Curved Spacetime (Riemannian Geometry) & Einstein Field Equations

October 2, 2019

0 Mathematical preliminaries

Here we review some concepts which apply to general differentiable manifolds before we specialize to Riemannian manifolds. Topics covered in this section can be found in any textbook on differential geometry and Riemannian geometry. Mathematicians introduced the concept of a manifold in order to generalize the Euclidean geometry in $\mathbb{R}^3$ or $\mathbb{R}^2$. One defines a manifold of dimension n, denoted M^n in the following way: A manifold, M^n is a topological space which is covered by a collection of open sets, say, U_i, $i = 1, 2, ...N$, for some positive integer, N, i.e.

$$M^n = \cup_{i=1}^{N} U_i$$

and each open set is homeomorphic to an open (set) ball in n-dimensional Euclidean hyperplane, $\mathbb{R}^n$,

$$U_i \cong \mathbb{R}^n.$$

A manifold is a topological space because one can have the notion of a neighborhood containing every "point" in the manifold, the neighborhoods being the open set, U_i's. As we have discussed in the class with the example of the two-sphere, S^2, one needs more than one coordinate systems or coordinate charts to cover a manifold (in case of sphere it was seen to be covered by at least 2 charts, one covering all points except the north pole, while the second chart covering everything except the south pole). However, in general case as we will talk about two such charts at a time. In the first chart we have points labeled by unprimed coordinates, $\{x^\mu\}$, while in the second chart we have the points labeled by primed coordinates, $\{x'^\mu\}$[1]. We then also demanded that for a point in the manifold which lies in the overlap region of of two such charts, the coordinates of the point in the two charts should be smooth functions of each other, i.e. if we label the same point, P in one chart by x and in another chart by x', then the coordinate transformation map

$$x'^\mu = x'^\mu(x) \tag{1}$$

is a ***smooth function*** of x and vice-versa. This smoothness means that all order derivatives exist (none of them diverge)

$$\frac{\partial^n x'}{\partial x^{\nu_1} \partial x^{\nu_2} \ldots \partial x^{\nu_n}} = \text{finite},$$

[1] The curly bracket, $\{x^\mu\}$ means that we actually have a list of objects indexed by μ,

$$\{x^\mu\} \equiv (x^0, x^1, x^2, \ldots)$$

Since the chart is homeomorphic to $\mathbb{R}^n$, the total number of objects in the list i.e. the total umber of distinct values that μ can take is also n. This is called the ***dimension*** of the manifold.

so that we have no problems while making Taylor series expansions of various kinds. Manifolds with such ***all order differentiable*** inter-coordinate transformation maps (C^∞) are called ***differentiable manifolds***. We will denote a differential manifold by the symbol M.

0.1 Calculus on Manifolds (Differential Geometry)

We are already familiar with multi-variable calculus on the Euclidean space, $\mathbb{R}^3$. To do calculus on an n-dimensional manifold, the first thing one has to introduce are ***functions*** on the manifold. The idea is if we pick at random, any point on the manifold, say P, there is a map which inputs P and spits out a real number in return, $f(P)$. This is a function, f on the manifold. Usually the point, P is labeled by its coordinates, so instead of $f(P)$ we will use the notation, $f(\{x^\mu\})$. Since there are n-number of coordinates which are all in the argument of the function, this is a multi-variable map, but following the abusive convention prevalent literature we will drop the curly variable, and denote the function as $f(x^\mu)$.

After introducing functions, the next natural thing is to introduce derivatives. But since such maps are multi-variable maps one needs to ask and answer "derivative in which direction?". To this end let's consider a one-dimensional curve (for example, the path of a point particle) c through the manifold and let's say that our friend, the point P which happens to lie on that curve. I should remind you that a 1D curve is parameterized by a parameter, call it λ, and each point on the curve the coordinates are functions of λ,

$$x^\mu = x^\mu(\lambda).$$

Now let's return to the question of taking derivatives of a function on the manifold and rephrase the question

"What is the rate of change of the function, f as we move along the curve $c(\lambda)$ at the point P?"

This rate of change of a function, f would be

$$\left.\frac{df}{d\lambda}\right|_P$$

i.e. ratio of change in the value of the function df as one moves and along the curve measured by change in the curve parameter value, $d\lambda$ around/about the point P. But $f(P)$ can be thought of as a function of the coordinates of P, x^μ_P. So using chain rule,

$$\left.\frac{df(x^\mu(\lambda))}{d\lambda}\right|_P = \left.\frac{dx^\mu(\lambda)}{d\lambda}\right|_P \frac{\partial f(x)}{\partial x^\mu}.$$

This can be written in a form,

$$\left.\frac{df(x^\mu(\lambda))}{d\lambda}\right|_P = V^\mu(x_P)\,\partial_\mu f$$

where,

$$V^\mu(x_P) = \left.\frac{dx^\mu(\lambda)}{d\lambda}\right|_P.$$

It is easy to realize these V^μ are nothing but components of the **tangent vector** along the curve at the point P.
(Recall similar stuff in your vector calculus course, where for a curve in 3d Euclidean space $\mathbb{R}^3$, you took a **directional derivative** of a function, $\phi(x)$ along the curve with tangent vector, $\hat{\boldsymbol{n}}$ by

$$\hat{\boldsymbol{n}} \cdot \boldsymbol{\nabla}\phi(x)$$

).

If we think about ***all*** one dimensional curves passing through the point P, we can think of the collection of the tangent vectors of each of those curves at P. One can show that this collection of tangent vectors form a linear vector space. This vector space located at P constitutes the ***Tangent Space*** of the manifold, M at the point P, denoted by, $T_P M$. Evidently a basis for this vector space, $T_P M$ is served by

$$\hat{e}_\mu = \frac{\partial}{\partial x^\mu} = \partial_\mu. \tag{2}$$

This basis of tangent vectors made out of partial derivative operators wrt to the coordinates is naturally called the ***coordinate basis.*** Since we are considering an n-dimensional manifold, the number of coordinates, x^μ is n. Then, the number of independent basis vectors, ∂_μ, of the tangent space is also n then i.e. $dim(T_P M) = n$.

However, note that I could have also conducted the analysis using a different coordinate system/chart covering the same point, P. In this new coordinate system, P is labeled by x'. This would mean that I am using a different coordinate basis,

$$\hat{e}'_\mu = \frac{\partial}{\partial x'^\mu} = \partial'_\mu$$

However the rate of change of the function along the curve, $\frac{df}{d\lambda}$ should be same in any coordinate system. Mathematically,

$$\frac{df}{d\lambda} = V^\nu \, \partial_\nu f = V'^\mu \, \partial'_\mu f.$$

One can check that this is only possible if the components of the tangent vector must transform under a general coordinate transformation (1) as follows,

$$V'^\mu(x') = \frac{\partial x'^\mu}{\partial x^\nu} V^\nu(x). \tag{3}$$

If one introduces a transformation matrix (also called the Jacobian matrix),

$$\underline{J^\mu{}_\nu(x)} = \begin{pmatrix} \frac{\partial x'^\mu(x)}{\partial x^\nu} & \cdots \\ \vdots & \end{pmatrix}, \tag{4}$$

then the above transformation can be written in a matrix form

$$V' = J.V$$

where, V is a **column vector** $\begin{pmatrix} V^1 \\ V^2 \\ \vdots \end{pmatrix}$ and so is V'. Similarly for the inverse transformation,

$$x^\mu = x^\mu(x'),$$

one can introduce the inverse Jacobian matrix,

$$\underline{\left(J^{-1}\right)^\mu{}_\nu} = \begin{pmatrix} \frac{\partial x^\mu(x')}{\partial x'^\nu} & \cdots \\ \vdots & \end{pmatrix}. \tag{5}$$

Analogously one can define, the dual vector space at each such point P on the manifold M, called the ***cotangent space*,** denoted by $T^*_P M$. Here, in brief, we recall the notion of the "dual (vector) space" , $V^*$of a given linear vector space, V . The dual space is defined to be the collection of all linear maps/functions which takes an input vectors, $v \in V$ can spits out a pure number (real or complex). Say ω is a member of the dual space, $\omega \in V^*$, and $v_{1,2} \in V$ are two vectors, then we have,

$$\begin{aligned} \omega[v_{1,2}] &= \text{real or complex number}, \\ \omega[v_1 + v_2] &= \omega[v_1] + \omega[v_2]. \end{aligned}$$

A typical example of such a linear map which turns vectors into pure numbers would be, $\theta^\mu :=$ which takes as the input a vector and it spits out the μ-th component". So if, $v = v^\alpha \hat{e}_\alpha$. Then,

$$\theta^\mu[v] = v^\mu.$$

One can easily check that this map is linear. In fact this example furnishes a basis for the dual space, $\left\{\hat{\theta}^\mu\right\}$, defined by their action on the basis vectors of the direct space,

$$\hat{\theta}^\mu[\hat{e}_\nu] = \delta^\mu_\nu.$$

In particular one can show that dual basis to the coordinate basis (2) in the direct space are given by coordinate differentials

$$\hat{\theta}^\mu = dx^\mu. \tag{6}$$

This follows from the following argument, under a coordinate change the duality relation should be unchanged, i.e.

$$\hat{\theta'}^\mu\left[\hat{e'}_\nu\right] = \hat{\theta}^\mu\left[\hat{e}_\nu\right] = \delta^\mu_\nu$$

For the coordinate basis, $\hat{e}_\nu = \partial_\nu$and $\hat{e'}_\nu = \partial'_\nu$, thus giving us,

$$\begin{gathered} \hat{\theta'}^\mu\left[\partial'_\nu\right] = \hat{\theta}^\mu\left[\partial_\nu\right], \\ \left(\frac{\partial x^\lambda}{\partial x'^\nu}\right)\hat{\theta'}^\mu\left[\partial_\lambda\right] = \delta^\lambda_\nu\,\hat{\theta}^\mu\left[\partial_\lambda\right] \end{gathered}$$

which leads to the transformation rule

$$\left(\frac{\partial x^\lambda}{\partial x'^\nu}\right)\hat{\theta'}^\mu = \delta^\lambda_\nu\,\hat{\theta}^\mu,$$

or,

$$\hat{\theta}'^{\mu} = \left(\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right) \hat{\theta}^{\nu}$$

It is now easy to see that the coordinate differentials satisfy such a transformation rule,

$$dx'^{\mu} = \left(\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right) \hat{dx}^{\nu}$$

Thus the dual space basis which is dual to the coordinate basis in the direct space are the coordinate differentials,

$$\hat{\theta}^{\mu} = dx^{\mu}.$$

In the coordinate basis, an element of the cotangent space at P, call it ω will be expanded,

$$\omega = \omega_{\mu}(x)\, dx^{\mu}.$$

Again, such an expression has to be invariant under general coordinate transformations, (1), i.e.

$$\omega_{\mu}(x)\; dx^{\mu} = \omega'_{\mu}(x')\; dx'^{\mu},$$

which is only possible if one has the transformation law

$$\omega_{\mu}(x) = \omega'_{\nu}(x')\; \frac{\partial x'^{\nu}}{\partial x^{\mu}}$$

or equivalently

$$\omega'_{\mu}(x') = \frac{\partial x^{\nu}}{\partial x'^{\mu}}\; \omega_{\nu}(x). \tag{7}$$

This can be again expressed in a matrix language,

$$\begin{aligned} \omega' &= \omega.J^{-1}, \\ \omega &= \omega'.J. \end{aligned}$$

where now the covariant vector, ω is a **row vector,** $\omega = \left(\omega^0, \omega^1, \omega^2, \ldots\right)$.

