Chapter 3

Curvature

Problem Set #3: 3.1, 3.3, 3.5, 3.13, 3.15 (Due Monday Oct. 28th)
Midterm-exam: October 30th

3.1 Covariant derivative

In the previous chapter we have shown that the partial derivative of a non-scalar tensor is not a tensor (see (2.34)). It does not transform as a tensor but one might wonder if there is a way to define another derivative operator which would transform as a tensor and would reduce to the partial derivative in Minkowski space (note that exterior derivative does transform as a tensor, but does not reduce to partial derivative in the limit of flat space). The desired derivative operator (called **covariant derivative** and denote by ∇) can be constructed by by enforcing certain properties such as linearity, i.e.

$$\nabla(T+S) = \nabla T + \nabla S, \tag{3.1}$$

and product rule, i.e.

$$\nabla(T \otimes S) = \nabla T \otimes S + T \otimes \nabla S. \tag{3.2}$$

One can show (see Wlad's book) that if the operator obeys product rule than it can be written as a partial derivative plus a linear correction whose coefficients are called the **connection coefficients**,

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda. \tag{3.3}$$

Then the transformation properties of $\Gamma^\nu_{\mu\lambda}$ (which does not have to and will not transform as a tensor) can be determined by demanding that $\nabla_\mu V^\nu$ transforms as a $(1,1)$ tensor, i.e.

$$\nabla_{\mu'} V^{\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \nabla_\mu V^\nu. \tag{3.4}$$

By substituting (3.3) in (3.4) we obtain

$$\begin{aligned}
\partial_{\mu'} V^{\nu'} + \Gamma^{\nu'}_{\mu'\lambda'} V^{\lambda'} &= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \left(\partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda \right) \\
\frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial}{\partial x^\mu} \left(\frac{\partial x^{\nu'}}{\partial x^\nu} V^\nu \right) + \Gamma^{\nu'}_{\mu'\lambda'} \left(\frac{\partial x^{\lambda'}}{\partial x^\lambda} V^\lambda \right) &= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial}{\partial x^\mu} V^\nu + \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \Gamma^\nu_{\mu\lambda} V^\lambda \\
\frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial}{\partial x^\mu} \frac{\partial x^{\nu'}}{\partial x^\nu} V^\nu + \Gamma^{\nu'}_{\mu'\lambda'} \frac{\partial x^{\lambda'}}{\partial x^\lambda} V^\lambda &= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \Gamma^\nu_{\mu\lambda} V^\lambda \\
\Gamma^{\nu'}_{\mu'\lambda'} V^{\lambda'} &= \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \Gamma^\nu_{\mu\lambda} V^{\lambda'} - \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial}{\partial x^\mu} \frac{\partial x^{\nu'}}{\partial x^\lambda} V^{\lambda'}
\end{aligned} \tag{3.5}$$

Since this must be true for all $V^{\lambda'}$ we get a transformation law for the connection coefficients

$$\Gamma^{\nu'}_{\mu'\lambda'} = \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \Gamma^\nu_{\mu\lambda} - \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial^2 x^{\nu'}}{\partial x^\mu \partial x^\lambda}. \tag{3.6}$$

Similarly we can show that a different set of coefficients $\tilde{\Gamma}^\lambda_{\mu\nu}$ should be used to define a covariant derivative of a one form,

$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu + \tilde{\Gamma}^\lambda_{\mu\nu} \omega_\lambda \tag{3.7}$$

where $\tilde{\Gamma}^\lambda_{\mu\nu}$ although so far unrelated does transforms exactly as $\Gamma^\lambda_{\mu\nu}$ in (3.6). To establish a relation between $\Gamma^\lambda_{\mu\nu}$ and $\tilde{\Gamma}^\lambda_{\mu\nu}$ we demand that the covariant derivative reduces to partial derivative for scalars

