no space swing with a simple number, ϕ. so $D\phi/du = d\phi/du$

Absolute derivative of covariant tensors

$D\phi/du = d\phi/du$ so let $\phi = \lambda^a\mu_a$ then

$$\frac{D\phi}{du} = \frac{d\phi}{du} = \frac{D(\lambda^a\mu_a)}{du} = \mu_a\frac{D\lambda^a}{du} + \lambda^a\frac{D\mu_a}{du}$$

$$\frac{d(\lambda^a\mu_a)}{du} = \mu_a\Big(\frac{d\lambda^a}{du} + \lambda^b\Gamma^a_{bc}\frac{dx^c}{du}\Big) + \lambda^a\frac{D\mu_a}{du}$$

$$\lambda^a\frac{d\mu_a}{du} + \mu_a\frac{d\lambda^a}{du} = \mu_a\frac{d\lambda^a}{du} + \mu_a\lambda^b\Gamma^a_{bc}\frac{dx^c}{du} + \lambda^a\frac{D\mu_a}{du}$$

$$\lambda^a\frac{D\mu_a}{du} = \lambda^a\frac{d\mu_a}{du} - \mu_a\lambda^b\Gamma^a_{bc}\frac{dx^c}{du}$$

do some index manipulation and get

$$\frac{D\mu_a}{du} = \frac{d\mu_a}{du} - \Gamma^b_{ac}\mu_b\frac{dx^c}{du}$$

Absolute derivative of higher order tensors

For example, to get the absolute derivative of a mixed tensor τ^a_b then look at the special case where $\tau^a_b = \lambda^a\mu_b$

$$\frac{D\tau^a_b}{du} = \lambda^a\frac{D\mu_b}{du} + \mu_b\frac{D\lambda^a}{du} = \lambda^a(\frac{d\mu_b}{du} - \Gamma^d_{bc}\mu_d\frac{dx^c}{du}) + \mu_b(\frac{d\lambda^a}{du} + \Gamma^a_{dc}\lambda^d\frac{dx^c}{du})$$

$$= \frac{d(\lambda^a\mu_b)}{du} + \mu_b\Gamma^a_{dc}\lambda^d\frac{dx^c}{du} - \lambda^a\Gamma^d_{bc}\mu_d\frac{dx^c}{du} = \frac{d(\tau^a_b)}{du} + \Gamma^a_{dc}\tau^d_b - \Gamma^d_{bc}\tau^a_d\frac{dx^c}{du}$$

Covariant derivative

$$\lambda^a_{;c} = \frac{\partial\lambda^a}{\partial x^c} + \Gamma^a_{bc}\lambda^b$$

rules are as above, so for a second order contravariant tensor we have

$$\tau^{ab}{}_{;c} = (\lambda^a\mu^b)_{;c} = \lambda^a{}_{;c}\mu_b + \lambda^a\mu_{b;c}$$

$$= \frac{\partial \lambda^a}{\partial x^c}\mu^b + \Gamma^a_{dc}\lambda^d\mu^b + \lambda^a\frac{\partial \mu^b}{\partial x^c} + \lambda^a\Gamma^b_{dc}\mu^d = \partial_c\tau^{ab} + \Gamma^a_{dc}\tau^{db} + \Gamma^b_{dc}\tau^{ad}$$

similarly for covariant

$$\lambda_{ab;c} = \frac{\partial \lambda_{ab}}{\partial x^c} - \Gamma^d_{ac}\lambda_{db} - \Gamma^d_{bc}\lambda_{ad}$$

Covariant differention forms a tensor field of 1 higher covariant order than the original tensor field.

Covariant derivative of the metric

In getting the Christoffel symbols (section 3.4) in terms of the metric we had

$$\frac{\partial g_{ab}}{\partial x^c} = \frac{\partial \vec{e_a}.\vec{e_b}}{\partial x^c} = \frac{\partial \vec{e_a}}{\partial x^c}.\vec{e_b} + \vec{e_a}.\frac{\partial \vec{e_b}}{\partial x^c}$$
$$= \Gamma^d_{ac}\vec{e_d}.\vec{e_b} + \vec{e_a}.\vec{e_d}\Gamma^d_{bc} = \Gamma^d_{ac}g_{db} + \Gamma^d_{bc}g_{ad}$$

rearrange this to get

$$\frac{\partial g_{ab}}{\partial x^c} - \Gamma^d_{ac}g_{db} - \Gamma^d_{bc}g_{ad} = 0$$

but this is the definition of the covariant derivative of a second order covarient tensor. so $g_{ab;c} = 0$, and we can similarly get that $g^{ab}_{;c} = 0$. This means that index raising and lowering can be swapped in and out of covariant differention. $R^a = g^{ab}R_b$. Then $R^a_{;c} = (g^{ab}R_b)_{;c} = g^{ab}_{;c}R_b + g^{ab}R_{b;c} = g^{ab}R_{b;c}$

Covariant derivative

$$\lambda^a{}_{;c} = \frac{\partial \lambda^a}{\partial x^c} + \Gamma^a_{bc}\lambda^b$$

Covariant differention forms a tensor field of 1 higher covariant order than the original tensor field. Rules are the same as for absolute derivative.