At this point that , it is easy to see we can cook up quantities using tangent vectors and dual vectors, which are invariant under coordinate transformations such as,

$$\omega[V] = \omega_{\mu} V^{\mu}.$$

In matrix shorthand,

$$\omega[V] = \omega.V,$$

which is fine because the rhs being a matrix product of a row vector and a column vector is a pure number: $(1, n) \times (n, 1) = (1, 1)$.

Theorem: ***Any two distinct finite dimensional linear vector spaces,*** V_1 ***and*** V_2 ***which have the same dimension, are isomorphic to each other.***

Corollary: *A vector space, V and it's dual vector space, V^* are isomorphic to each other. This isomporphism allows one to map/identify each and every vector, v to it's image vector dual vector, v^*:*

$$v^* \leftrightarrow v. \tag{8}$$

Using this map, one can define an inner product of two vectors, say v_1 and v_2 as follows:

$$(v_1, v_2) \equiv (v_1^*)_\mu \, v_2^\mu. \tag{9}$$

In fact, one can use this to define the norm of vector, $||v||$, by taking the inner product of a vector with itself,

$$||v||^2 = (v^*)_\mu \, v^\mu.$$

We realize this isomorphism between the vector and dual vector i.e. (8), in component form, by specifying a quantity called the ***metric***, $g_{\mu\nu}$

$$(v^*)_\mu = g_{\mu\nu} v^\nu. \tag{10}$$

Specifying a metric, turns the vector space into a metric space. In particular, specifying a metric for the tangent space, $T_P M$ at each point, P in a differentiable manifold turns it into a ***Riemannian Manifold***.

0.2 Lie derivative or Lie brackets

One can rewrite the transformation law of vectors and dual vectors, (3) and (7), in active transformation form as follows. Under a diffeomorphism/ general coordinate transformation,

$$x^\mu \to x'^\mu(x) = x^\mu + \xi^\mu(x),$$

a vector transforms as,

$$\begin{aligned} V^\mu(x) \to V'^\mu(x') &= \frac{\partial x'^\mu}{\partial x^\nu} V^\nu(x), \\ &= V^\mu(x) + \xi^\mu{}_{,\nu} V^\nu(x). \\ \implies V'^\mu(x) &= V^\mu(x-\xi) + \xi^\mu{}_{,\nu} V^\nu(x-\xi). \end{aligned}$$

So, one defines the Lie derivative, $\mathcal{L}_\xi V$ by the active transformation,

$$\begin{aligned} V'^\mu(x) - V^\mu(x) &= V^\mu(x-\xi) + \xi^\mu{}_{,\nu} V^\nu(x-\xi) - V^\mu(x) \\ &= V^\mu(x) - \xi^\nu \partial_\nu V^\mu(x) + O\left(\xi^2\right) + \xi^\mu{}_{,\nu} V^\nu(x-\xi) - V^\mu(x) \\ &= -\left(\xi^\nu \partial_\nu V^\mu - V^\nu \partial_\nu \xi^\mu\right) + O\left(\xi^2\right) \\ &\equiv -\left(\mathcal{L}_\xi V\right)^\mu + O\left(\xi^2\right) \end{aligned}$$

Similarly, from the transformation rule one can define the Lie-derivative of an one-form,

$$\begin{aligned}
\omega'_\mu(x) - \omega_\mu(x) &= \frac{\partial x^\nu}{\partial x'^\mu}\,\omega_\nu(x-\xi) - \omega_\mu(x) \\
&= -\xi^\nu \partial_\nu \omega_\mu - \omega_\nu \partial_\mu \xi^\nu + O\left(\xi^2\right) \\
&\equiv -\left(\mathcal{L}_\xi \omega\right)_\mu + O\left(\xi^2\right)
\end{aligned}$$

Although we can define a Lie-derivative for a one-form, we do not define the Lie-bracket for an one-form.

One can generalize this process to general (p,q)-type tensors. The results are summarized here,

$$\begin{aligned}
\left(\mathcal{L}_\xi V\right)^\mu &= \xi^\nu \partial_\nu V^\mu - V^\nu \partial_\nu \xi^\mu \\
\left(\mathcal{L}_\xi \omega\right)_\mu &= \xi^\nu \partial_\nu \omega_\mu + \omega_\nu \partial_\mu \xi^\nu \\
\left(\mathcal{L}_\xi T\right)^{\mu_1 \ldots \mu_p}{}_{\nu_1 \ldots \nu_q} &= \xi^\nu \partial_\nu\, T^{\mu_1 \ldots \mu_p}{}_{\nu_1 \ldots \nu_q} - T^{\rho \ldots \mu_p}{}_{\nu_1 \ldots \nu_q} \partial_\rho \xi^{\mu_1} - T^{\mu_1 \rho \ldots \mu_p}{}_{\nu_1 \ldots \nu_q} - \ldots \\
&\quad + T^{\mu_1 \ldots \mu_p}{}_{\rho \ldots \nu_q} \partial_{\nu_1} \xi^\rho + T^{\mu_1 \ldots \mu_p}{}_{\nu_1 \rho \ldots \nu_q} \partial_{\nu_2} \xi^\rho + \ldots
\end{aligned}$$

It is important to note that the Lie derivative, doesn't change the rank of the tensor on which it is acting, i.e. Lie derivative of a vector is a vector, Lie derivative of a one-form is another one-form, Lie derivative of a scalar is a scalar as well, and so on.

Homework:
a. Show that for a scalar the Lie derivative is just the directional derivative:

$$\mathcal{L}_\xi \phi = \xi^\mu \partial_\mu \phi.$$

b. Show that the Lie derivative obeys the Leibniz rule for differentiation,

$$\mathcal{L}_\xi \left(V_1 \otimes V_2\right) = V_1 \otimes \left(\mathcal{L}_\xi V_2\right) + V_1 \otimes \left(\mathcal{L}_\xi V_2\right).$$

where, $V_{1,2}$ are tangent space vectors.

The action of Lie-derivative on vectors, also allows us to define a new binary operation in the Tangent space, namely, the ***Lie bracket*** of any two vectors,

$$\left[V_1, V_2\right] \equiv \mathcal{L}_{V_1} V_2$$

Homework:
Show that the Lie-bracket obeys the following three properties
a. Linearity:

$$\left[\xi, V_1 + V_2\right] = \left[\xi, V_2\right] + \left[\xi, V_2\right].$$

b. Anti-symmetry:

$$\left[V_1, V_2\right] = -\left[V_2, V_1\right].$$

c. Jacobi Identity:

$$[V_1,[V_2,V_3]]+[V_2,[V_3,V_1]]+[V_3,[V_1,V_2]]=0.$$

One can show that the commutator of two successive diffeomorphisms, η and ξ, is a third diffeomorphism given by the Lie bracket of η and ξ.

$$\mathcal{L}_\eta\left(\mathcal{L}_\xi V\right)-\mathcal{L}_\xi\left(\mathcal{L}_\eta V\right)=\mathcal{L}_{[\eta,\xi]}V.$$

In math (differential geometry) literature, the notion of a Lie derivative follows naturally from **integral curves of a (smooth) vector field**. These curves are such that the tangent vector at each point on the curve coincides with the vector field. Such integral curves represent a map from the manifold to itself, which further induces a map between tangent spaces (and cotangent spaces) all along the curve[2] . The Lie derivative of a smooth vector field, Y along the integral curve of another smooth vector field X is defined to be the Lie derivative, $\mathcal{L}_X Y$ by taking the difference of the vector field Y with the induced map (pushforward) vector of Y along an integral curve of X:

$$\mathcal{L}_X Y=\lim_{\delta\to 0}\frac{Y(x+\delta)-f_*Y(x+\delta)}{\delta}$$

Here x and $x+\delta$ are two points on the integral curve of X, given by the curve equation, $f(x)=x+\delta$. Using the formula for pushforward maps, one can check and the above equation coincides with the expression found before,

$$\mathcal{L}_X Y=X^\nu\partial_\nu Y^\mu-Y^\nu\partial_\nu X^\mu.$$

Once we have defined Lie derivatives for vector fields, one can generalize this notion to define Lie derivatives for dual vectors and higher rank tensors by demanding that the Lie derivative obeys Leibniz rule and that the Lie derivative must act on a scalar like a directional partial derivative. Thus given a smooth vector field, say X on can introduce a notion of differentiation along the integral curves of X. However, in many situations one might not have *any* smooth vector field supported on the manifold (e.g. Hairy ball theorem for compact manifolds). Thus we will opt for a different notion of differentiation in the coming sections, namely the covariant derivative by introducing an independent additional data, dubbed as the ***affine connection*** for the manifold (including compact manifolds such as S^2).

[2] In general, a (smooth) map between points on two different (or different points on the same) manifold(s) in turn induces a map between tangent (cotangent) spaces at the two points, popularly referred to as ***pushforward (pullback)*** maps. Let's say f is a smooth map from the manifold, $\mathbb{M}^m$ to the manifold $\mathbb{N}^n$. Then the *pushforward* map $f_*:T_P\mathbb{M}\to T_{f(P)}\mathbb{N}$ is defined by,

$$[f_*V(y)]^\mu=V^\nu(x)\,\frac{\partial y^\mu}{\partial x^\nu}.$$

Here x^μ, y^μ are the coordinates of the point P in $\mathbb{M}$ and $f(P)$ on $\mathbb{N}$. V is a vector field defined on $\mathbb{M}$. The pullback map then follows,

$$[f_*\omega(y)]_\mu=\omega_\nu\frac{\partial x^\nu}{\partial y^\mu}.$$

1 Riemannian Manifolds

Mathematically, the curved spacetime of general relativity is a *Riemannian manifold* i.e. a differential manifold, M, which comes with an extra feature called the metric (g), which allows us to define distances. A Riemannian manifold is usually denoted as (M, g). Two nearby points in a Riemannian manifold with coordinates, $\{x^\mu\}$ and $\{x^\mu + dx^\mu\}$ are separated by a "distance" aka spacetime interval,

$$ds^2 = g_{\mu\nu}(x)\, dx^\mu\, dx^\nu. \tag{11}$$

The same information is also presented in a matrix form i.e. metric g is written as a matrix, with the element in the μ-th row and ν-th column being $g_{\mu\nu}$

$$\underline{g_{\mu\nu}} = \begin{array}{cc} & \nu\text{-th column} \\ & \downarrow \\ \mu\text{–th row}\rightarrow & \begin{pmatrix} & & \\ & g_{\mu\nu} & \\ & & \end{pmatrix} \end{array}$$

The metric, $g_{\mu\nu}$ can be taken to be symmetric in its indices by rewriting (11) a bit, as we show in the next few lines

$$\begin{aligned} ds^2 &= g_{12}dx^1dx^2 + g_{21}dx^2dx^1 + ... \\ &= (g_{12} + g_{21})\, dx^1dx^2 + \dots \\ &= \tilde{g}_{12}dx^1dx^2 + \tilde{g}_{12}dx^1dx^2 + \dots \\ &= 2\widetilde{g_{12}}dx^1dx^2 \end{aligned}$$

where

$$\tilde{g}_{12} = \tilde{g}_{21} = \frac{g_{12} + g_{21}}{2}$$

.