$$\nabla_\mu \phi = \partial_\mu \phi. \tag{3.8}$$

Thus

$$\begin{aligned}
\partial_\mu \left(\omega_\lambda V^\lambda \right) &= \nabla_\mu \left(\omega_\lambda V^\lambda \right) \\
\left(\partial_\mu \omega_\lambda \right) V^\lambda + \left(\partial_\mu V^\lambda \right) \omega_\lambda &= \left(\nabla_\mu \omega_\lambda \right) V^\lambda + \left(\nabla_\mu V^\lambda \right) \omega_\lambda \\
\left(\partial_\mu \omega_\lambda \right) V^\lambda + \left(\partial_\mu V^\lambda \right) \omega_\lambda &= \left(\partial_\mu \omega_\lambda \right) V^\lambda + \tilde{\Gamma}^\sigma_{\mu\lambda} \omega_\sigma V^\lambda + \left(\partial_\mu V^\lambda \right) \omega_\lambda + \Gamma^\sigma_{\mu\lambda} \omega_\sigma V^\lambda \\
\Gamma^\sigma_{\mu\lambda} \omega_\sigma V^\lambda &= -\tilde{\Gamma}^\sigma_{\mu\lambda} \omega_\sigma V^\lambda.
\end{aligned} \tag{3.9}$$

Since ω_σ and V^λ are arbitrary

$$\Gamma^\sigma_{\mu\lambda} = -\tilde{\Gamma}^\sigma_{\mu\lambda}. \tag{3.10}$$

and therefore

$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu - \Gamma^\lambda_{\mu\nu} \omega_\lambda. \tag{3.11}$$

The covariant derivatives of an arbitrary rank tensor are given by

$$\begin{aligned}\nabla_\sigma T^{\mu_1 \dots \mu_k}{}_{\mu_k \dots \nu_l} &= \partial_\sigma T^{\mu_1 \dots \mu_k}{}_{\mu_k \dots \nu_l} + \Gamma^{\mu_1}_{\sigma\lambda} T^{\lambda\mu_2 \dots \mu_k}{}_{\nu_1 \dots \nu_l} + \dots \\ &\quad - \Gamma^{\lambda}_{\sigma\nu_1} T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} - \dots . \end{aligned} \tag{3.12}$$

In general relativity the $4^3 = 64$ independent connection coefficients of $\Gamma^\lambda_{\mu\nu}$ are uniquely specified by the metric $g_{\mu\nu}$. This is accomplished by demanding that the connection coefficient is **torsion-free**, i.e.

$$\Gamma^\lambda_{\mu\nu} = \Gamma^\lambda_{(\mu\nu)}, \tag{3.13}$$

and **metric-compatible**, i.e.

$$\nabla_\rho g_{\mu\nu} = 0. \tag{3.14}$$

The torsion-free implies, for example, that the antisymmetrized covariant derivative is also the exterior derivative, i.e.

$$\nabla_{[\mu}\omega_{\nu]} = \partial_{[\mu}\omega_{\nu]} - \Gamma^\lambda_{[\mu\nu]}\omega_\lambda = \partial_{[\mu}\omega_{\nu]}.$$

The metric-compatibility implies a number of nice properties. First of all the covariant derivative of inverse metric also vanishes, i.e.

$$\nabla_\rho g^{\mu\nu} = 0 \tag{3.15}$$

and thus the raising an lowering operators commute with covariant derivative

$$\nabla_\rho V^\mu = \nabla_\rho \left(g^{\mu\nu} V_\nu \right) = g^{\mu\nu} \nabla_\rho V_\nu. \tag{3.16}$$

Moreover the torsion-free and metric-compatible properties of the connection single out a unique connection known as the Christoffel (or Levi-Civita) connection of **Christoffel symbol**. The formula for the Christoffel symbol can be derived by from

