Introduction to General Relativity, AST 5220

Håvard Tveit Ihle

January 15, 2018

1 Tensors

The main motivation for using tensors is to write down equations that are invariant under coordinate transformations, that is, the equations look the same in any coordinate system. The reason we want to write down such equations is that we want our physical theories to be independent of the coordinates we use to describe (or express) these theories. Hence we will here study coordinate transformations, and how objects change onder such transformations in order to learn what tensors are, and why they are so useful.

1.1 3D-rotations

We want to understand what tensors are, and why they are so important, but let us first start with something more familiar.

Here are some equations we all know and love:

$$m\vec{a} = \vec{F}, \tag{1}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}. \tag{2}$$

These equations (vector equations) are nice because, in a sense, they are "always true". We may not always agree about what a_x is (e.g. if my x-axis points in a different direction than yours), but we will both agree that (1) is true.

Why is this?

I claim this is because $\vec{a}$ transforms in exactly the same way as $\vec{F}$ under rotations. Specifically, if:

$$\vec{x} \underbrace{\rightarrow}_{\text{"goes to"}} \vec{\tilde{x}} = \hat{R}\vec{x}, \tag{3}$$

then

$$\vec{a} \rightarrow \vec{\tilde{a}} = \hat{R}\vec{a}, \tag{4}$$

$$\vec{F} \rightarrow \vec{\tilde{F}} = \hat{R}\vec{F}, \tag{5}$$

ensuring that

$$m\vec{\tilde{a}} = \vec{\tilde{F}}, \tag{6}$$

where by $\vec{x}$ "goes to" $\vec{\tilde{x}}$, we mean that we change coordinates from $\vec{x}$ to $\vec{\tilde{x}}$. $\vec{\tilde{x}}$ is then given as a function of the old coordinates, here just by $\vec{x}$ multiplied by a rotation matrix $\hat{R}$.

The rotation matrix $\hat{R}$ is given by e.g. (for rotations in the $x - y$ plane):

$$\hat{R} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{7}$$

Einstein Summation Convention

It is useful to introduce the Einstein Summation Convention. This may seem a bit unnecessary at this point, but it will make our lives much simpler when we get to more complicated tensor expressions soon.

The basic idea is to write all tensors[a] on index form and to not bother writing summing signs. For example, eq 3 will become:

$$x^i \to \tilde{x}^i = R^i{}_j x^j \equiv \sum_j R^i{}_j x^j. \tag{8}$$

This convention works because (as we will learn) in tensor equations, indices are always summed in a very specific way.

The Einstein Summation Convention can be summarized in the following rules:

- If an index is repeated **on the same side** of an equation this index is summed over. I.e. the index j in eq 8.
- Always sum one upper and one lower index. If you find yourself wanting to sum over two upper or two lower indices, or find yourself wanting to sum more than two repeated indices, then something has probably gone wrong!
- Indices that are summed over (called *dummy indices*) can be changed to another index symbol. I.e. in the expression $T^i{}_j F^{jk}$, j can be changed to l, giving us $T^i{}_l F^{lk}$, which is an entirely equivalent expression. This often needs to be done to prevent using the same symbol for multiple repeated indices.
- The free indices (the ones that are **not** summed over) have to match in each term of a tensor equation. I.e.

  $$T^{ij} V_j = W^i F_k{}^k \tag{9}$$

  is fine, but

  $$T^{ij} V_j = W^l F_k{}^k, \tag{10}$$

  $$T^{ij} = W^i F_k{}^k, \tag{11}$$

  are not valid tensor equations. In the first case, eq. 10, the free index, i, on the left hand side does not show up on the right hand side, likewise the free index l only shows up on the right hand side. In the second case, eq. 11, the index j on the left does not have a match on the right.
- We use latin indices $i, j, k, l..$ to denote purely spatial indices. These take the values 1, 2, 3, denoting the three spatial dimensions. We use greek indices $\mu, \nu, \rho, \sigma..$ to denote spacetime indices. These take the values 0, 1, 2, 3, where 0 denotes the timelike dimension and 1, 2, 3 the spatial dimensions.

[a] We have not defined what a tensor is yet, but for now you can just think of matrices and vectors. We will soon enough learn exactly what a tensor is!

1.2 Vectors and Tensors in spacetime

What about vectors and tensors in spacetime? This is what we are really interested in. What is the relevant transformation for these tensors (analogous to rotations in space)?

In general relativity (GR), the transformations that the theory is unchanged under, are completely general coordinate transformations.