From now on we will take the metric to be in symmetric form only. In general for a vector space, the metric is defined to be a $(0, 2)$ rank tensor satisfying the following conditions:

$$g(U, V) = g(V, U)(\text{symmetry}) \tag{12}$$

$$g(U, U) \geq 0, (\text{norm positivity}) \tag{13}$$

where the equality holds iff $U = 0$, the zero vector, i.e. a *symmetric nondegenerate* (0,2) rank tensor.

However for tangent space vectors in Lorentzian spacetimes (see insert next page), norm positivity condition (13) is not true and instead one has to modify it little bit to define a nondegenerate metric in the following manner: if $g(U, V) = 0, \forall V$, then $U = 0$, the zero vector. One can check that this criterion disallows zero eigenvalues of the metric and keeps it invertible.

Normally, the square of the distance between two points in space is a positive quantity, unless the points are coincident in which case the distance is zero:

$$g_{\mu\nu} dx^\mu dx^\nu \geq 0$$

and hence in most spaces the metric tensor as a matrix would only have positive eigenvalues. However, in physics we need to include time as a coordinate, and we have seen before in special relativity, the square of the "spacetime distance" between two points which are *timelike separated* (i.e. have same spatial coordinates, $d\mathbf{x} = 0$ but different times of occurrences, $dt \neq 0$), is negative,

$$ds^2 = -dt^2 + d\mathbf{x}^2 = -dt^2.$$

Such spaces which allow for negative eigenvalues of the metric, $g_{\mu\nu}$ are called *Pseudo-Riemannian manifolds* by mathematicians and *Lorentzian spacetimes* by physicists. They are defined by,

$$\begin{aligned} g_{\mu\nu} dx^\mu dx^\nu &> 0, \text{ for spacelike separated points} \\ g_{\mu\nu} dx^\mu dx^\nu &= 0, \text{ for coincident or "lightlike/null" separated points} \\ g_{\mu\nu} dx^\mu dx^\nu &< 0, \text{ for timelike separated points} \end{aligned}$$

The squared line element, (11) must amount to the exact same quantity, regardless of whatever coordinate chart (x or x') we use to compute it. This implies

$$g'_{\mu\nu}(x')\, dx'^\mu\, dx'^\nu = g_{\mu\nu}(x)\, dx^\mu\, dx^\nu,$$

or equivalently,

$$g'_{\mu\nu}(x') = \frac{\partial x^\alpha}{\partial x'^\mu}\, \frac{\partial x^\beta}{\partial x'^\mu}\, g_{\alpha\beta}(x). \tag{14}$$

Thus the metric turns out to be a tensor, to be precise, a symmetric $(0, 2)$ rank tensor. As such it can be expanded in the tensor product basis for (0,2) tensors, namely, $dx^\mu \otimes dx^\nu$ in the following way

$$g = g_{\mu\nu}(x)\, dx^\mu \otimes dx^\nu.$$

However unlike in the Minkowski case, it is not an invariant tensor anymore - it changes to a new function g'.

Note that the transformation law (14) for the metric tensor components can also be expressed in a matrix notation using the inverse Jacobian,

$$g' = J^{-T}\, g\, J^{-1}. \tag{15}$$

where $J^{-T} \equiv \left(J^{-1}\right)^T$.

Homework: Consider two dimensional flat space (just space and no time), also known as 2d Euclidean space, $\mathbb{R}^2$. One can either use the *Cartesian* coordinates to label points in $\mathbb{R}^2$ i.e. (x, y), or one can use the plane polar coordinate system (r, θ) to label (almost) all the points in $\mathbb{R}^2$. The distance between two points in $\mathbb{R}^2$ say (r, θ) and $(r + dr, \theta + d\theta)$ is given by the familiar "squared line element",

$$ds^2 = dr^2 + r^2 \, d\theta^2.$$

So the metric, g is given by the components,

$$g_{rr} = 1, \; g_{\theta\theta} = r^2, \; g_{r\theta} = g_{\theta r} = 0.$$

In matrix form,

$$\underline{g_{\mu\nu}} = \begin{pmatrix} g_{rr} & g_{r\theta} \\ g_{\theta r} & g_{\theta\theta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}.$$

Check that under the coordinate transformations,

$$\begin{aligned} x &= r \, \cos\theta, \\ y &= r \, \sin\theta, \end{aligned}$$

the metric transforms to,

$$\underline{g_{\mu\nu}} = \begin{pmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

i.e. to say, the metric in Cartesian coordinates is,

$$ds^2 = dx^2 + dy^2.$$

(Hint: First construct the Jacobian matrix, $J = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix}$ and use the transformation law (15)).

1.1 Lowering and raising indices

The metric tensor allows us to map contravariant vectors to covariant vectors, the so called "lowering the index",

$$V_\mu(x) = g_{\mu\nu}(x) V^\nu(x)$$

while the inverse metric, $g^{\mu\nu}$, defined by,

$$g^{\mu\nu}(x) \, g_{\nu\lambda}(x) = \delta^\mu_\lambda$$

can be used to convert covariant vectors to contravariant vectors by the "raising the index" operation,

$$\omega^\mu(x) = g^{\mu\nu}(x) \, \omega_\nu(x).$$

1.2 Invariant tensors

From the homework, it is clear that the metric is not an invariant tensor under a general coordinate transformation. In the homework we transformed from plane polar to Cartesian coordinate system and the metric expressions (matrix) looked completely different. But one can check that the Kronecker delta continues to be an invariant tensor,

$$\begin{aligned}
\delta^\mu_\nu \to \delta'^\mu_\nu &= \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\nu}\,\delta^\alpha_\beta \\
&= \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial x'^\nu} \\
&= \frac{\partial x'^\mu}{\partial x'^\nu} \\
&= \delta^\mu_\nu .
\end{aligned}$$

What about the Levi Civita symbol, $\varepsilon_{\mu\nu\rho\sigma}$? Is this a tensor at all? and if it is a tensor, is it an invariant tensor under general coordinate transformations? Recall that we had defined, the Levi-Civita symbol by the rule:

$$\varepsilon_{\mu\nu\rho\sigma} = (-)^P. \tag{16}$$

where, P demotes the number of swaps one needs to turn 0123 to $\mu\nu\rho\sigma$. Similarly, we defined another Levi-Civita symbol with upstairs indices,

$$\varepsilon^{\mu\nu\rho\sigma} = \eta^{\mu\alpha}\eta^{\nu\beta}\eta^{\rho\gamma}\eta^{\sigma\delta}\varepsilon_{\alpha\beta\gamma\delta}.$$

The Levi-Civita symbol turned out to be invariant tensor under Lorentz transformations. But in the following we will see that unfortunately under the general coordinate transformation, the Levi-Civita symbol is non even a tensor, let alone an invariant tensor (in fact it is something we call a tensor-density). But before we proceed we also need to consider the transformation of the metric determinant, g, or more suitably the determinant of the inverse of metric, $g^{\mu\nu}$ under a GCT,

$$\begin{aligned}
\left(g'\right)^{-1} &= \frac{1}{4!}\varepsilon_{\mu\nu\rho\sigma}\ \varepsilon_{\alpha\beta\gamma\delta}\ g'^{\mu\alpha}\ g'^{\nu\beta}\ g'^{\rho\gamma}\ g'^{\sigma\delta} \\
&= \frac{1}{4!}\varepsilon_{\mu\nu\rho\sigma}\ \varepsilon_{\alpha\beta\gamma\delta}\ \left(J^\mu{}_{\bar\mu}\ J^\alpha{}_{\bar\alpha}\ g^{\bar\mu\bar\alpha}\right)\left(J^\nu{}_{\bar\nu}\ J^\beta{}_{\bar\beta}\ g^{\bar\nu\bar\beta}\right)\left(J^\rho{}_{\bar\rho}\ J^\gamma{}_{\bar\gamma}\ g^{\bar\rho\bar\gamma}\right)\left(J^\sigma{}_{\bar\sigma}\ J^\delta{}_{\bar\delta}\ g^{\bar\sigma\bar\delta}\right) \\
&= \frac{1}{4!}\left(\varepsilon_{\mu\nu\rho\sigma}\ J^\mu{}_{\bar\mu}J^\nu{}_{\bar\nu}J^\rho{}_{\bar\rho}J^\sigma{}_{\bar\sigma}\right)\left(\varepsilon_{\alpha\beta\gamma\delta}\ J^\alpha{}_{\bar\alpha}J^\beta{}_{\bar\beta}J^\gamma{}_{\bar\gamma}J^\delta{}_{\bar\delta}\right) g^{\bar\mu\bar\alpha}\ g^{\bar\nu\bar\beta}\ g^{\bar\rho\bar\gamma}\ g^{\bar\sigma\bar\delta} \\
&= J^2\ \frac{1}{4!}\varepsilon_{\bar\mu\bar\nu\bar\rho\bar\sigma}\ \varepsilon_{\bar\alpha\bar\beta\bar\gamma\bar\delta}\ g^{\bar\mu\bar\alpha}\ g^{\bar\nu\bar\beta}\ g^{\bar\rho\bar\gamma}\ g^{\bar\sigma\bar\delta} \\
&= J^2 g^{-1}.
\end{aligned} \tag{17}$$

Here, J is the determinant of the Jacobian. This implies,

$$g' = J^{-2}g. \tag{18}$$

Such quantities which transform under GCT with various factors/powers of the Jacobian determinant, J hanging around are called ***tensor densities***. Now we are in a position to figure out the transformation rule for ε under a GCT. Again, our starting point is the definition of the determinant of the Jacobian inverse (we are denoting the determinant of the Jacobian matrix by J),

$$\varepsilon_{\alpha\beta\gamma\delta}\, J^{-1} = \varepsilon_{\mu\nu\rho\sigma} \left(J^{-1}\right)^{\mu}{}_{\alpha} \left(J^{-1}\right)^{\nu}{}_{\beta} \left(J^{-1}\right)^{\rho}{}_{\gamma} \left(J^{-1}\right)^{\sigma}{}_{\delta} \tag{19}$$

Now look carefully at the rhs of the above equation. It looks like we have a transformation rule of ε as a $(0,4)$-type tensor. But the LHS is not ε but has a factor of J hanging around. So the Levi-Civita symbol is not a tensor but a tensor density as well. Similarly from the determinant identity,

$$\varepsilon^{\mu\nu\rho\sigma}\, J = J^{\mu}{}_{\alpha}\, J^{\nu}{}_{\beta}\, J^{\rho}{}_{\gamma}\, J^{\sigma}{}_{\delta}\, \varepsilon^{\alpha\beta\gamma\delta},$$

we note that he Levi-Civita symbol with all four upstairs indices, $\varepsilon^{\alpha\beta\gamma\delta}$ is a tensor-density as well.