so for a scalar $\phi = \lambda^a\mu_a$, then $\phi_{;c} = \partial_c\phi$ (ie no spatial terms as a number is just a number) so

$$\phi_{;c} = \partial_c\phi = (\lambda^a\mu_a)_{;c} = \lambda^a\mu_{a;c} + \mu_a\lambda^a{}_{;c}$$
$$\lambda^a\partial_c\mu_a + \mu_a\partial_c\lambda^a = \lambda^a\mu_{a;c} + \mu_a(\partial_c\lambda^a + \Gamma^a{}_{bc}\lambda^b)$$
$$\lambda^a\partial_c\mu_a = \lambda^a\mu_{a;c} + \mu_a\Gamma^a{}_{bc}\lambda^b$$

do index switching on the last term to get it all onto λ^a so we can cancel it

$$\lambda^a \partial_c \mu_a = \lambda^a \mu_{a;c} + \mu_b \Gamma^b{}_{ac} \lambda^a$$

$$\mu_{a;c} = \partial_c \mu_a - \Gamma^b{}_{ac} \mu_b$$

Higher order: $\tau_{ab} = \lambda_a \mu_b$

$$\tau_{ab;c} = (\lambda_a \mu_b)_{;c} = \lambda_{a;c} \mu_b + \lambda_a \mu_{b;c}$$

$$= (\partial_c \lambda_a - \Gamma^d_{ac} \lambda_d) \mu_b + \lambda_a (\partial_c \mu_b - \Gamma^d_{bc} \mu_d)$$
$$= \partial_c (\lambda_a \mu_b) - \Gamma^d_{ac} \lambda_d \mu_b - \lambda_a \Gamma^d_{bc} \mu_d$$
$$= \partial_c \tau_{ab} - \Gamma^d_{ac} \tau_{db} - \Gamma^d_{bc} \tau_{ad}$$

This is exactly the same as the expression we got from taking the derivative of the metric tensor so $g_{ab;c} = 0$

Hence $\lambda_{a;c} = (g_{ab} \lambda^b)_{;c} = g_{ab;c} \lambda^b + g_{ab} \lambda^b{}_{;c} = g_{ab} \lambda^b{}_{;c}$ so raising and lowering indices can be done either before or after taking a covariant derivative.

4 Tensor and Physical Curvature

So we've seen that gravity=curvature. So it would be nice to have a way to quantify curvature. There is one (but it isn't nice!).

4.1 Riemann Curvature Tensor

Key thing about curvature is that neighbouring geodesics get further apart (or closer together) at a rate depending on the local curvature. Flat space, no forces, geodesics are straight lines. The distance between two geodesics then increases/decreases linearly as a function of path length along the geodesic. So the second derivative of the distance between them is zero. For curved space then the geodesics don't separate linearly - the second derivative will not be zero, and this gives us a way to quantify the curvature of spacetime.

Take two geodesics γ and $\tilde{\gamma}$ with coodinates $x^a(u)$ and $\tilde{x}^a(u)$. The distance between them is $\zeta^a(u) = \tilde{x}^a(u) - x^a(u) =$. These are geodesics so

$$\frac{d^2 x^a}{du^2} + \Gamma^a_{bc} \frac{dx^b}{du} \frac{dx^c}{du} = 0$$

$$\frac{d^2\tilde{x}^a}{du^2} + \tilde{\Gamma}^a_{bc}\frac{d\tilde{x}^b}{du}\frac{d\tilde{x}^c}{du} = 0$$

but

$$\tilde{\Gamma}^a_{bc} = \Gamma^a_{bc} + \frac{\partial \Gamma^a_{bc}}{\partial x^d}(\tilde{x}^d - x^d) = \Gamma^a_{bc} + \partial_d \Gamma^a_{bc}\zeta^d$$

while subtracting the geodesics gives

$$\frac{d^2\tilde{x}^a}{du^2} + \tilde{\Gamma}^a_{bc}\frac{d\tilde{x}^b}{du}\frac{d\tilde{x}^c}{du} - \frac{d^2x^a}{du^2} - \Gamma^a_{bc}\frac{dx^b}{du}\frac{dx^c}{du} = 0$$

$$\frac{d^2(x^a+\zeta^a)}{du^2} + (\Gamma^a_{bc} + \partial_d\Gamma^a_{bc}\zeta^d)\frac{d(x^b+\zeta^b)}{du}\frac{d(x^c+\zeta^c)}{du} - \frac{d^2x^a}{du^2} - \Gamma^a_{bc}\frac{dx^b}{du}\frac{dx^c}{du} = 0$$

Let dot denote derivative with respect to path length, so this becomes

$$\ddot{x^a} + \ddot{\zeta}^a + \Gamma^a_{bc}\dot{x^b}\dot{x^c} + \Gamma^a_{bc}\dot{x^b}\dot{\zeta^c} + \Gamma^a_{bc}\dot{x^c}\dot{\zeta^b} + \Gamma^a_{bc}\dot{\zeta^b}\dot{\zeta^c} +$$

$$\partial_d\Gamma^a_{bc}\zeta^d\dot{x^b}\dot{x^c} + \partial_d\Gamma^a_{bc}\zeta^d\dot{x^b}\dot{\zeta^c} + \partial_d\Gamma^a_{bc}\zeta^d\dot{x^c}\dot{\zeta^b} + \partial_d\Gamma^a_{bc}\zeta^d\dot{\zeta^b}\dot{\zeta^c} - \ddot{x^a} - \Gamma^a_{bc}\dot{x^b}\dot{x^c} = 0$$

keep only first order terms in ζ and $\dot{\zeta}$ so ignore $\zeta^a\zeta^b$, $\zeta^a\dot{\zeta}^b$ and $\dot{\zeta}^a\dot{\zeta}^b$ - this is equivalent to saying the geodesics are originally close together. NB we keep the second derivative because if we had $y = x^2$ then $\dot{y} = 2x$ and $\ddot{y} = 2$. so if x is small then x and $\dot{x}$ are small but there is nothing you can do about 2!! so we keep second derivatives.