In spacetime we have a set of four coordinates

$$x^\mu = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \qquad \left(\underbrace{= \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}}_{\substack{\text{for special relativity} \\ \text{in cartesian coord.}}} \right). \tag{12}$$

A coordinate transformation is then

$$x^\mu \to \tilde{x}^\mu = \begin{pmatrix} \tilde{x}^0(x^0, x^1, x^2, x^3) \\ \tilde{x}^1(x^0, x^1, x^2, x^3) \\ \tilde{x}^2(x^0, x^1, x^2, x^3) \\ \tilde{x}^3(x^0, x^1, x^2, x^3) \end{pmatrix} \equiv \tilde{x}^\mu(x). \tag{13}$$

We define a spacetime vector[1], v^μ, as something that transforms in the following way under a transformation of the coordinates[2]:

$$v^\mu \to \tilde{v}^\mu = \underbrace{\frac{\partial \tilde{x}^\mu}{\partial x^\nu}}_{\text{Jacobian matrix}} v^\nu. \tag{14}$$

For now you will just have to take my word that this is a good definition for the transformation of a vector (but you can check yourself that e.g. the transformations in eq. 4 and 5 are special cases of this transformation rule). Note that under this definition x^μ does not transform like a vector! This is different from what we usually have under 3D rotations.

There is also another kind of vector, w_μ, with the index downstairs. This transforms slightly differently

$$w_\mu \to \tilde{w}_\mu = \underbrace{\frac{\partial x^\nu}{\partial \tilde{x}^\mu}}_{\text{inverse Jacobian}} w_\nu. \tag{15}$$

The Jacobian and it's inverse combine to give the Kronecker-delta, ${\delta^\mu}_\sigma = {\delta_\sigma}^\mu \equiv \delta^\mu_\sigma$, which is the tensor equivalent of the identity matrix

$$\frac{\partial \tilde{x}^\mu}{\partial x^\nu} \frac{\partial x^\nu}{\partial \tilde{x}^\sigma} = \delta^\mu_\sigma, \tag{16}$$

which essentially follows from the chain rule.

Let's look at the transformation of the combination

$$v^\mu w_\mu \to \tilde{v}^\mu \tilde{w}_\mu = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} v^\nu \frac{\partial x^\sigma}{\partial \tilde{x}^\mu} w_\sigma = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} \frac{\partial x^\sigma}{\partial \tilde{x}^\mu} v^\nu w_\sigma = \delta^\sigma_\nu v^\nu w_\sigma = v^\sigma w_\sigma = v^\mu w_\mu. \tag{17}$$

[1] When we say "vector" here (and later for tensors) we usually mean vector field, $v^\mu(x)$. That is, a field with a vector (or tensor) at each point in spacetime. The spacetime equivalents of 3D vector fields like the electric or magnetic field.

[2] For a comparison of different types of transformations, see figure 4 in the appendix.

We see that the combination $v^\mu w_\mu$ is invariant under a general coordinate transformation. Such a quantity we call a scalar, in general a scalar, ϕ, is something that transforms like

$$\phi \to \tilde{\phi} = \phi. \tag{18}$$

In general, a tensor is something with indices

$$T^{\mu\nu\cdots}{}_{\alpha\beta\cdots}, \tag{19}$$

transforming like

$$T^{\mu\nu\cdots}{}_{\alpha\beta\cdots} \to \tilde{T}^{\mu\nu\cdots}{}_{\alpha\beta\cdots} = \frac{\partial \tilde{x}^\mu}{\partial x^\rho}\frac{\partial \tilde{x}^\nu}{\partial x^\sigma}\cdots\frac{\partial x^\lambda}{\partial \tilde{x}^\alpha}\frac{\partial x^\delta}{\partial \tilde{x}^\beta}\cdots T^{\rho\sigma\cdots}{}_{\lambda\delta\cdots}. \tag{20}$$

So each upper index transforms with a Jacobian matrix, while each lower index transforms with an inverse Jacobian. So vectors with upper or lower indices are just tensors with one index.

1.3 Derivatives in spacetime

Since we have our nice spacetime coordinates, x^μ, it seems natural to define the spacetime derivative

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right). \tag{21}$$

By defining this derivative with a lower index we have sort of hinted that this should transform like a vector with a lower index, but this needs to be shown.