1.3 "New and Improved" Levi-Civita tensor : Invariant Levi-Civita tensor for curved spacetime

Recalling the way the determinant of the inverse metric transformed, we can perhaps cook up an invariant antisymmetric tensor,

$$\epsilon_{\mu\nu\rho\sigma} \overset{?}{\equiv} \frac{1}{\sqrt{g}}\, \varepsilon_{\mu\nu\rho\sigma}$$

In Lorentzian spacetimes, $g < 0$, so we need to put an extra minus sign inside before we take a square root. So,

$$\epsilon_{\mu\nu\rho\sigma} \equiv \frac{1}{\sqrt{-g}}\, \varepsilon_{\mu\nu\rho\sigma}. \tag{20}$$

Perhaps similarly we can define another separate invariant tensor from the $\varepsilon^{\alpha\beta\gamma\delta}$, say

$$\tilde{\epsilon}^{\alpha\beta\gamma\delta} = \sqrt{-g}\, \varepsilon^{\alpha\beta\gamma\delta}$$

What about when we try to raise all 4 indices to turn it into, $\epsilon^{\mu\nu\rho\sigma}$? Let's raise the indices and see what happens,

$$\begin{aligned}
\epsilon^{\mu\nu\rho\sigma} &= g^{\mu\alpha} g^{\nu\beta} g^{\rho\gamma} g^{\sigma\delta}\, \epsilon_{\alpha\beta\rho\sigma} \\
&= \frac{1}{\sqrt{-g}}\, g^{\mu\alpha} g^{\nu\beta} g^{\rho\gamma} g^{\sigma\delta}\, \varepsilon_{\alpha\beta\rho\sigma} \\
&= \frac{1}{\sqrt{-g}}\, \left(-\varepsilon^{\mu\nu\rho\sigma}\, g\right) \\
&= \sqrt{-g}\, \tilde{\varepsilon}^{\mu\nu\rho\sigma} \\
&= \tilde{\epsilon}^{\alpha\beta\gamma\delta}!
\end{aligned}$$

So we don't need a separate completely antisymmetric tensor with the tilde!

$$\epsilon^{\mu\nu\rho\sigma} = \sqrt{-g}\, \tilde{\varepsilon}^{\alpha\beta\gamma\delta}. \tag{21}$$

1.4 Invariant Volume element

Recall that we defined the volume element by the 4-form in Minkowski spacetime

$$d^4x = dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3.$$

Again, we are interested in knowing how does this volume element transform under GCT. To do this, we first rewrite it as using the Levi Civita symbol, $\varepsilon_{\mu\nu\rho\sigma}$

$$d^4x = \varepsilon_{\mu\nu\rho\sigma}\, dx^\mu \otimes dx^\nu \otimes dx^\rho \otimes dx^\sigma,$$

In the primed coordinate frame,

$$\begin{aligned} d^4x' &= \varepsilon_{\mu\nu\rho\sigma}\, dx'^\mu \otimes dx'^\nu \otimes dx'^\rho \otimes dx'^\sigma \\ &= \varepsilon_{\mu\nu\rho\sigma}\, J^\mu{}_\alpha\, J^\nu{}_\beta\, J^\rho{}_\gamma\, J^\sigma{}_\delta\, dx^\alpha \otimes dx^\beta \otimes dx^\gamma \otimes dx^\delta \\ &= \varepsilon_{\alpha\beta\gamma\delta}\, J\, dx^\alpha \otimes dx^\beta \otimes dx^\gamma \otimes dx^\delta \\ &= J\, d^4x. \end{aligned} \tag{22}$$

So the volume element is a tensor-density as well. (This is the proof of the volume element transformation you learn in multi-variable calculus). From this transformation of the volume element, (22) and recalling the transformation rule for the metric determinant, (18), we see that we can define an invariant volume element,

$$dV \equiv \sqrt{-g} d^4x. \tag{23}$$

2 Covariant Calculus: Affine connection

In the case of Minkowski spacetime, $\mathbb{R}^{1,3}$, we have seen that the derivative of a (p,q)-type tensor, $\partial_\mu T$ is a tensor as well under Poincaré transformations - in fact the derivative turns it into a $(p,q+1)$-type tensor. This fact was crucial when we wanted to make equations of physics covariant under Poincaré transformations because most equations in physics express “rate of change” of physical quantities,

$$\partial_\mu T = X$$

where X is a $(p,q+1)$-type tensor while, T is a (p,q)-type tensor .

Now the question we are confronted with at this point is - do derivatives of tensor fields in a generic (Pseudo)-Riemannian spacetime, transform as tensors under a *general* coordinate transformation (GCT). To this end let’s check how a derivative of a vector field transforms under a GCT,

$$\begin{aligned} \partial_\mu V^\nu(x) \to \partial'_\mu V'^\nu(x') &= \frac{\partial x^\alpha}{\partial x'^\mu}\, \partial_\alpha \left(\frac{\partial x'^\nu}{\partial x^\beta}\, V^\beta(x) \right) \\ &= \frac{\partial x^\alpha}{\partial x'^\mu}\, \frac{\partial x'^\nu}{\partial x^\beta}\, \partial_\alpha V^\beta(x) + \frac{\partial x^\alpha}{\partial x'^\mu}\, \frac{\partial^2 x'^\nu}{\partial x^\alpha \partial x^\beta}\, V^\beta(x). \end{aligned} \tag{24}$$

So we observe that the derivative of a vector ***does not*** transform homogeneously under a GCT, and in fact contains a non-tensor piece, namely, the second term

$$\frac{\partial x^\alpha}{\partial x'^\mu}\, \frac{\partial^2 x'^\nu}{\partial x^\alpha \partial x^\beta}\, V^\beta(x) \tag{25}$$

in addition to nice tensor like piece,

$$\frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\beta}\,\partial_\alpha V^\beta(x).$$

So as of now we cannot write down rate equations in a covariant form because the derivative fails to be covariant. For example, the continuity equation,

$$\partial_\mu j^\mu = 0 \rightarrow \left(\partial'_\mu j'^\mu + \text{non-tensor junk}\right) = 0,$$

changes to something unpleasant in appearance.

How do we remedy the situation? The idea is that maybe we can define/cook up a ***covariant derivative***, ∇, by adding some non-tensor piece, say $\Gamma^\nu_{\mu\lambda}$ to an ordinary partial derivative which will cancel out the non-tensor piece (25), which arises when carry out do a GCT , i.e. define,

$$\nabla_\mu V^\nu \equiv (\nabla_\mu V)^\nu = \partial_\mu V^\nu + \underline{\Gamma^\nu_{\mu\lambda}\,V^\lambda}. \tag{26}$$

and demand that under a GCT,

$$\nabla_\mu V^\nu \rightarrow \nabla'_\mu V'^\nu = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\beta}\,\nabla_\alpha V^\beta. \tag{27}$$

Let's see what kind of constraint does this covariance condition impose on the transformation rule of the quantity, Γ under a GCT. To this end we compute the LHS of the above transformation law (27),

$$\begin{aligned}\nabla'_\mu V'^\nu &= \partial'_\mu V'^\nu + \Gamma'^\nu_{\mu\lambda}\,V'^\lambda \\ &= \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\beta}\,\partial_\alpha V^\beta + \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial^2 x'^\nu}{\partial x^\alpha \partial x^\beta}\,V^\beta + \Gamma'^\nu_{\mu\lambda}\,\frac{\partial x'^\lambda}{\partial x^\beta}V^\beta,\end{aligned}$$

while the RHS of (27) is,h

$$\frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\beta}\,\nabla_\alpha V^\beta = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\beta}\,\left(\partial_\alpha V^\beta + \Gamma^\beta_{\alpha\gamma}V^\gamma\right).$$

Equating the last two expressions, we get,

$$\frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial^2 x'^\nu}{\partial x^\alpha \partial x^\beta}\,V^\beta + \Gamma'^\nu_{\mu\lambda}\,\frac{\partial x'^\lambda}{\partial x^\beta}V^\beta = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\beta}\,\Gamma^\beta_{\alpha\gamma}V^\gamma,$$

or, simplifying things a bit

$$\left(\Gamma'^\nu_{\mu\lambda}\frac{\partial x'^\lambda}{\partial x^\gamma} - \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\beta}\,\Gamma^\beta_{\alpha\gamma} + \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial^2 x'^\nu}{\partial x^\alpha \partial x^\gamma}\right)\,V^\gamma = 0.$$

Since this equality holds for arbitrary V^γ, the quantity inside the parenthesis must be zero,

$$\Gamma'^\nu_{\mu\lambda}\frac{\partial x'^\lambda}{\partial x^\gamma} - \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\beta}\,\Gamma^\beta_{\alpha\gamma} + \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial^2 x'^\nu}{\partial x^\alpha \partial x^\gamma} = 0$$

or, multiplying both sides by $\frac{\partial x^\gamma}{\partial x'^\rho}$, we get the transformation rule for Γ,

$$\Gamma'^{\nu}_{\mu\rho} = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\gamma}{\partial x'^\rho}\frac{\partial x'^\nu}{\partial x^\beta}\,\Gamma^\beta_{\alpha\gamma} \underline{- \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\gamma}{\partial x'^\rho}\frac{\partial^2 x'^\nu}{\partial x^\alpha \partial x^\gamma}}. \tag{28}$$

Thus we see that Γ itself transforms non-covariantly, the second piece which is underlined is an inhomogeneous piece. The first piece (not underlined) is of course a pure tensor like transformation piece. Some remarks about this object, Γ and the related covariant derivative, ∇, are in order:

1. $\Gamma^\mu_{\nu\lambda}$ is called the ***affine connection***. The name comes from the idea that Γ maps vectors from distinct tangent vector spaces of nearby points (affine) on the manifold. The word affine is derived from the word affinity which means "closeness" or "kinship". Heuristically, the affine connection measures the rate of change of the basis vectors as one moves from one tangent space located at P to the tangent space located at a neighboring P'. Thus if $\hat{e}_\nu$ is a basis vector in the tangent space, then
$$\nabla_\mu \hat{e}_\nu \equiv \Gamma^\lambda_{\mu\nu}\hat{e}_\lambda.$$
(Check that this implies the following rule for covariant derivative of dual/contangent space basis vectors
$$\nabla_\mu \hat{\theta}^\nu = -\Gamma^\nu_{\mu\lambda}\hat{\theta}^\lambda.$$
)

2. This object is non-unique, we could only pin this connection down up to its transformation rule (28), not the exact expression. **Homework:** Show that adding a pure tensor, a $(1,2)$-type tensor, $T^\mu_{\nu\lambda}$ to an affine connection, $\Gamma^\mu_{\nu\lambda}$ gives us another new affine connection, $\bar{\Gamma}^\mu_{\nu\lambda}$ i.e. show that,
$$\bar{\Gamma}^\mu_{\nu\lambda} \equiv \Gamma^\mu_{\nu\lambda} + T^\mu_{\nu\lambda}$$
also transforms like (28). An equivalent statement is: The difference of two affine connections is a (1,2) rank pure tensor.