$$\ddot{\zeta}^a + \Gamma^a_{bc}\dot{x}^b\dot{\zeta}^c + \Gamma^a_{bc}\dot{\zeta}^b\dot{x}^c + \partial_d\Gamma^a_{bc}\zeta^d\dot{x}^b\dot{x}^c = 0$$

We have this term $\ddot{\zeta}^a$. We are looking for a tensor equation, and we'd get something with this term that transformed as a tensor if we had $D^2\zeta^a/du^2$.

$$\frac{D^2\zeta^a}{du^2} = \frac{D}{du}(\dot{\zeta}^a + \Gamma^a_{bc}\zeta^b\dot{x}^c) = \frac{D}{du}\dot{\zeta}^a + \frac{D}{du}(\Gamma^a_{bc}\zeta^b\dot{x}^c)$$

$$= \ddot{\zeta}^a + \Gamma^a_{bc}\dot{\zeta}^b\dot{x}^c + \frac{d}{du}(\Gamma^a_{bc}\zeta^b\dot{x}^c) + \Gamma^a_{ef}(\Gamma^e_{bc}\zeta^b\dot{x}^c)\dot{x}^f$$

$$= \ddot{\zeta}^a + \Gamma^a_{bc}\dot{\zeta}^b\dot{x}^c + \frac{d\Gamma^a_{bc}}{du}\zeta^b\dot{x}^c + \Gamma^a_{bc}\dot{\zeta}^b\dot{x}^c + \Gamma^a_{bc}\zeta^b\ddot{x}^c + \Gamma^a_{ef}\Gamma^e_{bc}\zeta^b\dot{x}^c\dot{x}^f$$

We can substitute for $\ddot{x}$ from the geodesic equation along γ, and take the christoffel symbol derivative with respect to coordinates rather than the affine parameter

$$= \ddot{\zeta}^a + 2\Gamma^a_{bc}\dot{\zeta}^b\dot{x}^c + \partial_d\Gamma^a_{bc}\zeta^b\dot{x}^c\dot{x}^d - \Gamma^a_{bc}\zeta^b\Gamma^c_{ef}\dot{x}^e\dot{x}^f + \Gamma^a_{ef}\Gamma^e_{bc}\zeta^b\dot{x}^c\dot{x}^f$$

so we then have an expression for $\ddot{\zeta}$ as

$$\ddot{\zeta}^a = \frac{D^2\zeta^a}{du^2} - 2\Gamma^a_{bc}\dot{\zeta}^b\dot{x}^c - \partial_d\Gamma^a_{bc}\zeta^b\dot{x}^c\dot{x}^d + \Gamma^a_{bc}\zeta^b\Gamma^c_{ef}\dot{x}^e\dot{x}^f - \Gamma^a_{ef}\Gamma^e_{bc}\zeta^b\dot{x}^c\dot{x}^f$$

Substitute for $\ddot{\zeta}^a$ into our original equation for the separation of two geodesics

$$\ddot{\zeta}^a + \Gamma^a_{bc}\dot{x}^b\dot{\zeta}^c + \Gamma^a_{bc}\dot{\zeta}^b\dot{x}^c + \partial_d\Gamma^a_{bc}\zeta^d\dot{x}^b\dot{x}^c = 0$$

$$\frac{D^2\zeta^a}{du^2} - 2\Gamma^a_{bc}\dot{\zeta}^b\dot{x}^c - \partial_d\Gamma^a_{bc}\zeta^b\dot{x}^c\dot{x}^d + \Gamma^a_{bc}\zeta^b\Gamma^c_{ef}\dot{x}^e\dot{x}^f - \Gamma^a_{ef}\Gamma^e_{bc}\zeta^b\dot{x}^c\dot{x}^f + \Gamma^a_{bc}\dot{x}^b\dot{\zeta}^c$$
$$+\Gamma^a_{bc}\dot{\zeta}^b\dot{x}^c + \partial_d\Gamma^a_{bc}\zeta^d\dot{x}^b\dot{x}^c = 0$$

$$\frac{D^2\zeta^a}{du^2} - \Gamma^a_{bc}\dot{\zeta}^b\dot{x}^c - \partial_d\Gamma^a_{bc}\zeta^b\dot{x}^c\dot{x}^d + \Gamma^a_{bc}\zeta^b\Gamma^c_{ef}\dot{x}^e\dot{x}^f - \Gamma^a_{ef}\Gamma^e_{bc}\zeta^b\dot{x}^c\dot{x}^f + \Gamma^a_{bc}\dot{x}^b\dot{\zeta}^c + \partial_d\Gamma^a_{bc}\zeta^d\dot{x}^b\dot{x}^c = 0$$