In order to see how the derivative transforms we need to act on different stuff, and then transform to new coordinates. The simplest thing to act on is a scalar, ϕ. So the question is: Does $\partial_\mu\phi$ transform like a vector?

$$\partial_\mu\phi \to \underbrace{\tilde{\partial}_\mu}_{\equiv\partial/\partial\tilde{x}^\mu}\,\tilde{\phi} = \underbrace{\frac{\partial x^\nu}{\partial \tilde{x}^\mu}\partial_\nu}_{\text{chain rule}}\,\phi. \tag{22}$$

We see that indeed $\partial_\mu\phi$ transforms as a vector with a lower index!

What about the derivative of a vector?

$$\partial_\mu A^\nu \to \tilde{\partial}_\mu\tilde{A}^\nu = \frac{\partial x^\alpha}{\partial \tilde{x}^\mu}\partial_\alpha\left(\frac{\partial \tilde{x}^\nu}{\partial x^\beta}A^\beta\right) \tag{23}$$

$$= \underbrace{\frac{\partial x^\alpha}{\partial \tilde{x}^\mu}\frac{\partial \tilde{x}^\nu}{\partial x^\beta}\partial_\alpha A^\beta}_{\text{regular tensor term}} + \underbrace{\frac{\partial x^\alpha}{\partial \tilde{x}^\mu}\frac{\partial^2 \tilde{x}^\nu}{\partial x^\alpha \partial x^\beta}A^\beta}_{\text{annoying extra term!}}. \tag{24}$$

We see that $\partial_\mu A^\nu$ actually does **not** transform as a tensor! The first term looks exactly like we expect a tensor to transform, but we also have the second term which ruins the fun.

This is pretty bad! We want to write all our physical theories as tensor equations to ensure that the physics is invariant under coordinate transformations, but we can't really write all of physics without using derivatives (you can try!).

What we can do is to define a new quantity (also not a tensor), $\Gamma^\mu{}_{\alpha\beta}$, called the *connection coefficients*, *Christoffel symbols* or just *Gamma's*, such that

$$\Gamma^\mu{}_{\alpha\beta} \to \tilde{\Gamma}^\mu{}_{\alpha\beta} = \frac{\partial \tilde{x}^\mu}{\partial x^\lambda}\frac{\partial x^\rho}{\partial \tilde{x}^\alpha}\frac{\partial x^\sigma}{\partial \tilde{x}^\beta}\Gamma^\lambda{}_{\rho\sigma} - \frac{\partial x^\rho}{\partial \tilde{x}^\alpha}\frac{\partial x^\sigma}{\partial \tilde{x}^\beta}\frac{\partial^2 \tilde{x}^\mu}{\partial x^\rho \partial x^\sigma}. \tag{25}$$

Using this new quantity we can define a new derivative, ∇_μ, called the *covariant derivative*

$$\nabla_\mu \phi \equiv \partial_\mu \phi, \tag{26}$$

$$\nabla_\mu A^\nu \equiv \underbrace{\overbrace{\partial_\mu A^\nu}^{\text{Not a tensor}} + \overbrace{\Gamma^\nu{}_{\mu\lambda} A^\lambda.}^{\text{Not a tensor}}}_{\text{A tensor!}} \tag{27}$$

The covariant derivative of a vector now transforms as a tensor

$$\begin{aligned}
\nabla_\mu A^\nu \to \tilde{\nabla}_\mu \tilde{A}^\nu &= \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial \tilde{x}^\nu}{\partial x^\beta} \partial_\alpha A^\beta + \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial^2 \tilde{x}^\nu}{\partial x^\alpha \partial x^\beta} A^\beta \\
&+ \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial \tilde{x}^\nu}{\partial x^\beta} \Gamma^\beta{}_{\alpha\lambda} A^\lambda - \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial^2 \tilde{x}^\nu}{\partial x^\alpha \partial x^\beta} A^\beta \qquad (28) \\
&= \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial \tilde{x}^\nu}{\partial x^\beta} \left(\partial_\alpha A^\beta + \Gamma^\beta{}_{\alpha\lambda} A^\lambda \right) \qquad (29) \\
&= \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial \tilde{x}^\nu}{\partial x^\beta} \left(\nabla_\alpha A^\beta \right). \qquad (30)
\end{aligned}$$

Likewise, we can show that

$$\nabla_\mu A_\nu \equiv \partial_\mu A_\nu - \Gamma^\lambda{}_{\mu\nu} A_\lambda \tag{31}$$

also transforms as a tensor (note the minus sign!).