$$0 = \nabla_\rho g_{\mu\nu} = \partial_\rho g_{\mu\nu} - \Gamma^\lambda_{\rho\mu} g_{\lambda\nu} - \Gamma^\lambda_{\rho\nu} g_{\mu\lambda} \tag{3.17}$$
$$0 = \nabla_\mu g_{\nu\rho} = \partial_\mu g_{\nu\rho} - \Gamma^\lambda_{\mu\nu} g_{\lambda\rho} - \Gamma^\lambda_{\mu\rho} g_{\nu\lambda} \tag{3.18}$$
$$0 = \nabla_\nu g_{\rho\mu} = \partial_\nu g_{\rho\mu} - \Gamma^\lambda_{\nu\rho} g_{\lambda\mu} - \Gamma^\lambda_{\nu\mu} g_{\rho\lambda} \tag{3.19}$$

by subtracting (3.18) and (3.19) from (3.17) and using (3.13),

$$\partial_\rho g_{\mu\nu} - \partial_\mu g_{\nu\rho} - \partial_\nu g_{\mu\rho} = \left(\Gamma^\lambda_{\rho\mu} g_{\lambda\nu} - \Gamma^\lambda_{\mu\rho} g_{\nu\lambda} \right) + \left(\Gamma^\lambda_{\rho\nu} g_{\mu\lambda} - \Gamma^\lambda_{\nu\rho} g_{\lambda\mu} \right) - \left(\Gamma^\lambda_{\mu\nu} g_{\lambda\rho} + \Gamma^\lambda_{\nu\mu} g_{\rho\lambda} \right) \tag{3.20}$$
$$\partial_\rho g_{\mu\nu} - \partial_\mu g_{\nu\rho} - \partial_\nu g_{\mu\rho} = -2\Gamma^\lambda_{\mu\nu} g_{\lambda\rho} \tag{3.21}$$

or

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\rho} \left(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}\right). \tag{3.22}$$

In Minkowski space described by Cartesian all of the Christoffel symbols vanish, but this does not have to be the case in curvilinear coordinates. As an example consider a two dimensional Euclidean space described by polar coordinates with metric

$$ds^2 = (\mathrm{d}r)^2 + r^2 (\mathrm{d}\theta)^2 . \tag{3.23}$$

The non-vanishing components of the inverse metric are

$$g^{rr} = 1 \tag{3.24}$$
$$g^{\theta\theta} = r^{-2} \tag{3.25}$$

and for example the connection coefficient

$$\begin{aligned}
\Gamma^{r}_{rr} &= \frac{1}{2} g^{r\rho} \left(\partial_r g_{r\rho} + \partial_r g_{\rho r} - \partial_\rho g_{rr}\right) = \\
&= \frac{1}{2} g^{rr} \left(2\partial_r g_{r\rho} - \partial_\rho g_{rr}\right) + \frac{1}{2} g^{r\theta} \left(2\partial_r g_{r\rho} - \partial_\rho g_{rr}\right) = \\
&= \frac{1}{2} 1 \left(2\partial_r g_{rr} - \partial_r g_{rr}\right) + \frac{1}{2} 0 \left(2\partial_r g_{r\sigma} - \partial_\sigma g_{rr}\right) = 0
\end{aligned} \tag{3.26}$$

but

$$\begin{aligned}
\Gamma^{r}_{\theta\theta} &= \frac{1}{2} g^{r\rho} \left(\partial_\theta g_{\theta\rho} + \partial_\theta g_{\rho\theta} - \partial_\rho g_{\theta\theta}\right) = \\
&= \frac{1}{2} g^{rr} \left(2\partial_\theta g_{\theta r} - \partial_r g_{rr}\right) + \frac{1}{2} g^{r\theta} \left(2\partial_\theta g_{\theta\theta} - \partial_\theta g_{\theta\theta}\right) = \\
&= \frac{1}{2} 1 \left(2\partial_\theta g_{\theta r} - \partial_r g_{rr}\right) + \frac{1}{2} 0 \left(2\partial_\theta g_{\theta\rho} - \partial_\rho g_{\theta\theta}\right) = \\
&= -\frac{1}{2} \partial_r g_{rr} = -r.
\end{aligned} \tag{3.27}$$

It is a straightforward exercise to find all other coefficients,

$$\Gamma^{r}_{\theta r} = \Gamma^{r}_{r\theta} = 0 \tag{3.28}$$
$$\Gamma^{\theta}_{rr} = 0 \tag{3.29}$$
$$\Gamma^{\theta}_{r\theta} = \Gamma^{\theta}_{\theta r} = \frac{1}{r} \tag{3.30}$$
$$\Gamma^{\theta}_{\theta\theta} = 0. \tag{3.31}$$

Just like one can make the connection coefficients to be non-zero in flat space it is possible to make the connection coefficients to vanish at some point curved space but not everywhere.