But the Christoffel symbols are symmetric so $\Gamma^a_{bc}\dot{\zeta}^b\dot{x}^c = \Gamma^a_{cb}\dot{\zeta}^c\dot{x}^b = \Gamma^a_{bc}\dot{\zeta}^c\dot{x}^b$

$$\frac{D^2\zeta^a}{du^2} - \partial_d\Gamma^a_{bc}\zeta^b\dot{x}^c\dot{x}^d + \Gamma^a_{bc}\zeta^b\Gamma^c_{ef}\dot{x}^e\dot{x}^f - \Gamma^a_{ef}\Gamma^e_{bc}\zeta^b\dot{x}^c\dot{x}^f + \partial_d\Gamma^a_{bc}\zeta^d\dot{x}^b\dot{x}^c = 0$$

swap indices around on the last term $\partial_d\Gamma^a_{bc}\zeta^d\dot{x}^b\dot{x}^c = \partial_b\Gamma^a_{dc}\zeta^b\dot{x}^d\dot{x}^c$ so

$$\frac{D^2\zeta^a}{du^2} + (\Gamma^a_{be}\Gamma^e_{cd} - \partial_d\Gamma^a_{bc} - \Gamma^a_{ed}\Gamma^e_{bc} + \partial_b\Gamma^a_{dc})\zeta^b\dot{x}^c\dot{x}^d = 0$$

So here we have something which should equal zero if the space is flat, and not zero if the space is curved. This can all be stuck together into the Riemann curvature tensor

$$\frac{D^2\zeta^a}{du^2} + R^a_{cbd}\zeta^b\dot{x}^c\dot{x}^d = 0$$

where

$$R^a_{cbd} = (\Gamma^a_{be}\Gamma^e_{cd} - \partial_d\Gamma^a_{bc} - \Gamma^a_{ed}\Gamma^e_{bc} + \partial_b\Gamma^a_{dc})$$

Geodesics in flat space maintain their separation. Those in curved space do not.

4.2 Newtonian gravity meaning of curvature tensor

In Newtonian gravity then two particles initially at rest at the same horizontal distance x but at vertical distance $\pm y$ (where $y \ll x$) from a star of mass M are accelerated radially by the gravitational force $\sim GM/x^2$. Since they are in free fall they are travelling on geodesics, and these geodesics are getting closer together as the particles fall towards the star. Resolving the radial force on one of the stars into a force along the same direction as the other radial force, and one perperdicular to it then there is a force towards the other particle of size $2GM/x^2 y/x$ and the distance apart is $\sim 2y$. Thus there is an acceleration term between the two particles of magnitude which is pulling them together so

$$\frac{d^2y}{dt^2} = -\frac{GM}{x^3}y$$

Convert this now to path length using dirty astrophysicists approximations so $ds^2/dt^2 \sim c^2$

$$\frac{d^2y}{ds^2}\Big(\frac{ds}{dt}\Big)^2 = \frac{d^2y}{ds^2}c^2 = -\frac{GM}{x^3}y$$

$$\frac{d^2x}{ds^2} = -\frac{GM}{c^2x^3}y$$

Compare this with the curvature tensor

$$\frac{D^2\zeta^a}{du^2} = -R^a_{cbd}\zeta^b\dot{x}^c\dot{x}^d$$

Biggest term again is for $c = d = 0$ so

$$\frac{D^2\zeta^a}{du^2} \approx -R^a_{0b0}\zeta^b c^2$$

ζ^a is the separation and its $(0, 0, y, 0)$ as both particles have the same t, x and z then there is only separation in y so this becomes

$$\frac{D^2y}{du^2} \approx -R^y_{0y0}yc^2$$

Again we see explicitally that the tidal forces which we'd associate with gravity can be made simply from curved spacetime. The curvature tensor in some sense tells us about TIDAL forces - how gravity changes from point to point.

And in flat space, then all the $R^a{}_{cbd} = 0$ so $D^2\zeta^a/du^2 = 0$ and geodesics separate as $\zeta^a = Au + B$ i.e. in flat space, geodesics either maintain a constant separation (parallel lines) or separate/approach linearly. So with no gravity things continue in staright lines at constant velocity!

4.3 More on the Curvature tensor

We've seen above that the Riemann curvature tensor is teh way to describe curvature in all its gory detail. Hence it also describes other curvature dependent quantities such as the swing in the tensor components if its parallel transported around a little loop about a point P. If the looped components are λ^a(loop) and the old are λ^a then

$$\lambda^a(\text{loop}) - \lambda^a = \Delta\lambda^a = -\frac{1}{2}(R^a{}_{bcd})_P \lambda^b f^{cd}$$

where f^{cd} is a coordinate area term.

Since curvature also comes into the covariant derivative, then its the Riemann curvature tensor which gives the difference in order of derivatives..

$$\lambda^a{}_{;bc} - \lambda^a{}_{;cb} = -R^a{}_{dbc}\lambda^d$$

BUT WHAT IT PHYSICALLY MEANS ISN'T ANY OF THESE!! IT PLYSICALLY MEANS TIDAL FORCES FROM GRAVITY!

4.4 Stress-Energy Tensor

So the curvature of space is related to tidal forces, ie to M/R^3 which is the density of mass. We want to get some tensor equation (valid in any frame) such that curvature=gravity, so we want some tensor equation linking the Riemann curvature tensor and mass density. But it can't just be mass density. Einstein said that mass and energy are equivalent so in fact ALL forms of energy (not just mass) should gravitate. And in fact if this were so it would make some sense of SR. If you accelerate a particle, giving it more and more energy, then as you approach c then the mass increases rather than the velocity. Which is all a bit arbitrary unless all forms of energy gravitate because then as we add energy we increase the inertial (i.e. gravitational) mass. So what we want is a tensor equation linking energy density with curvature.