The covariant derivative of a general tensor is then defined as you might expect, e.g.

$$\nabla_\mu T^{\alpha\beta}{}_\gamma \equiv \partial_\mu T^{\alpha\beta}{}_\gamma + \Gamma^\alpha{}_{\mu\lambda} T^{\lambda\beta}{}_\gamma + \Gamma^\beta{}_{\mu\lambda} T^{\alpha\lambda}{}_\gamma - \Gamma^\lambda{}_{\mu\gamma} T^{\alpha\beta}{}_\lambda. \tag{32}$$

So the covariant derivative of a tensor comes with one Γ-term for each index of the tensor, and the terms corresponding to an upper index come with a positive sign, while the terms corresponding to a lower index come with a negative sign.

We introduce the following simplified notation for the "regular" and covariant derivatives of tensors

$$T^{\alpha\beta}{}_{,\mu} \equiv \partial_\mu T^{\alpha\beta}, \tag{33}$$

$$T^{\alpha\beta}{}_{;\mu} \equiv \nabla_\mu T^{\alpha\beta}. \tag{34}$$

Great, we have now found a derivative that we can use on tensors and still end up with tensors. However, we still have no way to actually calculate $\Gamma^\mu{}_{\alpha\beta}$. It will turn out that the Γ's carry information about the curvature of spacetime (as well as information about the coordinates used), so we will have to talk about curvature before we can learn how to actually calculate the Γ's.

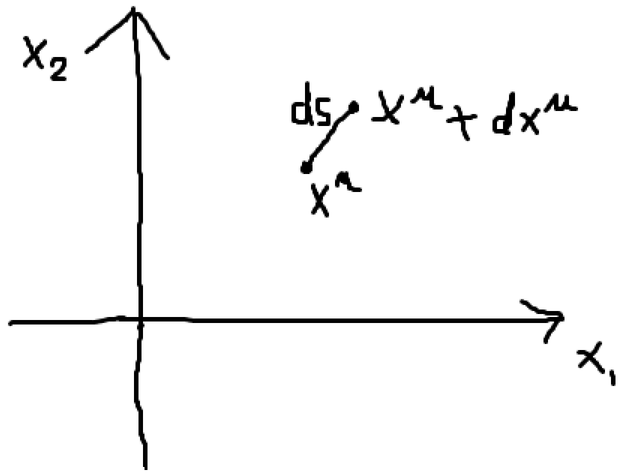

Figure 1: The metric tensor defines the spacetime interval, ds, between events x^μ and $x^\mu + dx^\mu$.

2 Curvature

2.1 The metric tensor

There is a special tensor, called the metric

$$g_{\mu\nu} = g_{\nu\mu}. \tag{35}$$

This tensor defines the infinitesimal spacetime interval, ds (see Figure 1), between two events (points), x^μ and $x^\mu + dx^\mu$ given by

$$ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu. \tag{36}$$

Let us look at some examples:

- The spatial interval on a 2-D sphere (with coordinates (θ, ϕ)) is given by

$$\begin{aligned} ds^2 &= g_{\theta\theta} d\theta^2 + g_{\theta\phi} d\theta d\phi + g_{\phi\theta} d\phi d\theta + g_{\phi\phi} d\phi^2 & (37) \\ &= d\theta^2 + \sin^2\theta d\phi^2. & (38) \end{aligned}$$

  I.e.

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}. \tag{39}$$

- In spacetime, the first example is the completely flat spacetime, called *Minkowski space*, which is the spacetime of special relativity

$$ds^2 = -dt^2 + \delta_{ij} dx^i dx^j, \tag{40}$$

  or

$$g_{\mu\nu} = \eta_{\mu\nu} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{41}$$

- As another example, take the (spatially) flat FRW metric

$$ds^2 = -dt^2 + a^2(t)\delta_{ij}dx^i dx^j, \tag{42}$$

implying a metric

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & a^2 & 0 & 0 \\ 0 & 0 & a^2 & 0 \\ 0 & 0 & 0 & a^2 \end{pmatrix}, \tag{43}$$

and the inverse metric (written just as g with upper indices)

$$g^{\mu\nu} \equiv (g^{-1})^{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1/a^2 & 0 & 0 \\ 0 & 0 & 1/a^2 & 0 \\ 0 & 0 & 0 & 1/a^2 \end{pmatrix}. \tag{44}$$