Since the covariant derivatives of a vector is

$$\nabla_\mu V^\mu = \partial_\mu V^\mu + \Gamma^\mu_{\mu\lambda} V^\lambda \tag{3.32}$$

and (one can also show)

$$\Gamma^\mu_{\mu\nu} = \frac{1}{\sqrt{|g|}} \partial_\nu \sqrt{|g|} \tag{3.33}$$

then we obtain a useful expression

$$\nabla_\mu V^\mu = \frac{1}{\sqrt{|g|}} \partial_\mu \left(\sqrt{|g|} V^\mu \right). \tag{3.34}$$

3.2 Parallel transport

Up until now we were not able to compare different tangent vectors at different points since they where not elements of the same vector space. With the help of the connection coefficients we can continuously move the tangent vectors (or higher rank tensors) from on point to another (or **parallel transport**), but the resulting vector will usually depend on the path along which it was moved. (Think about a parallel transport of a vector on the surface of a sphere to convince yourself that the parallel transported vector would depend on the path.) This is a generic property of curved spaces which is why it makes no sense to ask what is a relative velocity of two particles in two distinct points. In fact interpreting the cosmological expansion of space by galaxies receding away at a speed defined by the redshift is incorrect and can lead to paradoxes involving superluminal velocities. Of course in cosmology nothing is receding away, but the metric between galaxies changes which causes the light between the object to change the wavelength (i.e. to redshift).

In flat space the parallel transport of a tensor along a parametrized curve $x^\mu(\lambda)$ is given by the requirement

$$\left(\frac{D}{d\lambda} T\right)^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} = \frac{dx^\mu}{d\lambda} \partial_\mu T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} = 0 \tag{3.35}$$

which is generalized to curved spaces as

$$\left(\frac{D}{d\lambda} T\right)^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} = \frac{dx^\mu}{d\lambda} \nabla_\mu T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} = 0. \tag{3.36}$$

This is the **parallel transport equation** which, for example, take the following form for vectors

$$\frac{D}{d\lambda}V^{\mu} = \frac{d}{d\lambda}V^{\mu} + \Gamma^{\mu}_{\sigma\rho}\frac{dx^{\sigma}}{d\lambda}V^{\rho} = 0. \tag{3.37}$$

It follows, for example, that the inner product of two parallel transported vectors is preserved, i.e.

$$\frac{D}{d\lambda}\left(g_{\mu\nu}V^{\mu}W^{\nu}\right) = \frac{D}{d\lambda}\left(g_{\mu\nu}\right)V^{\mu}W^{\nu} + g_{\mu\nu}\frac{D}{d\lambda}\left(V^{\mu}\right)W^{\nu} + g_{\mu\nu}V^{\mu}\frac{D}{d\lambda}\left(VW^{\nu}\right) = 0. \tag{3.38}$$

Next we will obtain a formal solution of the parallel transport equation (3.37). Our task is to find the so-called **parallel propagator** matrix $P^{\mu}{}_{\rho}(\lambda_0, \lambda)$ along trajectory $\gamma(\lambda)$ such that

$$V^{\mu}(\lambda) = P^{\mu}{}_{\rho}(\lambda, \lambda_0)V^{\rho}(\lambda_0). \tag{3.39}$$