First we will see how to characterise energy density in SR. In SR we generally dealt with individual particles. where $E = \gamma m_o c^2$ which is the particle energy, $u^\mu = dx^\mu/d\tau$ is world velocity, $v^\mu = dx^\mu/dt$ is coordinate velocity and $p^\mu = m_0 u^\mu$ is the 4 momentum of the particle. A stationary particle has $p^\mu = m_0(c, \vec{0})$ while a moving one has $p^\mu = m_0\gamma(c, \vec{v}) = (E/c, \vec{p})$. We want to generalise these to look at density. This is a problem because of course density is a frame dependent quantity. The total number of particles in a given volume must be invariant but the volume changes by length contraction along the direction of motion, so if n_0 is the number density in the rest frame then $n_0\gamma$ is the number density as measured in a moving frame. But if we are now moving, then there is a flux of particles across the surfaces. i.e. we take two separate concepts density and flux, and make them into a single 4-vector the number-flux 4-vector $\tilde{N} = n_0\tilde{u}$ so this has components $N^\mu = n_0 u^\mu$. This is a 4-vector as it transforms in the right way for a contravariant tensor. In the rest frame then $N^\mu = n_0 dx^\mu/d\tau$. In a moving frame then $N^{\nu'} = ndx^{\nu'}/d\tau = n_0 dx^{\nu'}/dx^\mu dx^\mu/d\tau = dx^{\nu'}/dx^\mu N^\mu$. Alternatively we could just have said that we know $u^\mu = dx^\mu/d\tau$ is a 4 vector and so since N^μ is related to it only by invariant quantities (rest frame density) then this must also be a 4-vector.

So that is density done. But we wanted energy density. Lets stick in the rest frame for the time being, then if the particles have no velocity relative to each other then the energy density is simply $m_0 c^2 n_0 = \rho_0 c^2$ where ρ_0 is the mass density in the rest frame. But if we then go to another frame we have to transform the energy AND the density - so it transforms by two factors of γ. So this needs to be a second order contravariant tensor. How about $T^{\mu\nu} = N^\mu p^\nu$ - in the rest frame then this would have only one component $T^{00} = n_0 c m_0 c = \rho_0 c^2$. That sounds like a good start. So $T^{\mu\nu} = N^\mu p^\nu = n u^\mu m_0 u^\nu = \rho_0 u^\mu u^\nu$.

This works if we are only dealing with **dust** i.e. something where the particles have no internal motions, and no stresses or heat conduction or anything complicated. But in general a gas HAS internal motion - particles are moving with respect to each other so there is no rest frame as such for each particle, only a rest frame for the gas as a whole.

We can instead define a general $T^{\mu\nu}$ which is the flow of the μ component of 4-momentum across a surface of constant ν. Then in the rest frame we have T^{00} as flux of 0th component of 4 momentum (E/c) across a surface of constant ct - energy density, cT^{0i} is energy flow per unit area along the i direction - this would be heat conduction. $T^{i0}c$ is the density of the i component of momentum. T^{ii} is the flow of the ith momentum component per unit area in the i direction - i.e. pressure across the i plane. T^{ij} is the flow of the ith momentum component across the j plane - these are viscous drags.

In the rest frame the viscous drags have to be equal (or the fluid as a whole rotates and its not a rest frame) so $T^{ij} = T^{ji}$. And if heat flows along a direction i then that energy will carry momentum and $T^{0i} = T^{i0}$ So the general stress-energy tensor is symmetric in the rest frame - so it must be **symmetric** everywhere.

A **perfect fluid** is one in which there is no heat conduction or viscous stresses. Then in the rest frame

$$T^{\mu\nu} = \mathrm{diag}(\rho_0 c^2, p, p, p)$$

This makes a lot of sense - for a relativistic gas the mean energy per particle is 3kT, so the total rest frame energy density $\rho_0 c^2 + n_0 3kT = \rho_0 c^2 + 3p$. We can already see something really fun. PRESSURE contributes to energy density, and energy density curves space. So for the interior of a star, as the fuel runs out then the star contracts and gravity gets stronger. so we require a larger pressure to hold the star up. But this presssure adds to the energy density, so increases gravity, so more pressure is needed so .... black holes!

Lets write the stress energy tensor for a perfect fluid as the sum of tensors

$$T^{\mu\nu} = (\rho_0 + p/c^2)u^\mu u^\nu - p\eta^{\mu\nu}$$

so we can see its a tensor as its made out of things that are either invariant (scalars) or tensors.

4.5 Conservation laws

The stress energy tensor for a perfect fluid in special relativity.

$$T^{\mu\nu} = (\rho_0 + p/c^2)u^\mu u^\nu - p\eta^{\mu\nu}$$

is a really compact way to count up all the contributions to the energy density, and its a tensor so it works in all (inertial) frames.

But the stress energy tensor also embodies more information. Energy is conserved. Rate of change of energy is $l^3 \partial T^{00}/\partial t$. But this change is produced by net energy inflow/outflow along the 6 faces of the cube. The energy flow in is $l^2 c T^{01}(x)$ and then energy flow out is $l^2 c T^{01}(x+l)$ so net inward flow is $-l^3 c \partial T^{01}/\partial x$. There are simialr contributions from the other 2 pairs of faces so we have

$$l^3 \frac{\partial T^{00}}{\partial t} = -l^3 c \frac{\partial T^{01}}{\partial x} - l^3 c \frac{\partial T^{02}}{\partial y} - l^3 c \frac{\partial T^{03}}{\partial z}$$

$$\frac{\partial T^{00}}{\partial t} + \frac{\partial T^{0i}}{\partial x^i} = 0$$

$$\partial_\alpha T^{0\alpha} = T^{0\alpha}{}_{,\alpha} = 0$$

Momentum is also conserved. and that leads to

$$\partial_\alpha T^{i\alpha} = T^{i\alpha}{}_{,\alpha} = 0$$

In other words, conservation of both momentum and energy are embodied in

$$\partial_\alpha T^{\beta\alpha} = T^{\beta\alpha}{}_{,\alpha} = 0$$

as the stress-energy tensor is symmetric then there is only one divergance $\partial_\alpha T^{\beta\alpha} = \partial_\alpha T^{\alpha\beta}$.