3. The covariant derivative, ∇ does obey the necessary conditions it needs to satisfy to be called a derivative i.e. linearity,
$$\nabla_\mu \left(V_1 + V_2\right)^\nu = \nabla_\mu V_1^\nu + \nabla_\mu V_2^\nu,$$
and Leibnitz rule,
$$\nabla_\lambda \left(V_1^\mu V_2^\nu\right) = V_1^\mu \nabla_\lambda V_2^\nu + V_2^\nu \nabla_\lambda V_2^\mu.$$

4. We note that the ordinary derivative of a scalar, $\partial_\mu \phi(x)$ does transform like a vector, so we must have that the covariant derivative of a scalar must reduce to its ordinary partial derivative,
$$\nabla_\mu \phi \equiv \partial_\mu \phi$$

5. Leibniz rule for covariant derivative plus the fact that ∇ reduces to a ordinary partial derivative, ∂ when acted on scalars immediately yields the action of covariant derivative on a cotangent space vector or one-form aka a covariant vector. Leibniz rule gives,

$$\begin{aligned}\nabla_\mu\left(V^\nu\omega_\nu\right) &= V^\nu\nabla_\mu\omega_\nu+\omega_\nu\nabla_\mu V^\nu \\ &= V^\nu\nabla_\mu\omega_\nu+\omega_\nu\left(\partial_\mu V^\nu+\Gamma^\nu_{\mu\lambda}\,V^\lambda\right),\end{aligned}$$

while scalar property gives,

$$\nabla_\mu\left(V^\nu\omega_\nu\right)=\partial_\mu\left(V^\nu\omega_\nu\right)=V^\nu\partial_\mu\omega_\nu+\omega_\nu\partial_\mu V^\nu.$$

Equating the two expressions,

$$V^\nu\nabla_\mu\omega_\nu+\omega_\nu\,\Gamma^\nu_{\mu\lambda}\,V^\lambda=V^\nu\partial_\mu\omega_\nu,$$

or,

$$\left(\nabla_\mu\omega_\nu-\partial_\mu\omega_\nu+\Gamma^\lambda_{\mu\nu}\omega_\lambda\right)V^\nu=0.$$

Since this equality holds for arbitrary V^ν, we must have,

$$\nabla_\mu\omega_\nu-\partial_\mu\omega_\nu+\Gamma^\lambda_{\mu\nu}\omega_\lambda=0,$$

or

$$\nabla_\mu\omega_\nu=\partial_\mu\omega_\nu-\Gamma^\lambda_{\mu\nu}\omega_\lambda \tag{29}$$

6. Finally by creating higher rank tensors by taking tensor products of contravariant and covariant vectors and using Leibniz rule we can arrive at the expression for covariant derivative of general rank tensors:

$$\begin{aligned}\nabla_\lambda\,T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q} &= \partial_\lambda\,T^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q}+\Gamma^{\mu_1}_{\lambda\alpha}\,T^{\alpha\mu_2\ldots\mu_p}{}_{\nu_1\ldots\nu_q}+\Gamma^{\mu_2}_{\lambda\alpha}\,T^{\mu_1\alpha\ldots\mu_p}{}_{\nu_1\ldots\nu_q}+\ldots \\ &\quad -\Gamma^\beta_{\lambda\nu_1}\,T^{\mu_1\ldots\mu_p}{}_{\beta\nu_2\ldots\nu_q}-\Gamma^\beta_{\lambda\nu_2}\,T^{\mu_1\ldots\mu_p}{}_{\nu_1\beta\ldots\nu_q}-\ldots\end{aligned} \tag{30}$$

3 Levi Civita connection

In a Riemannian manifold, (M,g), the specification of the metric singles out a very special affine connection which is a function of the metric and its derivative, $\Gamma^\lambda_{\mu\nu}=\Gamma^\lambda_{\mu\nu}(g,\partial g)$. This is called the Levi-Civita connection. It is defined by the following two properties,

1. Symmetry under exchange of the two downstairs indices:

$$\Gamma^\lambda_{\mu\nu}=\Gamma^\lambda_{\nu\mu}.$$

Otherwise known as the ***Torsion-free*** condition.

2. The metric tensor remains covariantly constant,

$$\nabla_\mu g_{\nu\rho} = 0.$$

Otherwise known as the ***metric compatibility*** condition.

Solving the above pair of equations, we get the Levi-Civita connection,

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\rho} \left(\partial_\mu g_{\rho\nu} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}\right). \tag{31}$$

Homework: Derive the above formula from the torsion-free and metric compatibility condition.

4 Point Particle in curved spacetime : Geodesic equation

The action for the point particle moving through curved spacetime is,

$$I_{pp}\left[x(\lambda)\right] = -m \int d\lambda \sqrt{-g_{\mu\nu}\left(x(\lambda)\right) \frac{dx^\mu(\lambda)}{d\lambda} \frac{dx^\nu(\lambda)}{d\lambda}}$$

where λ is some parameter labeling the curve. The Lagrangian is,

$$L\left(x^\mu, \frac{dx^\nu}{d\lambda}\right) = -m \sqrt{-g_{\mu\nu}\left(x(\lambda)\right) \frac{dx^\mu(\lambda)}{d\lambda} \frac{dx^\nu(\lambda)}{d\lambda}}, \tag{32}$$

and the Lagrange's equation of motion is,

$$\frac{\partial L}{\partial x^\mu} = \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{x}^\mu}\right), \dot{x} \equiv \frac{dx}{d\lambda}.$$

Plugging in the form of the Lagrangian (32), we get

$$-\frac{m}{2\sqrt{-g_{\alpha\beta}\left(x\right)\dot{x}^\alpha \dot{x}^\beta}} \left(-\partial_\mu g_{\rho\sigma} \dot{x}^\rho \dot{x}^\sigma\right) = \frac{d}{d\lambda}\left(\frac{m\, 2\, g_{\mu\rho}\dot{x}^\rho}{2\sqrt{-g_{\alpha\beta}\left(x\right)\dot{x}^\alpha \dot{x}^\beta}}\right) \tag{33}$$

Define a new parameter, the proper time, τ given by,

$$\frac{d\tau}{d\lambda} = \sqrt{-g_{\alpha\beta}\left(x\right)\dot{x}^\alpha \dot{x}^\beta},$$

so that,

$$\frac{d}{d\lambda} = \frac{d\tau}{d\lambda}\frac{d}{d\tau} = \left(\sqrt{-g_{\alpha\beta}\left(x\right)\dot{x}^\alpha \dot{x}^\beta}\right)\frac{d}{d\tau}.$$

In this parametrization the equation of motion, (33), looks like,

$$
\begin{aligned}
-\frac{m}{2}\left(-\partial_\mu g_{\rho\sigma}\frac{\dot{x}^\rho}{\sqrt{-g_{\alpha\beta}(x)\,\dot{x}^\alpha\dot{x}^\beta}}\frac{\dot{x}^\sigma}{\sqrt{-g_{\alpha\beta}(x)\,\dot{x}^\alpha\dot{x}^\beta}}\right) &= \frac{d}{d\tau}\left(\frac{m\,2\,g_{\mu\rho}\dot{x}^\rho}{2\sqrt{-g_{\alpha\beta}(x)\,\dot{x}^\alpha\dot{x}^\beta}}\right),\\
\Longrightarrow\ \frac{m}{2}\partial_\mu g_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} &= \frac{d}{d\tau}\left(m\,g_{\mu\rho}\frac{dx^\rho}{d\tau}\right)\\
&= m\left(\frac{d}{d\tau}g_{\mu\rho}\frac{dx^\rho}{d\tau}+g_{\mu\rho}\frac{d^2x^\rho}{d\tau^2}\right)\\
&= m\left(\partial_\nu g_{\mu\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau}+mg_{\mu\rho}\frac{d^2x^\rho}{d\tau^2}\right).
\end{aligned}
$$

Multiplying both sides by $\frac{g^{\lambda\mu}}{m}$, we get,

$$\frac{d^2x^\lambda}{d\tau^2}+g^{\lambda\mu}\partial_\nu g_{\mu\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau}-\frac{1}{2}g^{\lambda\mu}\partial_\mu g_{\rho\sigma}\partial_\nu g_{\mu\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau}=0,$$

which can be rearranged a bit to look like,

$$\frac{d^2x^\lambda}{d\tau^2}+\frac{1}{2}g^{\lambda\mu}\left(\partial_\nu g_{\mu\rho}+\partial_\rho g_{\nu\mu}-\partial_\mu g_{\nu\rho}\right)\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau}=0.$$

But the quantity in the parenthesis can be identified to be the Levi-Civita connection, hence we have the equation of motion of a point particle to be,

$$\frac{d^2x^\lambda}{d\tau^2}+\Gamma^\lambda_{\nu\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau}=0. \tag{34}$$

This is also called ***geodesic equation***. Geodesics are defined to be curves of extremal distances (stationary variation) in a Riemannian manifold. Since the point particle action is proportional to the path-length, extremizing the particle action gives rise to particle trajectories along curves of least path-length or extremal path-length.

4.1 Parallel transport

We have claimed earlier that the affine connection(s) allows us to compare vectors in different tangent spaces. One can then use the affine connection to ask the question, whether two vectors in the two different tangent spaces can be "aligned" or "parallel" to each other. Another way of phrasing the same question is, how can one use the affine connection to transport a given vector from one tangent space to another in a manner "parallel to itself". To answer this we consider a vector, V at the point P. Suppose we want to transport this vector V living in T_PM to another point P' along some path connecting P toP'. If $\xi^\mu=\frac{dx^\mu}{d\lambda}$ be the tangent vector along the path (curve) PP', then for parallel transport, the change in the vector after being transported on the curve by an parameter interval, $d\lambda$ is,

$$(\xi^\mu\nabla_\mu V^\nu)=0.$$

Expanding out everything,

$$
\begin{aligned}
\frac{dx^\mu}{d\lambda}\left(\partial_\mu V^\nu+\Gamma^\nu_{\mu\rho}V^\rho\right) &= 0\\
\Longrightarrow\ \frac{dV^\nu}{d\lambda}+\Gamma^\nu_{\mu\rho}V^\rho\frac{dx^\mu}{d\lambda} &= 0.
\end{aligned} \tag{35}
$$

In particular we can choose $V^\mu = \frac{dx^\mu}{d\lambda}$ i.e. the tangent vector to some yet undetermined curve itself. Then the equation of motion of this self-parallel motion i.e. parallel transporting this tangent vector along itself is given by,

$$\frac{d^2x^\nu}{d\lambda^2} + \Gamma^\nu_{\mu\rho}\frac{dx^\mu}{d\lambda}\frac{dx^\rho}{d\lambda} = 0. \tag{36}$$

Now we know in Euclidean geometry only type of curve is obtained parallel transporting along its tangent - the straight line. So the above equation gives the definition of "straight line" in curved spacetimes. Note that the connection appearing in the self-parallel equation can be ANY affine connection, not necessarily the Levi-Civita connection. However the connection appearing in the geodesic motion is the Levi-Civita connection.

5 Geodesic Deviation equation: Riemann Tensor

We consider a family of geodesics i.e. a one parameter family of geodesics, $x^\mu(\tau, s)$, where s is the label for the different geodesics and τ of course is the proper time labeling points on the same geodesic. Now at the beginning of the course I claimed that the curvature of spacetime is measured by examining whether the geodesics are converging, diverging or remaining parallel. In this section we will study this. Few things we need are:

- tangent vector along the geodesic,

$$T^\mu = \frac{\partial x^\mu(\tau, s)}{\partial \tau}.$$

- geodesic separation/deviation vector,

$$S^\mu = \frac{\partial x^\nu(\tau, s)}{\partial s},$$

  i.e. the vector connecting points on two different geodesics but those points have the same value of τ is,

$$\sim S^\mu\, \delta s$$

  .

Note that since the order of partial derivatives do not matter,

$$\frac{\partial T^\mu}{\partial s} = \frac{\partial S^\mu}{\partial \tau} = \frac{\partial^2 x^\mu(\tau, s)}{\partial \tau \partial s}.$$

So the first natural candidate for how quick the geodesics converging or diverging as we progress along the geodesic i.e. move along τ is perhaps acceleration of the geodesic deviation i.e.

$$\propto \frac{\partial^2 S^\mu}{d\tau^2}.$$

But as we will see this is not quite appropriate. The reason goes as follows. Recall that for parallel straights lines in flat space, we naively expect the tangent vector to remain unchanged as we move along the straight line, i.e.

$$\frac{\partial T^\mu}{\partial \tau} = 0,$$

BUT we have seen that in polar coordinates this is not so, in fact the proper equation is,

$$T^\nu \nabla_\nu T^\mu = 0.$$

So taking the cue from this, we can modify our claim, the geodesic separation rate should be measured perhaps by the quantity

$$(T.\nabla)^2 S^\mu,$$

instead of $\frac{\partial^2 S^\mu}{\partial \tau^2}$.