If we define a transition matrix

$$A^{\mu}{}_{\rho}(\lambda) \equiv -\Gamma^{\mu}_{\sigma\rho}\frac{dx^{\sigma}}{d\lambda} \tag{3.40}$$

then (3.37) can be written as a Schrodinger equation

$$\frac{d}{d\lambda}V^{\mu} = A^{\mu}{}_{\rho}V^{\rho}. \tag{3.41}$$

By substituting (3.39) into (3.41) we get

$$\frac{d}{d\lambda}\left[P^{\mu}{}_{\rho}(\lambda, \lambda_0)V^{\rho}(\lambda_0)\right] = A^{\mu}{}_{\sigma}\left[P^{\sigma}{}_{\rho}(\lambda, \lambda_0)V^{\rho}(\lambda_0)\right] \tag{3.42}$$

$$\frac{d}{d\lambda}P^{\mu}{}_{\rho}(\lambda, \lambda_0) = A^{\mu}{}_{\sigma}P^{\sigma}{}_{\rho}(\lambda, \lambda_0). \tag{3.43}$$

By integrating both side we get

$$P^{\mu}{}_{\rho}(\lambda, \lambda_0) = \delta^{\mu}_{\rho} + \int_{\lambda_0}^{\lambda} A^{\mu}{}_{\sigma}(\eta)P^{\sigma}{}_{\rho}(\eta, \lambda_0)d\eta \tag{3.44}$$

which can be solved by iteration

$$P^{\mu}{}_{\rho}(\lambda, \lambda_0) = \delta^{\mu}_{\rho} + \int_{\lambda_0}^{\lambda} A^{\mu}{}_{\rho}(\eta_1)d\eta_1 + \int_{\lambda_0}^{\lambda}\int_{\lambda_0}^{\eta_1} A^{\mu}{}_{\sigma}(\eta_1)A^{\sigma}{}_{\rho}(\eta_2)d\eta_1 d\eta_2 + ... \tag{3.45}$$

or in matrix notation

$$P(\lambda, \lambda_0) = 1 + \int_{\lambda_0}^{\lambda} A(\eta_1)d\eta_1 + \int_{\lambda_0}^{\lambda}\int_{\lambda_0}^{\eta_2} A(\eta_2)A(\eta_1)d\eta_1 d\eta_2 + ... \tag{3.46}$$

which can be simplified using a path-ordered product of matrices, $\mathcal{P}[A(\eta_n)A(\eta_{n-1})...A(\eta_1)]$, as

$$P(\lambda, \lambda_0) = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{\lambda_0}^{\lambda} \mathcal{P}[A(\eta_n)...A(\eta_1)] d\eta_1 d\eta_2 ... \tag{3.47}$$

which is the series expansion of an exponential,

$$P(\lambda, \lambda_0) = \mathcal{P} \exp \left(\int_{\lambda_0}^{\lambda} A(\eta) d\eta \right). \tag{3.48}$$

Turning back to the components notation gives us

$$P^{\mu}{}_{\nu}(\lambda, \lambda_0) = \mathcal{P} \exp \left(- \int_{\lambda_0}^{\lambda} \Gamma^{\mu}_{\sigma\nu} \frac{dx^{\sigma}}{d\eta} d\eta \right). \tag{3.49}$$

In quantum field theory the same formula is known as Dyson's formula which is due to the fact that (3.41) is mathematically equivalent to the Schrodinger equation. The parallel transport transformation around a closed loop is called the holonomy of the connection around the loop. For the metric-compatible connections the group of holonomy transformations at a point is a Lorentz group. Note that the knowledge of the holonomy group at each point is sufficient to determine the metric and thus might question what is more fundamental: holonomies of all loops or metric at all point. In the canonical quantum gravity the metric is treated as fundamental, when in the loop quantum gravity the holonomies are more fundamental.