4.6 Strong equivalence principle - how to go from SR to GR

Any physical law which can be expressed in tensor notation in SR has EXACTLY the same form in a locally inertial frame of a curve spacetime (GR)! Locally inertial frames (free fall frames) are equivalent to saying we choose to look at a small piece of spacetime, where we can approximate any curved space to a locally flat space. In this choice of coordinates (local cartesian coordinates) then all the christoffel symbols are zero so standard diferentiation and covariant differentiation are the same. BUT once we go outside of this the curvature becomes apparent. So to make our tensor equations work in any general frame we need to swap standard differentiation (which is NOT a real tensor) ($\partial/\partial x^c = \partial_c =_{,c}$ with covariant differentiation $_{;c}$ (which IS a tensor). So this is sometimes called the comma goes to semicolon rule!

So if we wanted to generalise $T^{\mu\nu}$ and its conservation laws to GR we'd swap the metric tensor from flat space to generalised curved space so for a perfect fluid

$$T^{\mu\nu} = (\rho_0 + p/c^2)u^\mu u^\nu - pg^{\mu\nu}$$

and take the covariant deriavative rather than the partial derivative

$$T^{\beta\alpha}{}_{;\alpha} = 0$$

4.7 Tensors associated with the Riemann curvature tensor

Well, we have a bit of a problem. We want a tensor equation which has curvature=energy density. The stress energy tensor is second order, symmetric, with covariant derivative of zero. So we are looking for a way to denote curvature which is also a second order symmetric tensor with covariant derivative of zero. But the Riemann curvature tensor which completely embodies all the information about the curvature of space is a 4th order tensor. There is no way that a second order tensor can be equal to a 4th order one. So we need to CONTRACT the Riemann curvature tensor, preferably without losing any information about the spatial curvature!!!

If we were to contract R^a_{bcd} we could sum over one of the covariant indices with the contravariant one. But which covariant index - in principle $R^a_{acd} \neq R^a_{bad} \neq R^a_{bca}$. However, the Riemann curvature tensor has a large number of symmetries so instead of having N^4 components it actually has *only*(!) $N^2(N^2-1)/12$.

We could derive these out using the full form of the Riemann curvature tensor but its very long and tedious so instead we are going to use a trick. We can always define locally inertial coordinates, those in which the metric tensor is given by $\eta_{\alpha\beta}$ and the Christoffel symbols are zero, BUT THEIR DERIVATIVES ARE NOT - we can't transform gravity away in a global sense. I'd call this the primed frame but the equations will get very messy. BUT IN WHAT FOLLOWS WE ARE LOOKING AT THE COMPONENTS OF THE CURVATURE TENSOR IN THE LOCALLY INERTIAL FRAME.

$$R^a{}_{bcd} = (\partial_c \Gamma^a{}_{bd} - \partial_d \Gamma^a{}_{bc} + \Gamma^e{}_{bd}\Gamma^a{}_{ec} - \Gamma^e{}_{bc}\Gamma^a{}_{ed})$$
$$= \partial_c \Gamma^a{}_{db} - \partial_d \Gamma^a{}_{cb} \text{ in the locally inertial frame}$$

Cyclic Identity

So you can see that so if we cyclically perturb the covariant indices then we have

$$R^a{}_{cdb} = \partial_d \Gamma^a{}_{bc} - \partial_b \Gamma^a{}_{dc}$$
$$R^a{}_{dbc} = \partial_b \Gamma^a{}_{cd} - \partial_c \Gamma^a{}_{bd}$$

so add all these, remembering that christoffel symbols are symmetric

$$R^a{}_{bcd} + R^a{}_{cdb} + R^a{}_{dbc} = \partial_c \Gamma^a{}_{db} - \partial_d \Gamma^a{}_{cb} + \partial_d \Gamma^a{}_{bc} - \partial_b \Gamma^a{}_{dc} + \partial_b \Gamma^a{}_{cd} - \partial_c \Gamma^a{}_{bd} = 0$$

But $R^a{}_{bcd} + R^a{}_{cdb} + R^a{}_{dbc} = 0$ is a TENSOR equation. So even though we derived it in the local frame, it must be true in ALL frames!

Bianchi Identity

Take the covariant derivative of the Riemann curvature tensor - but in a frame where the Christoffel symbols are zero then this is the same as the normal derivative!

$$R^a{}_{bcd;e} = \partial_e(\partial_c \Gamma^a{}_{bd} - \partial_d \Gamma^a{}_{bc}) = \partial_e \partial_c \Gamma^a{}_{bd} - \partial_e \partial_d \Gamma^a{}_{bc}$$

cyclically permuting c, d, e gives

$$R^a{}_{bde;c} = \partial_c \partial_d \Gamma^a{}_{be} - \partial_c \partial_e \Gamma^a{}_{bd}$$

$$R^a{}_{bec;d} = \partial_d \partial_e \Gamma^a{}_{bc} - \partial_d \partial_c \Gamma^a{}_{be}$$

So

$$R^a{}_{bcd;e} + R^a{}_{bde;c} + R^a{}_{bec;d} = \partial_e \partial_c \Gamma^a{}_{bd} - \partial_e \partial_d \Gamma^a{}_{bc} + \partial_c \partial_d \Gamma^a{}_{be} - \partial_c \partial_e \Gamma^a{}_{bd} + \partial_d \partial_e \Gamma^a{}_{bc} - \partial_d \partial_c \Gamma^a{}_{be} = 0$$

and again, its a tensor equation, so holds in all frames not just the locally inertial frame.