The metric tensor carries all the information about the curvature of spacetime, and it is the dynamical variable that is "solved for" in GR using the Einstein equation (which we will get to). This also means that the Christoffel symbols are determined by the metric[3]

$$\Gamma^\mu{}_{\alpha\beta} = \frac{1}{2} g^{\mu\nu} \left(g_{\alpha\nu,\beta} + g_{\beta\nu,\alpha} - g_{\alpha\beta,\nu} \right). \tag{45}$$

Another useful thing we can do with the metric tensor is to move indices up and down

$$V_\mu \equiv g_{\mu\nu} V^\mu, \tag{46}$$

$$T_\lambda{}^\mu \equiv g^{\mu\nu} T_{\lambda\nu}, \tag{47}$$

and so on. Since g is a tensor, the new tensors we get by moving indices up or down will also transform as we would expect based on the new positions of the indices.

Sometimes we wish to take a tensor with a given number of indices and make another tensor from it that has fewer indices. We also want to make sure to do this in a way such that the resulting object still transforms like a tensor.

One way to do this is to set an upper index equal to a lower index, and sum over all values. E.g.

$$S_\mu \equiv S_{\mu\lambda}{}^\lambda. \tag{48}$$

This is called *contracting* two indices, and it works out because upper and lower indices transform with the inverse transformations to each other.

What if we want to contract two lower or two upper indices? We can't sum over these (remember that we are only allowed to sum one upper and one lower index!), however, we can use the metric (or inverse metric). E.g.

$$T_\lambda \equiv g_{\mu\nu} T_\lambda{}^{\mu\nu}. \tag{49}$$

Note that we can not make a tensor with only one index less than the original tensor (using contraction), we always need to contract two indices at a time.

[3] Here we have assumed $\nabla_\mu g_{\alpha\beta} = 0$ (often called metric compatibility) and that the antisymmetric part of the connection coefficients, called the torsion, vanish, $\Gamma^\mu{}_{\alpha\beta} - \Gamma^\mu{}_{\beta\alpha} = 0$. The first of these assumptions is simply a common and useful convention, while the second is an important, but also very common, physical assumption.

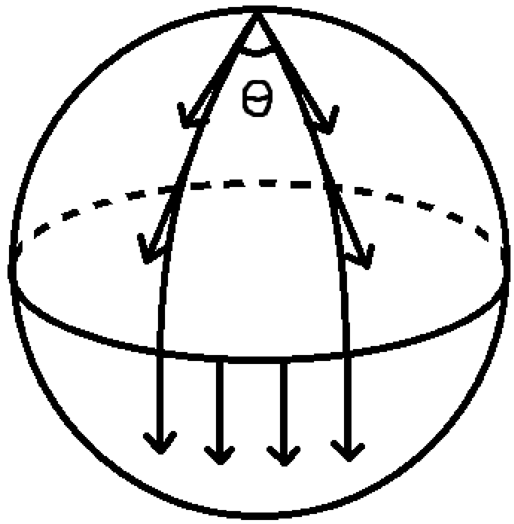

Figure 2: Parallel transport of a vector around a loop on a sphere. Even though the vector always stays parallel to itself, it has changed directions when it returns to it's original position.

2.2 Parallel transport and the Riemann curvature tensor

In flat spacetime *parallel transport*, that is, moving a vector from one point to another along a curve while keeping it constant, seems simple enough, we just move it without changing it. If we are using Cartesian coordinates we can just keep the components of the vector constant.

Written in mathematics, moving a vector along a curve $x^\mu(\lambda)$ while keeping the components constant, is ensured by requiring that the derivative with respect to λ vanishes,

$$\frac{d}{d\lambda}V^\mu = \frac{dx^\nu}{d\lambda}\frac{\partial}{\partial x^\nu}V^\mu = 0, \tag{50}$$

or just $V^\mu(x(\lambda_1)) = V^\mu(x(\lambda_0))$, where λ is a parameter monotoncally increasing along the curve.

If we are in curved spacetime (or curved space), this is not so trivial. This can be seen, for example, by moving a vector around on the surface of a sphere while keeping it parallel to itself.

If you start at the north pole with a vector pointing towards the equator, and you move it along the surface to the equator, then along the equator some angle θ and then back up to the north pole in a straight line, all while keeping the direction of the vector parallel to itself, you will see that the vector has changed! It now points in a different direction (in fact, the direction differs by the angle θ, that you moved along the equator.). See Fig. 2.

In a sense this is bad. We cannot uniquely compare (add, subtract, take dot product etc) vectors defined at different points in spacetime. We could of course parallel transport one vector to the other, but the vector you end up with would depend on the path you took between the points, so the combination would be entirely arbitrary.