3.3 Geodesics

Geodesic connecting two points can be defined as a path which parallel transports its own tangent vector, i.e.

$$\frac{D}{d\lambda} \frac{dx^{\mu}}{d\lambda} = \frac{d^2 x^{\mu}}{d\lambda^2} + \Gamma^{\mu}_{\rho\sigma} \frac{dx^{\rho}}{d\lambda} \frac{dx^{\sigma}}{d\lambda} = 0. \tag{3.50}$$

This is the geodesic equation which is obtained by substituting the tangent vector $\frac{dx^{\mu}}{d\lambda}$ into (3.37). For the metric-compatible connection coefficients the geodesic is also the path of the largest proper time (there is no smallest proper time since one can always construct a path of zero length made out of null segments) defined as

$$\tau[x] = \int d\lambda \sqrt{-g_{\mu\nu} \frac{dx^{\mu}}{d\lambda} \frac{dx^{\nu}}{d\lambda}} \tag{3.51}$$

that can be obtained using variational derivative, i.e.

$$\frac{\delta\tau[x]}{\delta x^\mu} = 0. \tag{3.52}$$

By Taylor expanding the metric

$$g_{\mu\nu}(x+\delta x) \approx g_{\mu\nu}(x) + \delta x^\sigma \partial_\sigma g_{\mu\nu}(x) \tag{3.53}$$

we can vary the proper time we get

$$\begin{aligned}
\frac{\delta\tau}{\delta x^\sigma} &= \frac{\delta\tau[x+\delta x]-\delta\tau[x]}{\delta x^\sigma} = \int d\lambda \frac{\sqrt{-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}} - \sqrt{-\left(g_{\mu\nu}+\delta x^\rho\partial_\rho g_{\mu\nu}\right)\frac{d(x^\mu+\delta x^\mu)}{d\lambda}\frac{d(x^\nu+\delta x^\nu)}{d\lambda}}}{\delta x^\sigma} \\
&= \int d\lambda \frac{\sqrt{-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}} - \sqrt{-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} - \left(\delta x^\rho\partial_\rho g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} - g_{\mu\nu}\frac{d\delta x^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} - g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{d\delta x^\nu}{d\lambda}\right)}}{\delta x^\sigma} \\
&= \int d\lambda \frac{\sqrt{-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}\left[1-\sqrt{1+\left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{-1}\left(-\delta x^\rho\partial_\rho g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} - 2g_{\mu\nu}\frac{d\delta x^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)}\right]}{\delta x^\sigma} \\
&= \int d\lambda \frac{\sqrt{-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}\left[\frac{1}{2}\left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{-1}\left(-\delta x^\rho\partial_\rho g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} - 2g_{\mu\nu}\frac{d\delta x^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)\right]}{\delta x^\sigma} \\
&= \int d\lambda \frac{\left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{-1/2}\left(-\frac{1}{2}\delta x^\rho\partial_\rho g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} - g_{\mu\nu}\frac{d\delta x^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)}{\delta x^\sigma}
\end{aligned} \tag{3.54}$$

By changing the integration variable from λ to proper time

$$d\lambda = d\tau\left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{-1/2} \tag{3.55}$$

we get

$$\begin{aligned}
\frac{\delta\tau}{\delta x^\sigma} &= \int d\tau \frac{\left(-\frac{1}{2}\delta x^\rho\partial_\rho g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} - g_{\mu\nu}\frac{d\delta x^\mu}{d\tau}\frac{dx^\nu}{d\tau}\right)}{\delta x^\sigma} \\
&= \int d\tau \frac{\left(-\frac{1}{2}\partial_\rho g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\delta x^\rho + \frac{d}{d\tau}\left(g_{\mu\nu}\frac{dx^\nu}{d\tau}\right)\delta x^\mu\right)}{\delta x^\sigma} \\
&= \int d\tau\left(-\frac{1}{2}\partial_\sigma g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} + \frac{d}{d\tau}\left(g_{\sigma\nu}\frac{dx^\nu}{d\tau}\right)\right)
\end{aligned} \tag{3.56}$$