Index Symmetries

Start with the fully covariant quantity

$$R_{abcd} = g_{ae} R^e{}_{bcd}$$

$$R_{abcd} = g_{ae}[\partial_c \Gamma^e{}_{db} - \partial_d \Gamma^e{}_{cb}]$$

we can take the g_{ae} inside the derivatives and get

$$R_{abcd} = \partial_c \Gamma_{adb} - \partial_d \Gamma_{acb}$$

these Christoffel symbols can be written in terms of the derivatives of the metric (see lecture 7) where we had

$$2\Gamma_{abc} = \partial_b g_{ac} + \partial_c g_{ba} - \partial_a g_{bc}$$

so

$$2R_{abcd} = \partial_c[\partial_d g_{ab} + \partial_b g_{da} - \partial_a g_{db}] - \partial_d[\partial_c g_{ab} + \partial_b g_{ca} - \partial_a g_{cb}]$$
$$R_{abcd} = \frac{1}{2}(\partial_c \partial_b g_{da} - \partial_c \partial_a g_{db} - \partial_d \partial_b g_{ca} + \partial_d \partial_a g_{cb})$$

where we've used the fact that $g_{ab} = g_{ba}$ and that $\partial_d \partial_c g_{ab} = \partial_c \partial_d g_{ab}$.

So then we can see that

$$R_{abcd} = -R_{bacd}$$
$$R_{abcd} = -R_{abdc}$$
$$R_{abcd} = R_{cdab}$$

Again, these are TENSOR equations, so must be true in all frames.

so because of all these symmetries then we can contract the Riemann curvature tensor to a second order tensor and not lose TOO much information.

4.8 Ricci Tensor

The index symmetries have some important implications for $R^a{}_{bcd}$. If we are contracting over the first index, $R^a{}_{acd}$ then we can see that $R^a{}_{acd} = g^{ae} R_{eacd} = -g^{ae} R_{aecd} = -g^{ea} R_{aecd} = -R^e{}_{eacd} = -R^a{}_{acd}$. The only way this can be true in general is if $R^a{}_{acd} = 0$. So this is not a useful index to contract over!

So now we have the choice of contracting over index 3 or 4. But $R_{abcd} = -R_{abdc}$ so $R^a{}_{bad} = -R^a{}_{bda}$. So modulo a sign change then there is only one non–zero contraction of the Riemann curvature tensor, which we call the **Ricci tensor**.

$$R_{ab} = R^c_{abc}$$

NB there is no widely accepted convention for the sign of the Riemann curvature tensor, or the Ricci tensor, so check the sign conventions of whatever book you are reading.

4.9 The Einstein Field Equations

The Ricci tensor is a second order tensor about curvature while the stress-energy tensor is a second order tensor about the source of gravity (energy density). So how about $R^{\mu\nu} = \kappa T^{\mu\nu}$ as the equation we are after which is curvature=gravity ? This in fact was Einsteins first guess!

well, stress energy tensor is also symmetric, and has covariant derivative of zero. The Ricci tensor is also symmetric. Contract the cyclic identity $R^a{}_{bcd} + R^a{}_{cdb} + R^a{}_{dbc} = 0$ so

$$R^a{}_{bca} + R^a{}_{cab} + R^a{}_{abc} = 0$$

but $R^a{}_{acd} = 0$ from above so this leaves

$$R^a{}_{bca} + R^a{}_{cab} = 0$$

$$R^a{}_{bca} - R^a{}_{cba} = 0 \text{ from } R_{abcd} = -R_{abdc}$$

$$R_{bc} = R_{cb}$$

So the Ricci tensor is symmetric, as required. Incidentally this means that $R^a{}_b = R_b{}^a$ as $R^a{}_c = g^{ab}R_{bc} = g^{ab}R_{cb} = R_c{}^a$. So we have two out of the 3 qualities we need for the stress energy tensor (second order and symmetric). But what about the third - covariant derivative of zero ?

We can use the Bianchi identity to get this for the Ricci tensor

$$R^a{}_{bca;e} + R^a{}_{bae;c} + R^a{}_{bec;a} = 0$$

$$R_{bc;e} - R_{be;c} + R^a{}_{bec;a} = 0$$

we know (section 3.10) that we can take the metric in and out of covariant differentiation ie that raising and lowering of indices can go in and out of covariant differentiation. So lets raise the index b - multiply it by g^{fb}.

$$g^{fb}R_{bc;e} - g^{fb}R_{be;c} + g^{fb}R^a{}_{bec;a} = 0$$

$$R^f{}_{c;e} - R^f{}_{e;c} + R^{af}{}_{ec;a} = 0$$

contract over e and f ie set $e = f$

$$R^f{}_{c;f} - R^f{}_{f;c} + R^{af}{}_{fc;a} = 0$$

Define $R = R^a{}_a = R_a{}^a$ as the Riemann curvature scalar. Then we have

$$R^f{}_{c;f} - R_{;c} + R^{af}{}_{fc;a} = 0$$

$$R^f{}_{c;f} - R_{;c} + R^{fa}{}_{cf;a} = 0 \quad \text{from } R_{abcd} = -R_{abdc} = R_{badc}$$

$$R^f{}_{c;f} - R_{;c} + R^a{}_{c;a} = 0$$

$$2R^a_{c;a} - R_{;c} = 0$$

$$(R^a_c - \frac{1}{2}R\delta^a_c)_{;a} = 0$$

$$(g^{bc}R^a_c - \frac{1}{2}g^{bc}\delta^a_c R)_{;a} = 0$$

$$(R^{ba} - \frac{1}{2}g^{ba}R)_{;b} = 0$$

so the covariant derivative of the Ricci tensor IS NOT ZERO! But this funny combination of the ricci tensor and curvature scalar IS.