On the other hand, we now have a method to detect curvature!

If someone hands you a complicated metric, and you want to know if the spacetime corresponding to the metric is curved, or if it is secretly just flat spacetime and your friend was using some weird coordinates, you can just check this by moving a vector around in this spacetime

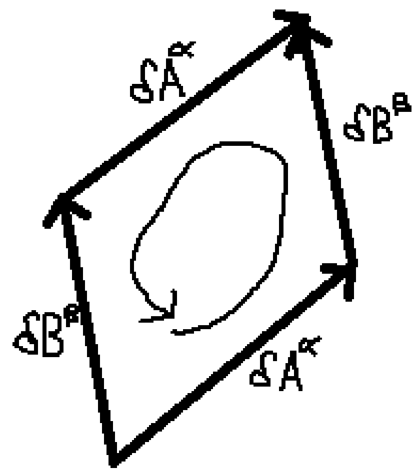

Figure 3: Moving a vector, V^μ, around loop along two infinitesimal vectors, δA^α and δB^β, and seeing how it changes due to the curvature, allows us to define the Riemann curvature tensor.

using parallel transport, and see if the vector changes. If the vector changes, the space has curvature, and if the vector stays the same, it means that the space is (probably) flat[4].

To generalize the equation for parallel transport (Eq. 50) to a case where we don't use Cartesian coordinates we just replace the spacetime derivative $\partial/\partial x^\nu$ by the covariant derivative ∇_ν. This way the equation becomes a valid tensor equation.

$$\frac{dx^\nu}{d\lambda}\nabla_\nu V^\mu = 0. \tag{51}$$

We argued that parallel transporting a vector around a curve and comparing to the original vector was a way for us to detect curvature. Let's apply this to a general infinitesimal curve, and see what we get. If we start with a vector V^μ and parallel transport it first along an infinitesimal vector δA^α, then along an infinitesimal vector δB^β, then backwards along δA^α and finally backwards along δB^β, so we are back where we started (see Fig 3).

The change in the vector V^μ from moving around this infinitesimal loop using parallel transport should then somehow contain information about the curvature of the spacetime. In addition we expect the change to be proportional to δA^α and δB^β, as well as V^μ itself. We then get:

$$V^\mu_{\text{after}} - V^\mu_{\text{before}} = R^\mu{}_{\nu\alpha\beta} V^\nu \delta A^\alpha \delta B^\beta, \tag{52}$$

where $R^\mu{}_{\nu\alpha\beta}$ is called the *Riemann curvature tensor*, or simply the *Riemann tensor*, and it tells us how spacetime is curved.

From Eq. 51 we have all the information we need to calculate $R^\mu_{\nu\alpha\beta}$ explicitly. It turns out, after a somewhat tedious derivation, that the Riemann tensor is given by the commutator of two covariant derivatives

$$R^\mu{}_{\nu\alpha\beta} V^\nu = [\nabla_\alpha, \nabla_\beta] V^\mu. \tag{53}$$

[4]Of course there are sometimes paths you can take in spacetime where your vector would happen to end up with the same components (for example if you move the vector in Fig. 2 around an angle $\theta = 2\pi$ at the equator, before you move it back up to the north pole.)

After another tedious derivation you can show that this reduces to

$$R^{\mu}{}_{\nu\alpha\beta} = \Gamma^{\mu}{}_{\nu\beta,\alpha} - \Gamma^{\mu}{}_{\nu\alpha,\beta} + \Gamma^{\mu}{}_{\alpha\lambda}\Gamma^{\lambda}{}_{\beta\nu} - \Gamma^{\mu}{}_{\beta\lambda}\Gamma^{\lambda}{}_{\alpha\nu}. \tag{54}$$

The Riemann tensor is then the ultimate test for flatness. If all the components of the Riemann tensor are zero, then the spacetime is flat.

3 General relativity

3.1 The Ricci tensor, the Ricci scalar and the Einstein equation

We can contract the first and third indices of the Riemann tensor in order to obtain the *Ricci tensor*

$$R_{\mu\nu} \equiv R^{\lambda}{}_{\mu\lambda\nu}. \tag{55}$$

Contracting the two indices if the Ricci tensor gives us the *Ricci scalar*

$$\mathcal{R} \equiv g^{\mu\nu} R_{\mu\nu}. \tag{56}$$

Together these tensors make up the *Einstein tensor*, $G_{\mu\nu}$, which is the left hand side of the *Einstein equation*

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \mathcal{R} = 8\pi G T_{\mu\nu}, \tag{57}$$

where $T_{\mu\nu}$ is the *Energy-momentum tensor*.