Before we go any further we note the following identity

$$\begin{aligned} (T.\nabla) S^\mu &= T^\nu \nabla_\nu S^\mu \\ &= T^\nu \left(\partial_\nu S^\mu + \Gamma^\mu_{\nu\rho} S^\rho\right) \\ &= \frac{\partial S^\mu}{\partial \tau} + \Gamma^\mu_{\nu\rho} T^\nu S^\rho \\ &= \frac{\partial T^\mu}{\partial s} + S^\rho \Gamma^\mu_{\rho\nu} T^\nu \\ &= S^\nu \nabla_\nu T^\mu \\ &= (S.\nabla) T^\mu \end{aligned} \tag{37}$$

where we have used chain rule,

$$T^\nu = \frac{\partial x^\nu(\tau, s)}{\partial \tau}, \quad T^\nu \partial_\nu = \frac{\partial x^\nu(\tau, s)}{\partial \tau} \frac{\partial}{\partial x^\nu(\tau, s)} = \frac{\partial}{\partial \tau}$$

to go from the second to third line. Then we used

$$\frac{\partial S^\mu}{\partial \tau} = \frac{\partial^2 x^\mu(\tau, s)}{\partial \tau \partial s} = \frac{\partial T^\mu}{\partial s}$$

to go from third to fourth line. Then finally we used the

$$\frac{\partial}{\partial s} = \frac{\partial x^\nu(\tau, s)}{\partial s} \frac{\partial}{\partial x^\nu(\tau, s)} = S^\nu \partial_\nu$$

and, the parallel transport definition,

$$S^\nu \partial_\nu T^\mu + S^\rho \Gamma^\mu_{\rho\nu} T^\nu = S^\nu \left(\partial_\nu T^\mu + S^\rho \Gamma^\mu_{\nu\rho} T^\rho\right) = S^\nu \nabla_\nu T^\mu,$$

to go from the fourth line to the fifth line.

Now let's go back to our object of interest, i.e. the geodesic separation acceleration $(T.\nabla)^2 S^\mu$. Now observe that, due to the identity (37)

$$(T.\nabla)^2 S = T.\nabla (T.\nabla S) = T.\nabla (S.\nabla T).$$

Next we fully expand this quantity, indices and all in full glory.

$$
\begin{aligned}
T.\nabla\left(S.\nabla T\right)^{\mu} &= T^{\nu}\nabla_{\nu}\left(S.\nabla T\right)^{\mu} \\
&= T^{\nu}\partial_{\nu}\left(S.\nabla T\right)^{\mu} + T^{\nu}\Gamma^{\mu}_{\nu\sigma}\left(S.\nabla T\right)^{\sigma} \\
&= \frac{\partial}{\partial\tau}\left(S^{\rho}\nabla_{\rho}T\right)^{\mu} + T^{\nu}\Gamma^{\mu}_{\nu\lambda}\left(S^{\rho}\nabla_{\rho}T\right)^{\lambda} \\
&= \frac{\partial}{\partial\tau}\left(S^{\rho}\partial_{\rho}T^{\mu} + S^{\rho}\Gamma^{\mu}_{\rho\sigma}T^{\sigma}\right) + T^{\nu}\Gamma^{\mu}_{\nu\lambda}\left(S^{\rho}\partial_{\rho}T^{\lambda} + S^{\rho}\Gamma^{\lambda}_{\rho\sigma}T^{\sigma}\right) \\
&= \frac{\partial}{\partial\tau}\left(\frac{\partial}{\partial s}T^{\mu} + S^{\rho}\Gamma^{\mu}_{\rho\sigma}T^{\sigma}\right) + T^{\nu}\Gamma^{\mu}_{\nu\lambda}\left(\frac{\partial}{\partial s}T^{\lambda} + S^{\rho}\Gamma^{\lambda}_{\rho\sigma}T^{\sigma}\right) \\
&= \frac{\partial^2 T^{\mu}}{\partial\tau\partial s} + \underbrace{\frac{\partial}{\partial\tau}\left(S^{\rho}\Gamma^{\mu}_{\rho\sigma}T^{\sigma}\right)} + T^{\nu}\Gamma^{\mu}_{\nu\lambda}\frac{\partial T^{\lambda}}{\partial s} + \Gamma^{\mu}_{\nu\lambda}\Gamma^{\lambda}_{\rho\sigma}T^{\nu}S^{\rho}T^{\sigma} \\
&= \frac{\partial^2 T^{\mu}}{\partial\tau\partial s} + \underbrace{\frac{\partial S^{\rho}}{\partial\tau}\Gamma^{\mu}_{\rho\sigma}T^{\sigma} + T^{\nu}S^{\rho}T^{\sigma}\Gamma^{\mu}_{\rho\sigma,\nu} + S^{\rho}\Gamma^{\mu}_{\rho\sigma}\frac{\partial T^{\sigma}}{\partial\tau}} \\
&\qquad\qquad + T^{\nu}\Gamma^{\mu}_{\nu\lambda}\frac{\partial T^{\lambda}}{\partial s} + \Gamma^{\mu}_{\nu\lambda}\Gamma^{\lambda}_{\rho\sigma}T^{\nu}S^{\rho}T^{\sigma}, \\
&= \frac{\partial^2 T^{\mu}}{\partial\tau\partial s} + \frac{\partial S^{\rho}}{\partial\tau}\Gamma^{\mu}_{\rho\sigma}T^{\sigma} + T^{\nu}\Gamma^{\mu}_{\nu\lambda}\frac{\partial T^{\lambda}}{\partial s} + S^{\rho}\Gamma^{\mu}_{\rho\sigma}\frac{\partial T^{\sigma}}{\partial\tau} \\
&\qquad\qquad + T^{\nu}S^{\rho}T^{\sigma}\Gamma^{\mu}_{\rho\sigma,\nu} + \Gamma^{\mu}_{\nu\lambda}\Gamma^{\lambda}_{\rho\sigma}T^{\nu}S^{\rho}T^{\sigma}.
\end{aligned} \tag{38}
$$

Note that in going from the second last to the last line, I have used,

$$
\frac{\partial S^{\rho}}{\partial\tau} = \frac{\partial T^{\rho}}{\partial s}.
$$

To male further progress, we need to simply the term,

$$
\frac{\partial^2 T^{\mu}}{\partial\tau\partial s}.
$$

To this end we note that, we take the following covariant directional derivative of both sides of the geodesic equation,

$$
S.\nabla\underbrace{\left(T.\nabla T = 0\right)} \implies S.\nabla\left(T.\nabla T\right) = 0,
$$

or in full blown index notation,

$$
\begin{aligned}
0 &= \left[S.\nabla\left(T.\nabla T\right)\right]^{\mu} \\
&= \left(S^{\rho}\nabla_{\rho}T.\nabla T\right)^{\mu} \\
&= S^{\rho}\partial_{\rho}\left(T.\nabla T\right)^{\mu} + S^{\rho}\Gamma^{\mu}_{\rho\beta}\underbrace{\left(T.\nabla T\right)^{\beta}} \\
&= \frac{\partial}{\partial s}\left(T^{\nu}\partial_{\nu}T^{\mu} + T^{\nu}\Gamma^{\mu}_{\nu\sigma}T^{\sigma}\right) + S^{\rho}\Gamma^{\mu}_{\rho\beta}\underbrace{0} \\
&= \frac{\partial}{\partial s}\left(\frac{\partial}{\partial\tau}T^{\mu} + T^{\nu}\Gamma^{\mu}_{\nu\sigma}T^{\sigma}\right) \\
&= \frac{\partial^2 T^{\mu}}{\partial s\partial\tau} + \frac{\partial T^{\rho}}{\partial s}\Gamma^{\mu}_{\nu\sigma}T^{\sigma} + T^{\nu}S^{\rho}T^{\sigma}\Gamma^{\mu}_{\nu\sigma,\rho} + T^{\nu}\Gamma^{\mu}_{\nu\sigma}\frac{\partial T^{\sigma}}{\partial s}.
\end{aligned}
$$

This implies,

$$\frac{\partial^2 T^\mu}{\partial s \partial \tau} = -\frac{\partial T^\rho}{\partial s}\Gamma^\mu_{\nu\sigma}T^\sigma - T^\nu S^\rho T^\sigma \Gamma^\mu_{\nu\sigma,\rho} - T^\nu \Gamma^\mu_{\nu\sigma}\frac{\partial T^\sigma}{\partial s}. \tag{39}$$

We plug this in the result for the geodesic deviation equation (38), and we get,

$$\begin{aligned} T.\nabla\left(S.\nabla T\right)^\mu &= -\frac{\partial T^\rho}{\partial s}\Gamma^\mu_{\nu\sigma}T^\sigma - T^\nu S^\rho T^\sigma \Gamma^\mu_{\nu\sigma,\rho} - T^\nu \Gamma^\mu_{\nu\sigma}\frac{\partial T^\sigma}{\partial s} \\ &\quad + \frac{\partial S^\rho}{\partial \tau}\Gamma^\mu_{\rho\sigma}T^\sigma + T^\nu \Gamma^\mu_{\nu\lambda}\frac{\partial T^\lambda}{\partial s} + S^\rho \Gamma^\mu_{\rho\sigma}\frac{\partial T^\sigma}{\partial \tau} \\ &\quad + T^\nu S^\rho T^\sigma \Gamma^\mu_{\rho\sigma,\nu} + \Gamma^\mu_{\nu\lambda}\Gamma^\lambda_{\rho\sigma}T^\nu S^\rho T^\sigma \\ &= T^\nu S^\rho T^\sigma\left(\Gamma^\mu_{\rho\sigma,\nu} - \Gamma^\mu_{\nu\sigma,\rho} + \Gamma^\mu_{\nu\lambda}\Gamma^\lambda_{\rho\sigma}\right) + S^\rho \Gamma^\mu_{\rho\lambda}\underbrace{\frac{\partial T^\lambda}{\partial \tau}}. \end{aligned} \tag{40}$$

Finally we replace the underbraced term using the geodesic equation,

$$\frac{\partial T^\lambda}{\partial \tau} + \Gamma^\lambda_{\nu\sigma}T^\nu T^\sigma = 0 \implies \frac{\partial T^\lambda}{\partial \tau} = -\Gamma^\lambda_{\nu\sigma}T^\nu T^\sigma.$$

and obtain the expression for the geodesic separation "acceleration rate"

$$\begin{aligned} \left(T.\nabla\right)^2 S^\mu = T.\nabla\left(S.\nabla T\right)^\mu &= T^\nu S^\rho T^\sigma\left(\Gamma^\mu_{\rho\sigma,\nu} - \Gamma^\mu_{\nu\sigma,\rho} + \Gamma^\mu_{\nu\lambda}\Gamma^\lambda_{\rho\sigma} - \Gamma^\mu_{\rho\lambda}\Gamma^\lambda_{\nu\sigma}\right) \\ &= T^\nu S^\rho T^\sigma R^\mu{}_{\sigma\nu\rho}, \end{aligned} \tag{41}$$

where we have introduced the ***Riemann Curvature Tensor*** ,

$$R^\mu{}_{\sigma\nu\rho} \equiv \Gamma^\mu_{\rho\sigma,\nu} - \Gamma^\mu_{\nu\sigma,\rho} + \Gamma^\mu_{\nu\lambda}\Gamma^\lambda_{\rho\sigma} - \Gamma^\mu_{\rho\lambda}\Gamma^\lambda_{\nu\sigma}. \tag{42}$$

This tensor is supposed to measure the curvature of the underlying manifold since it is proportional to the rate of change (acceleration) of separation of a family of geodesics (recall such a family of geodesics are analogous to parallel straight lines in flat space)

Note at this point we have not yet proven that the indexed object (42) which I declared to be the Riemann tensor is indeed a tensor, especially because it is made up of non-tensor object such as the Levi-Civita connection and the derivatives of the Levi-Civita connection. Presently we will see that the object (42) can be expressed in a form such that its tensor properties are completely apparent and this is how most texts on Riemannian Geometry introduce the Riemann tensor,

$$\left[\nabla_\nu, \nabla_\rho\right]\hat{e}_\sigma = R^\mu{}_{\sigma\nu\rho}\hat{e}_\mu. \tag{43}$$

where we have used the commutator,

$$\left[\nabla_\nu, \nabla_\rho\right]\hat{e}_\sigma \equiv \nabla_\nu\left(\nabla_\rho \hat{e}_\sigma\right) - \nabla_\rho\left(\nabla_\nu \hat{e}_\sigma\right).$$

Since the LHS of equation (43) is a pure tensor being made up of unit vector, $\hat{e}_\rho$ and covariant derivatives, the RHS must be a tensor as well. This immediately means $R^\mu{}_{\sigma\nu\rho}$ is a tensor.