So we have the Ricci Tensor, which is a symmetric second order tensor, but its divergance IS NOT zero. Instead we have this composite of the Ricci tensor and the curvature scalar which has divergance of zero. So its this whole thing which we could put as proportional to the stress-energy tensor.

$$R^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}R = \kappa T^{\alpha\beta}$$

These are the Einstein field equations. There are 10 of them - should be 4^2 but the tensors are SYMMETRIC. But this is not a unique way to relate curvature to gravity and there is no way to prove them a priori. In fact this was Einsteins second guess at it - his first was $R^{ab} = \kappa T^{ab}$ And its not the final guess either, even in this form. In effect what we have is a differential equation. So the solution has an integration constant in it. We get the equation above if this constant is zero. But the more general form is

$$R^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}R + \Lambda g^{\alpha\beta} = \kappa T^{\alpha\beta}$$

where Λ is the cosmological constant.

But lets stick with the simplest possible form of the simplest possible equation i.e. set $\Lambda = 0$. We can put the Einstein equations in slightly different form:

$$g_{\gamma\alpha}R^{\alpha\beta} - \frac{1}{2}g_{\gamma\alpha}g^{\alpha\beta}R = \kappa g_{\gamma\alpha}T^{\alpha\beta}$$

$$R^{\beta}_{\gamma} - \frac{1}{2}\delta^{\beta}_{\gamma}R = \kappa T^{\beta}_{\gamma}$$

contract over β and γ i.e. set $\beta = \gamma$ Then since $\delta^{\beta}_{\beta} = 4$ then

$$R - 2R = \kappa T$$

$$R = -\kappa T$$

So an alternative form (which we are going to use below) is

$$R^{\alpha\beta} = \kappa(T^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}T)$$

4.10 Solving the Einstein equations

Einstein thought it would never be done. His equation is a second order tensor equation - so represents 16 separate equations! Though the symmetry properties means there are 'only' 10 independent equations!!

But the way to solve it is not in full generality, but to pick a reql physical situation we want to represent. The simplest is static curved spacetime round a spherically symmetric mass while the rest of spacetime is empty. Schwarzchild did this by guessing the form the metric should have

$$c^2 d\tau^2 = A(r)dt^2 - B(r)dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2$$

so the $g_{\mu\nu}$ are not functions of t - field is static. And spherically symmetric as surfaces with r, t constant have $ds^2 = r^2(d\theta^2 + sin^2\theta d\phi^2)$. For EMPTY spacetime then $R_{\mu\nu} = 0$ NB just because the Ricci tensor is zero DOES NOT means that the Riemann curvature tensor components are zero (ie no curvature)!!

$$R_{\mu\nu} = \partial_\nu \Gamma^{\sigma}_{\mu\sigma} - \partial_\sigma \Gamma^{\sigma}_{\mu\nu} + \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\rho\nu} - \Gamma^{\rho}_{\mu\nu}\Gamma^{\sigma}_{\rho\sigma} = 0$$

Substitute in with $x^0 = t$, $x^1 = r$, $x^2 = \theta$, $x^3 = \phi$ and get (see homework)

$$R_{00} = -\frac{A''}{2B} + \frac{A'}{4B}(\frac{A'}{A} + \frac{B'}{B}) - \frac{A'}{rB} = 0$$

$$R_{11} = \frac{A''}{2A} - \frac{A'}{4A}(\frac{A'}{A} + \frac{B'}{B}) - \frac{B'}{rB} = 0$$
$$R_{22} = \frac{1}{B} - 1 + \frac{r}{2B}(\frac{A'}{A} - \frac{B'}{B}) = 0$$
$$R_{33} = R_{22}\sin^2\theta = 0$$

All the rest of the $R_{\mu\nu}$ are zero. The R_{33} is redundant information as its linearly related to R_{22} so we have 3 equations for our unknown A and B. Solve these using additional information that as $r \to \infty$ then $A \to c^2$ and $B \to 1$ so that it makes the flat space metric (in spherical polar coordinates). We also know that $A(r) = g_{tt} \approx c^2[1 - 2GM/(rc^2)]$ in the weak field limit (from lecture 8, NB there we were sensible and had $x^0 = ct$ whereas here we've had $x^0 = t$ so there is an extra factor of c^2). Put all these together and get

$$A = c^2(1 - 2\frac{GM}{rc^2}) \quad B = (1 - 2\frac{GM}{rc^2})^{-1}$$

so we have the line element

$$ds^2 = c^2 d\tau^2 = c^2(1 - \frac{2GM}{c^2 r})dt^2 - (1 - \frac{2GM}{c^2 r})^{-1}dr^2 - r^2 d\theta^2 - r^2\sin^2\theta d\phi^2$$

or set $GM/c^2 = m$

$$ds^2 = c^2 d\tau^2 = c^2(1 - \frac{2m}{r})dt^2 - (1 - \frac{2m}{r})^{-1}dr^2 - r^2 d\theta^2 - r^2\sin^2\theta d\phi^2$$