The Einstein equation relates the curvature of the universe, represented by $G_{\mu\nu}$, to the matter and energy content of the universe, represented by $T_{\mu\nu}$. So any distribution of matter and energy will create curvature in spacetime around it, in much the same way as any distribution of charge will create an electromagnetic field.

Let me note briefly that $G_{\mu\nu} = 0$ does *not* imply that spacetime is flat, it just means that there is no energy or momentum locally. In the same way that the electromagnetic field does not vanish outside a charge distribution, the curvature will also extend outside a distribution of energy. Curvature is only zero when all the components of the Riemann tensor vanishes ($R^{\mu}{}_{\nu\alpha\beta} = 0$).

In addition, there are gravitational wave solutions of the Einstein equations, which are solutions of the sourceless Einstein equation ($G_{\mu\nu} = 0$), analogous to the radiative solutions of the sourceless Maxwell equations ($\partial_{\mu} F^{\mu\nu} = 0$).

3.2 Minimal coupling and the Geodesic equation

We have found the equations governing the curvature of spacetime, the Einstein equation, but what about the rest of physics? How do we do physics in curved spacetime?

There is a simple prescription we can use to generalize physics to curved spacetime, called the *minimal-coupling principle*:

1. Take a law of physics valid in special relativity (SR).

2. Write it in tensorial form by making the replacements $\eta_{\mu\nu} \to g_{\mu\nu}$ and $\partial_{\mu} \to \nabla_{\mu}$.

Let us illustrate with some examples:

- Consider the conservation of energy and momentum. In SR this is equivalent to the four divergence of the energy momentum tensor vanishing:

  $$\partial_{\mu} T^{\mu\nu} = 0. \tag{58}$$

  In GR this equation becomes:

  $$\nabla_{\mu} T^{\mu\nu} = 0. \tag{59}$$

  Simple!

- As a more interesting example consider Newton's second law, which, in the absence of forces is given by

$$\frac{d^2x^\mu}{d\lambda^2} = 0. \tag{60}$$

This is not a tensor equation. The velocity, $\frac{dx^\mu}{d\lambda}$, is a vector, but the acceleration involves the derivative of a vector, which, as you may remember, does not transform as a tensor

$$\frac{d^2x^\mu}{d\lambda^2} = \frac{dx^\alpha}{d\lambda}\partial_\alpha \frac{dx^\mu}{d\lambda}, \tag{61}$$

where we used the chain rule.

Replacing the spacetime derivative by a covariant derivative gives us the general relativistic version Newtons second law (in the absence of forces)[5]

$$\frac{dx^\alpha}{d\lambda}\nabla_\alpha \frac{dx^\mu}{d\lambda} = \frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu{}_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0. \tag{62}$$

This is called the *geodesic equation*, and is the equation of motion for a particle moving in a curved spacetime, and will be used a lot by us during this course.

Looking at the left hand side of Eq. 62 we see that this can also be interpreted as the parallel transport equation (Eq. 51) for a curve where the velocity vector of the curve, $\frac{dx^\mu}{d\lambda}$, is parallelly transported along the curve itself. Such a curve will be in a sense as "straight as possible". What this means is that if you, at any point along the curve approximate spacetime as flat, and make a nice cartesian coordinate system, the curve will look straight in these coordinates. Such a "locally straight" curve is called a *geodesic*. So particles in curved spacetime move (in the absence of forces) along geodesics.

[5] If you have forces present, they can be added to the right hand side once they are written on covariant form.

4 Summary

We are now done with a short introduction to GR. Let us summarize the most important points:

- Tensors are great, because they change in specific and predictable ways under coordinate transformations, assuring that tensor equations look the same in any coordinate system. This is crucial for physics, because we don't want the physical laws to depend on what coordinate system we are using.