or by demanding that $\delta\tau/\delta x^\sigma$ vanishes for all variations

$$\begin{aligned}
-\frac{1}{2}\partial_\sigma g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}+\frac{d}{d\tau}\left(g_{\sigma\nu}\frac{dx^\nu}{d\tau}\right) &= 0\\
-\frac{1}{2}\partial_\sigma g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}+\partial_\mu g_{\sigma\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}+g_{\sigma\nu}\frac{d^2x^\nu}{d\tau^2} &= 0\\
g_{\sigma\nu}\frac{d^2x^\nu}{d\tau^2}+\frac{1}{2}\left(-\partial_\sigma g_{\mu\nu}+\partial_\mu g_{\sigma\nu}+\partial_\nu g_{\mu\sigma}\right)\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} &= 0\\
\frac{d^2x^\rho}{d\tau^2}+\frac{1}{2}g^{\rho\sigma}\left(-\partial_\sigma g_{\mu\nu}+\partial_\mu g_{\sigma\nu}+\partial_\nu g_{\mu\sigma}\right)\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} &= 0. \qquad (3.57)
\end{aligned}$$

The last equation is equivalent to the geodesic equation (3.50) for the Christoffel connection, but may different for more general connections. The equation describes the motion of unaccelarated particle independent on its mass, but it is also straightforward to include forces. For example, if the particle has a charge q and mass m, then the geodesic equation would be given by

$$\frac{d^2x^\mu}{d\tau^2}+\Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau}=\frac{q}{m}F^\mu{}_\nu\frac{dx^\nu}{d\tau}. \qquad (3.58)$$

Note that the geodesics equation (3.50) is valid not only for proper time, but for any **affine parameter** defined as linear function of proper time, i.e.

$$\lambda = a\tau + b \qquad (3.59)$$

and for non-linear parametrizations (3.50) would be modified

$$\frac{d^2x^\mu}{d\alpha^2}+\Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\alpha}\frac{dx^\sigma}{d\alpha}=f(\alpha)\frac{dx^\mu}{d\alpha} \qquad (3.60)$$

where $f(\alpha)$ is some function which depends on the parametrization. Conversely if (3.60) is satisfied along some curve, then one can always find an affine parameter λ for which (3.50) is satisfied.

In addition to parallel transport the geodesics $x^\mu(\lambda)$ passing through some point p can be used to map point from the tangent space T_p to a neighborhood of p. Such mapping is called the **exponential map**

$$\exp_p : T_p \to M \qquad (3.61)$$

which is defined as

$$\exp_p(k^\mu) = x^\nu\,(\lambda = 1) \qquad (3.62)$$

where

$$k^\mu = \frac{dx^\mu(\lambda=0)}{d\lambda} \qquad (3.63)$$

is the tangent vector at point p where we have set $\lambda = 0$. Of course the map is well defined and invertible on a subset of T_p sufficiently close to $k^\mu = 0$, and so (3.61) should not be taken literally. The range of the map can fail to be all of the manifold given that there can be points not connected by a geodesic, and the range of the map can fail to be all of the tangent space if the manifold is geodesically incomplete (has singularities or boundaries).

3.4 Riemann tensor

The role of the Riemann (or curvature) tensor is to be able to represent the features of the connection coefficients which would manifest the presence of curvature. Such manifestations include parallel lines remain parallel or parallel transports around closed loops change vectors. The change of parallel transported vectors depends on the loop, and for a local description of the curvature it would be more useful to consider parallel transport around infinitesimal loops.

We can guess what kind of object the Riemann tensor R should be by considering a closed loop defined by two vectors A^μ and B^μ. Then a transformation around the loop should produce a vectorial change δV^μ to a vector V^μ such that

$$\delta V^\rho \propto R^\rho{}_{\sigma\mu\nu} V^\sigma A^\mu B^\nu. \tag{3.64}$$

Thus we anticipate the Riemann tensor to be of rank (1,3). It should also be antisymmetric, i.e.

$$R^\rho{}_{\sigma\mu\nu} = -R^\rho{}_{\sigma\nu\mu}, \tag{3.65}$$

so that the parallel transport in opposite direction comes with a negative sign.