Check: Here we check that indeed the second definition (43) indeed gives us back the form (42) we obtained from the geodesic deviation equation. First we evaluate,

$$\begin{aligned} \nabla_\nu \left(\nabla_\rho \hat{e}_\sigma\right) &= \nabla_\nu \left(\Gamma^\lambda_{\rho\sigma} \hat{e}_\lambda\right) \\ &= \left(\nabla_\nu \Gamma^\lambda_{\rho\sigma}\right) \hat{e}_\lambda + \Gamma^\lambda_{\rho\sigma} \left(\nabla_\nu \hat{e}_\lambda\right) \\ &= \partial_\nu \Gamma^\lambda_{\rho\sigma}\, \hat{e}_\lambda + \Gamma^\lambda_{\rho\sigma} \Gamma^\sigma_{\nu\lambda}\, \hat{e}_\sigma. \end{aligned}$$

where we have used Leibniz rule from going from line 1 to line 2, and then in going from line 2 to line 3, we have used the face that covariant derivative of a pure number such as $\Gamma^\lambda_{\rho\sigma}$ is same as partial derivative,

$$\nabla_\nu \Gamma^\lambda_{\rho\sigma} = \partial_\nu \Gamma^\lambda_{\rho\sigma}.$$

Thus the lhs of (43)

$$\begin{aligned} \left[\nabla_\nu, \nabla_\rho\right] \hat{e}_\sigma &= \nabla_\nu \left(\nabla_\rho \hat{e}_\sigma\right) - \nabla_\rho \left(\nabla_\nu \hat{e}_\sigma\right) \\ &= \left(\partial_\nu \Gamma^\lambda_{\rho\sigma}\, \hat{e}_\lambda + \Gamma^\lambda_{\rho\sigma} \Gamma^\sigma_{\nu\lambda}\, \hat{e}_\sigma\right) - \left(\partial_\rho \Gamma^\lambda_{\nu\sigma}\, \hat{e}_\lambda + \Gamma^\lambda_{\nu\sigma} \Gamma^\sigma_{\rho\lambda}\, \hat{e}_\sigma\right) \\ &= \left(\partial_\nu \Gamma^\mu_{\rho\sigma} - \partial_\rho \Gamma^\mu_{\nu\sigma} + \Gamma^\lambda_{\rho\sigma} \Gamma^\mu_{\nu\lambda} - \Gamma^\lambda_{\nu\sigma} \Gamma^\mu_{\rho\lambda}\right) \hat{e}_\mu \\ &= R^\mu{}_{\sigma\nu\rho} \hat{e}_\mu. \end{aligned}$$

So we have just shown that the two definitions of the Riemann tensor are equivalent at the end of the day. Now although by definition the Riemann tensor is a $(1,3)$-type tensor, it is more favorable in the research literature as well as some modern textbooks to lower the upstairs index and make a $(0,4)$-type tensor, namely,

$$R_{\mu\nu\rho\sigma} = g_{\mu\lambda} R^\lambda{}_{\nu\rho\sigma}.$$

By abuse of language we will also call this $(0,4)$ tensor as the Riemann Curvature tensor.

Homework:

Compute the Riemann tensor components for a two-dimensional manifold, namely the 2-sphere, given by the metric (of radius R),

$$ds^2 = R^2 \left(d\theta^2 + \sin^2\theta\, d\phi^2\right).$$

You will first need to compute, the Levi-Civita connection, $\Gamma^\mu_{\nu\lambda}$ and then plug them in (42).

5.1 (Anti)Symmetry Properties of the Riemann Tensor

1. Antisymmetry of the last pair of indices: This is evident from the definition (42) itself

   $$R^\mu{}_{\sigma\underline{\nu\rho}} = -R^\mu{}_{\sigma\underline{\rho\nu}}.$$

   or equivalently,

   $$R_{\mu\sigma\underline{\nu\rho}} = -R_{\mu\sigma\underline{\rho\nu}}.$$

   This holds for any type of connection, Γ.

2. Antisymmetry of the last three indices (aka the **first Bianchi Identity**):

$$R^{\mu}{}_{\sigma\nu\rho} + R^{\mu}{}_{\nu\rho\sigma} + R^{\mu}{}_{\rho\sigma\nu} = 0. \tag{44}$$

or equivalently,

$$R_{\mu\sigma\nu\rho} + R_{\mu\nu\rho\sigma} + R_{\mu\rho\sigma\nu} = 0.$$

This result holds for a "Torsion-free" connection i.e. any connection which is symmetric,

$$\Gamma^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{\nu\mu}.$$

(Check this by evaluating

$$\left[\left[\nabla_\mu, \nabla_\nu\right], \nabla_\rho\right]\Phi = -3R^{\sigma}{}_{\rho\mu\nu}\partial_\sigma\Phi - 2\Gamma^{\alpha}_{[\mu\nu]}\nabla_\alpha(\partial_\rho\Phi) + 2\partial_\rho\left(\Gamma^{\alpha}_{[\mu\nu]}\partial_\alpha\Phi\right).$$

and then applying Jacobi identity for covariant derivatives on a scalar, i.e $[[\nabla_\mu, \nabla_\nu], \nabla_\rho] + [[\nabla_\nu, \nabla_\rho], \nabla_\mu] + [[\nabla_\rho, \nabla_\mu], \nabla_\nu]\Phi = 0$).

3. Antisymmetry wrt to the first pair of indices:

$$R_{\underline{\mu\sigma}\nu\rho} = -R_{\underline{\sigma\mu}\nu\rho}.$$

This only holds for Metric compatible and torsion-free connection, i.e. Levi Civita connection. (One can prove this by taking trace of $R^{\mu}{}_{\sigma\nu\rho}$ wrt to the first and second indices and then showing that it vanishes, i..e. $\delta^{\sigma}_{\mu}\, R^{\mu}{}_{\sigma\nu\rho} = 0$.)

4. Symmetry wrt to exchange of first pair with second pair of indices,

$$R_{\mu\nu|\rho\sigma} = R_{\rho\sigma|\mu\nu}.$$

This is a result of the first Bianchi identity and the symmetry wrt the last two pair of indices.

5. Antisymmetry wrt derivatives, aka the second Bianchi Identity:

$$\nabla_\mu R^{\lambda}{}_{\sigma\nu\rho} + \nabla_\nu R^{\lambda}{}_{\sigma\rho\mu} + \nabla_\rho R^{\lambda}{}_{\sigma\mu\nu} = 0. \tag{45}$$

This can be proved by using the Jacobi Identity for commutators,

$$\left([\nabla_\mu, [\nabla_\nu, \nabla_\rho]] + [\nabla_\nu, [\nabla_\rho, \nabla_\mu]] + [\nabla_\rho, [\nabla_\mu, \nabla_\nu]]\right) V^\lambda = 0.$$

First we simplify each term in the above sum in the LHS

$$\begin{aligned}
[\nabla_\mu, [\nabla_\nu, \nabla_\rho]]\, V^\lambda &= \nabla_\mu\left([\nabla_\nu, \nabla_\rho]\, V^\lambda\right) - [\nabla_\nu, \nabla_\rho]\,(\nabla_\mu V)^\lambda \\
&= \nabla_\mu\left(R^{\lambda}{}_{\sigma\nu\rho}V^\sigma\right) - R^{\lambda}{}_{\sigma\nu\rho}\nabla_\mu V^\sigma + R^{\sigma}{}_{\mu\nu\rho}\nabla_\sigma V^\lambda \\
&= \left(\nabla_\mu R^{\lambda}{}_{\sigma\nu\rho}\right)V^\sigma + R^{\sigma}{}_{\mu\nu\rho}\nabla_\sigma V^\lambda,
\end{aligned}$$

Then we add all three terms and next we use Bianchi Identity I, Eq. (44) . This will give us (45).

Using the first four identities, one can show that the **number of independent components of the Riemann Curvature tensor in D-spacetime dimensions is,**

$$\frac{D^2\left(D^2-1\right)}{12}.$$

Check: Naively, the Riemann Curvature tensor, $R_{\mu\nu\,\rho\sigma}$ has D^4 components since each of the four indices can take D values. But then it fails to take into account the symmetry and antisymmetry property of the Riemann Curvature tensor. The antisymmetry of the first pair of indices,$\mu\nu$ tells us that the first pair can take only $N = {}^{D}C_2$ values. So for the second pair of indices, $\rho\sigma$. But then the symmetry of exchanging the first pair, $\mu\nu$ with the second pair, $\rho\sigma$ tells us the number of independent components of this four tensor object could only be

$$\frac{N(N+1)}{2} = \frac{D(D-1)\left(D^2-D+2\right)}{8}.$$

But then there is still the first Bianchi Identity, which gives constraints if we pick four different values for $\mu\nu\rho\sigma$ i.e. ${}^{D}C_4$. We need to remove these ${}^{D}C_4$ from the possible, $N(N+1)/2$ which gives us the number of independent components of the Riemann tensor to be:

$$\frac{N(N+1)}{2} - {}^{D}C_4 = \frac{D^2(D^2-1)}{12}.$$

6 Einstein Field Equations

In general relativity, gravitation is the curvature of spacetime caused by a source which is the stress-energy-momentum tensor. So we would like some equation which has the cause i.e. $T_{\mu\nu}$ on the RHS and the effect i.e. some curvature related quantity on the LHS, i.e. related to the Riemann tensor, $R^{\lambda}{}_{\mu\rho\nu}$. But clearly we cannot have some equation like,

$$R^{\lambda}{}_{\mu\rho\nu} \stackrel{?}{=} T_{\mu\nu}$$

because the indices on both sides are not the same. So first and foremost we have to contract two indices in the lhs. So we introduce the ***Ricci tensor*** as a contraction of Riemann tensor:

$$(Ric)_{\mu\nu} \equiv R^{\lambda}{}_{\mu\lambda\nu}.$$

Note that this tensor is symmetric in it's indices,

$$(Ric)_{\mu\nu} = (Ric)_{\nu\mu}\,.$$

Check:

$$\begin{aligned}(Ric)_{\mu\nu} &= g^{\rho\sigma}R_{\rho\mu\sigma\nu}\\ &= g^{\sigma\rho}R_{\sigma\nu\rho\mu}\\ &= (Ric)_{\nu\sigma}\,.\end{aligned}$$