- In order to take derivatives of tensors (and for the result to transform as a tensor) we need to use the covariant derivative ∇_μ instead of the regular spacetime derivative ∂_μ. The covariant derivative of a tensor introduces a term with a connection coefficient (or Christoffel symbol) $\Gamma^\mu{}_{\alpha\beta}$ for each index of the tensor e.g.

$$\nabla_\mu T^{\alpha\beta}{}_\gamma = T^{\alpha\beta}{}_{\gamma;\mu} = T^{\alpha\beta}{}_{\gamma,\mu} + \Gamma^\alpha{}_{\mu\lambda} T^{\lambda\beta}{}_\gamma + \Gamma^\beta{}_{\mu\lambda} T^{\alpha\lambda}{}_\gamma - \Gamma^\lambda{}_{\mu\gamma} T^{\alpha\beta}{}_\lambda.$$

- The metric tensor $g_{\mu\nu}$ tells you how to measure distances in a given spacetime ($ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu$). The metric tensor, and it's inverse $g^{\mu\nu}$, can also be used to move indices of tensors up or down, and contract indices together e.g.

$$\begin{aligned} V_\mu &\equiv g_{\mu\nu} V^\mu, \\ T_\lambda{}^\mu &\equiv g^{\mu\nu} T_{\lambda\nu}, \\ S_\lambda &\equiv g_{\mu\nu} S_\lambda{}^{\mu\nu}. \end{aligned}$$

- The connection coefficients can be calculated from the first derivatives of the metric

$$\Gamma^\mu{}_{\alpha\beta} = \frac{1}{2} g^{\mu\nu} \left(g_{\alpha\nu,\beta} + g_{\beta\nu,\alpha} - g_{\alpha\beta,\nu} \right).$$

- The Riemann curvature tensor is the ultimate measure of the curvature of a spacetime

$$R^\mu{}_{\nu\alpha\beta} = \Gamma^\mu{}_{\nu\beta,\alpha} - \Gamma^\mu{}_{\nu\alpha,\beta} + \Gamma^\mu{}_{\alpha\lambda}\Gamma^\lambda{}_{\beta\nu} - \Gamma^\mu{}_{\beta\lambda}\Gamma^\lambda{}_{\alpha\nu}. \tag{63}$$

  Spactime is flat if and only if all the components of the Riemann curvature tensor are zero.

- From the Riemann curvature tensor we can get the Ricci tensor and Ricci scalar

$$\begin{aligned} R_{\mu\nu} &\equiv R^\lambda{}_{\mu\lambda\nu}, \\ \mathcal{R} &\equiv g^{\mu\nu} R_{\mu\nu}. \end{aligned}$$

  Together these tensors make up the Einstein tensor, $G_{\mu\nu}$, which is the left hand side of the Einstein equation

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \mathcal{R} = 8\pi G T_{\mu\nu}.$$

  $T_{\mu\nu}$ is the energy momentum tensor arising from all the content (i.e. matter and energy) of the spacetime. The Einstein equation is the equation of motion for the metric $g_{\mu\nu}$, and describes how matter and energy affects the curvature of spacetime.

- Particles in spacetime move (in the absence of forces) along "locally straight" curves, called geodesics, governed by the geodesic equation

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma^\mu{}_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} = 0.$$

Appendix

Comparison of different transformations

	3D - rotations	**Lorentz transformations**	**(General) Coordinate transformations**
How do coordinates transform?	$x^i \to \tilde{x}^i = R^i{}_j x^j$ For example: $R^i{}_j = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$x^\mu \to \tilde{x}^\mu = \Lambda^\mu{}_\nu x^\nu$ For example: $\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & & \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ or: $\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & R^i{}_j & \\ 0 & & & \end{pmatrix}$	$x^\mu \to \tilde{x}^\mu = \tilde{x}^\mu(x)$ Where: $\tilde{x}^\mu = \begin{pmatrix} \tilde{x}^0\left(x^0, x^1, x^2, x^3\right) \\ \tilde{x}^1\left(x^0, x^1, x^2, x^3\right) \\ \tilde{x}^2\left(x^0, x^1, x^2, x^3\right) \\ \tilde{x}^3\left(x^0, x^1, x^2, x^3\right) \end{pmatrix}$
How do vectors transform?	$v^i \to \tilde{v}^i = R^i{}_j v^j$	$v^\mu \to \tilde{v}^\mu = \Lambda^\mu{}_\nu v^\nu$	$v^\mu \to \tilde{v}^\mu = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} v^\nu$
Dot product?	$\mathbf{v} \cdot \mathbf{u} = \delta_{ij} v^i u^j$	$v \cdot u = \eta_{\mu\nu} v^\mu u^\nu$	$v \cdot u = g_{\mu\nu}(x) v^\mu u^\nu$

Figure 4: Comparison of different types of transformations, 3D-rotations, Lorentz transformations, and general coordinate transformations, and how vector transformations and invariant dot products are defined in each case.