Since the indices on both sides match, could we have an equation for general relativity such as,

$$(Ric)_{\mu\nu} \stackrel{?}{=} T_{\mu\nu}$$

modulo some proportionality constant. But even this equation cannot be correct, because the RHS being a conserved current, obeys

$$\nabla^{\mu} T_{\mu\nu} = 0,$$

while,

$$\nabla^{\mu} (Ric)_{\mu\nu} \neq 0.$$

Check:

We take the second Bianchi Identity, i.e.

$$\nabla_{\mu} R^{\lambda}{}_{\sigma\nu\rho} + \nabla_{\nu} R^{\lambda}{}_{\sigma\rho\mu} + \nabla_{\rho} R^{\lambda}{}_{\sigma\mu\nu} = 0.$$

and we multiply both sides by, δ^{ν}_{λ} to obtain

$$\nabla_{\mu} (Ric)_{\sigma\rho} + \nabla_{\lambda} R^{\lambda}{}_{\sigma\rho\mu} - \nabla_{\rho} (Ric)_{\sigma\mu} = 0,$$

to which we multiply both sides by $g^{\mu\sigma}$,

$$\begin{aligned} \nabla^{\mu} (Ric)_{\mu\rho} + \nabla^{\lambda} (R_{\lambda\sigma\rho\mu} g^{\mu\sigma}) - \nabla_{\rho} (Ric) &= 0 \\ \Longrightarrow \nabla^{\mu} (Ric)_{\mu\rho} + \nabla^{\lambda} (R_{\rho\mu\lambda\sigma} g^{\mu\sigma}) - \nabla_{\rho} (Ric) &= 0 \\ \Longrightarrow \nabla^{\mu} (Ric)_{\mu\rho} + \nabla^{\lambda} (R_{\mu\rho\sigma\lambda} g^{\mu\sigma}) - \nabla_{\rho} (Ric) &= 0 \\ \Longrightarrow \nabla^{\mu} (Ric)_{\mu\rho} + \nabla^{\lambda} (Ric)_{\rho\lambda} - \nabla_{\rho} (Ric) &= 0 \\ \nabla^{\mu} (Ric)_{\mu\rho} + \nabla^{\lambda} (Ric)_{\lambda\rho} - \nabla_{\rho} (Ric) &= 0. \end{aligned}$$

So at the end we have,

$$\nabla^{\mu} (Ric)_{\mu\rho} = \frac{1}{2} \nabla_{\rho} (Ric), \tag{46}$$

where (Ric) is the trace of the Ricci tensor, $(Ric) = g^{\mu\nu} (Ric)_{\mu\nu}$.

But notice using the fact that covariant derivative of the metric vanishes we can rewrite the RHS,

$$\nabla_{\rho} (Ric) = g_{\rho\mu} \nabla^{\mu} (Ric) = \nabla^{\mu} (g_{\rho\mu}\ (Ric)).$$

So then using this form in the RHS (46) we have,

$$\nabla^{\mu} (Ric)_{\mu\rho} = \nabla^{\mu} \left(\frac{1}{2} g_{\mu\rho}\ (Ric) \right),$$

or,

$$\nabla^{\mu} \left[(Ric)_{\mu\rho} - \frac{1}{2} g_{\mu\rho}\ (Ric) \right] = 0. \tag{47}$$

This leads us to define the ***Einstein Tensor***,

$$G_{\mu\nu} = (Ric)_{\mu\rho} - \frac{1}{2} g_{\mu\rho} \ (Ric), \tag{48}$$

which is divergence-free,

$$\nabla^{\mu} G_{\mu\nu} = 0.$$

So now we have a candidate equation,

$$G_{\mu\nu} = c \, T_{\mu\nu} \tag{49}$$

since both sides are:

- two-index object which are symmetric in the indices,
- both sides are divergence-free.

It so turns out that the demanding that the Eq. (49) reduce in the limit of weak gravitational fields and low velocities to the non-relativistic Newtonian form,

$$\boldsymbol{\nabla}.\mathbf{g} = -4\pi G_N \, \rho,$$

(where the $\boldsymbol{\nabla}$ is the three-dimensional gradient operator, $\mathbf{g}$ is the gravitational field, ρ is the mass-density and G_N is the Newton's universal gravitational constant), one needs to choose,

$$c = 8\pi G_N.$$

Thus we arrive at the final form of the ***Einstein's field equations for gravitation***,

$$(Ric)_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \ (Ric) = 8\pi G_N \, T_{\mu\nu}. \tag{50}$$

7 Gauss Theorem and Stokes Theorem for curved spacetimes: Conservation Laws

Now we face a little problem. General coordinate invariance demands that the continuity equation be changed to,

$$\partial_{\mu} j^{\mu} \to \nabla_{\mu} j^{\mu}.$$

However if we naively integrate this general covariant form of the continuity equation over all spacetime as we did in special relativity it does not give us a conservation law,

$$\begin{aligned}
0 = \int d^4x \, \nabla_{\mu} j^{\mu} \quad &= \quad \int d^4x \, \left(\partial_{\mu} j^{\mu} + \Gamma^{\mu}_{\mu\rho} j^{\rho}\right) \\
&= \quad \int d^4x \, \partial_{\mu} j^{\mu} + \int d^4x \, \Gamma^{\mu}_{\mu\rho} j^{\rho} \\
&= \quad \left(\int d^3\mathbf{x} \, j^0 \Big|_{t_f} - \int d^3\mathbf{x} \, j^0 \Big|_{t_f} \right) + \int dt \, \left(\int d^3\mathbf{x} \, \boldsymbol{\nabla} \cdot \mathbf{j} \right) + \int d^4x \, \Gamma^{\mu}_{\mu\rho} j^{\rho}.
\end{aligned}$$

If we define the charge to be the volume integral of the charge,

$$Q \equiv \int d^3\mathbf{x}\, j^0.$$

Then we have,

$$(Q(t_f) - Q(t_i)) + \int dt \left(\int d^3\mathbf{x}\, \boldsymbol{\nabla} \cdot \mathbf{j} \right) + \int d^4x\, \Gamma^{\mu}_{\mu\rho} j^{\rho} = 0$$

Using Gauss theorem in 3D, we see that the second term vanishes due to the surface integral at infinity since we assume there are no sources or sinks at spatial infinity ("localized sources" i.e. nothing is coming in or going out through spatial infinity),

$$\int d^3\mathbf{x}\, \boldsymbol{\nabla} \cdot \mathbf{j} = \oint\!\!\!\oint_{\infty} \mathbf{j} \cdot \hat{\mathbf{n}}\, dS = 0.$$

Thus, we get,

$$Q(t_f) = Q(t_i) - \int d^4x\, \Gamma^{\mu}_{\mu\rho} j^{\rho}.$$

i.e. initial and final charge is not the same or charge is not conserved. because of the term which contains the connection. In flat space, this term was not there and we had.,

$$Q(t_f) = Q(t_i),$$

i.e. conservation of charge.

So one wonders what went wrong, and the answer is that - the volume element we used, d^4x is not general covariant and instead we need to use,

$$\int d^4x\, \sqrt{-g}.$$

i.e. define charge to be,

$$Q = \int d^3\mathbf{x} \sqrt{-g}\, j^0$$

Let's see how this leads to resurrection of the conservation law. First we note the following identity,

$$\Gamma^{\mu}_{\mu\rho} = \frac{1}{2} g^{\mu\beta} g_{\mu\beta,\rho} = \frac{\partial_\rho \left(\sqrt{-g}\right)}{\sqrt{-g}} \tag{51}$$

Again we start by integrating the covariant continuity equation,

$$\begin{aligned}
0 &= \int d^4x\sqrt{-g}\,\nabla_\mu j^\mu \\
&= \int d^4x\sqrt{-g}\,\left(\partial_\mu j^\mu + \Gamma^\mu_{\mu\rho} j^\rho\right) \\
&= \int d^4x\sqrt{-g}\,\left(\partial_\rho j^\rho + \frac{\partial_\alpha\left(\sqrt{-g}\right)}{\sqrt{-g}}\, j^\rho\right) \\
&= \int d^4x\,\partial_\rho\left(\sqrt{-g}\,j^\rho\right) \\
&= \left(\int d^3\mathbf{x}\,\sqrt{-g}\,j^0\Big|_{t_f} - \int d^3\mathbf{x}\,\sqrt{-g}\,j^0\Big|_{t_f}\right) + \int dt \int d^3\mathbf{x}\,\boldsymbol{\nabla}\cdot\left(\sqrt{-g}\mathbf{j}\right) \\
&= Q(t_f) - Q(t_i) + \int dt \oiint_\infty \sqrt{-g}\,\mathbf{j}\cdot\hat{\mathbf{n}}\,dS \\
&= Q(t_f) - Q(t_i).
\end{aligned}$$

Thus now gave recovered our conservation law,

$$Q(t_f) = Q(t_i).$$

An important by-product of our discussion is the 4-dimensional version of the Gauss-theorem i.e.if we integrate over all space and all times, i.e. $t_f \to +\infty$, $t_i \to -\infty$.

$$\begin{aligned}
\int d^4x\sqrt{-g}\,\nabla_\mu j^\mu &= \int d^4x\,\partial_\rho\left(\sqrt{-g}\,j^\rho\right) \\
&= Q(+\infty) - Q(-\infty) + \int_{-\infty}^{\infty} dt \oiint_\infty \sqrt{-g}\,\mathbf{j}\cdot\hat{\mathbf{n}}\,dS.
\end{aligned}$$

So in the last line we have an integral over spatial and temporal infinity.

Caveat: This result cannot be generalized to tensors of arbitrary rank i.e.

$$\int d^4x\sqrt{-g}\,\nabla_\mu\left(T^{\mu\mu_2\ldots\mu_p}\right) \neq \int d^4x\,\partial_\mu\left(\sqrt{-g}\,T^{\mu\mu_2\ldots\mu_p}\right).$$

So there is <u>no Gauss theorem for arbitrary rank tensors</u>.

However one can generalize the Stokes theorem to arbitrary rank **antisymmetric tensors.** This is because one can show that for an antisymmetric tensor, say $A^{\mu\nu\ldots\sigma}$ which is antisymmetric in all its indices, $\mu, \nu, \ldots \sigma$

$$\int d^4x\sqrt{-g}\,\nabla_\mu A^{\mu\nu\ldots\sigma} = \int d^4x\,\partial_\mu\left(\sqrt{-g}\,A^{\mu\nu\ldots\sigma}\right)$$

and then again integrating over all space and time.

Homework:

Check that for a two-index antisymmetric tensor,

$$\sqrt{-g}\,\nabla_\mu A^{\mu\nu} = \partial_\mu\left(\sqrt{-g}\,A^{\mu\nu}\right).$$

Hint: Expand the covariant derivative and use the symmetry property of the downstairs indices of the Levi-Civita connection, $\Gamma^\mu_{\nu\lambda} = \Gamma^\mu_{\lambda\nu}$.