To obtain the exact expression for the Riemann tensor in terms of connection coefficients we consider instead the action of a commutator of two covariant derivatives on the vector, i.e.

$$[\nabla_\mu, \nabla_\nu] V^\rho. \tag{3.66}$$

Indeed, the covariant derivative of a vector in the direction of a parallel transport is zero by definition, and thus the covariant derivative measure by how much the vector changes compared to what it would have been if parallel transported. Then the commutation of two covariant derivatives (when contracted with A^μ and B^ν) measures the changes to a vector V^μ if it was first parallel transformed one way (first along A^μ and then along B^μ) compared to the other way (first along B^μ and then along A^μ). The commutator of a vector is easy to express in terms of connection coefficients

$$\begin{aligned}
[\nabla_\mu, \nabla_\nu]V^\rho &= 2\nabla_{[\mu}\nabla_{\nu]}V^\rho = 2\left(\partial_{[\mu}\nabla_{\nu]}V^\rho - \Gamma^\sigma_{[\mu\nu]}\nabla_\sigma V^\rho + \Gamma^\rho_{[\mu|\sigma|}\nabla_{\nu]}V^\sigma\right) = \\
&= 2\left(\partial_{[\mu}\partial_{\nu]}V^\rho + \left(\partial_{[\mu}\Gamma^\rho_{\nu]\sigma}\right)V^\sigma + \Gamma^\rho_{[\nu|\sigma|}\partial_{\mu]}V^\sigma +\right) - \\
&- 2\left(\Gamma^\sigma_{[\mu\nu]}\partial_\sigma V^\rho + \Gamma^\sigma_{[\mu\nu]}\Gamma^\rho_{\lambda\sigma}V^\lambda\right) + 2\left(\Gamma^\rho_{[\mu|\sigma|}\partial_{\nu]}V^\sigma + \Gamma^\rho_{[\mu|\sigma}\Gamma^\sigma_{\lambda|\nu]}V^\lambda\right) = \\
&= 2\left(\partial_{[\mu}\Gamma^\rho_{\nu]\sigma} + \Gamma^\rho_{[\mu|\lambda}\Gamma^\lambda_{\sigma|\nu]}\right)V^\sigma - 2\Gamma^\sigma_{[\mu\nu]}\nabla_\sigma V^\rho. \qquad (3.67)
\end{aligned}$$

The second term is a torsion tensor, i.e.

$$T_{\mu\nu}{}^\sigma = -2\Gamma^\sigma_{[\mu\nu]} \qquad (3.68)$$

which vanishes for Christoffel connections, and the first term defines the Riemann tensor,

$$R^\rho{}_{\sigma\mu\nu} = 2\left(\partial_{[\mu}\Gamma^\rho_{\nu]\sigma} + \Gamma^\rho_{[\mu|\lambda}\Gamma^\lambda_{\sigma|\nu]}\right) = \partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\sigma\nu} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\sigma\mu}. \quad (3.69)$$

One can check that it is a legitimate (1,3) tensor, although it is not immediately clear that it is the same tensor as in (3.64) (see Wald for details). It is also straightforward to determine the action of the commutator $[\nabla_\rho, \nabla_\sigma]$ on a tensor of arbitrary rank,

$$[\nabla_\rho, \nabla_\sigma]X^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} = -2\Gamma^\lambda_{[\rho\sigma]}\nabla_\lambda X^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} + R^{\mu_1}{}_{\lambda\rho\sigma}X^{\lambda\mu_2\ldots\mu_k}{}_{\nu_1\ldots\nu_l}\ldots - R^\lambda{}_{\nu_1\rho\sigma}X^{\mu_1\ldots\mu_k}{}_{\lambda\nu_2\ldots\nu_l}\ldots. \qquad (3.70)$$

Sometimes it useful to express the relevant quantities in a coordinate independent way. Then one can think of the (1,2) torsion tensor as a map from two vectors to a third vector,

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] \qquad (3.71)$$

and the (1,3) Riemann tensor as a map from three vectors to a fourth vector,

$$R(X,Y)Z = \nabla_X\nabla_Y Z - \nabla_X\nabla_Y Z - \nabla_{[X,Y]}Z \qquad (3.72)$$

where the commutator of two vectors is

$$[X,Y]^\mu \equiv X^\lambda\partial_\lambda Y^\mu - Y^\lambda\partial_\lambda X^\mu \qquad (3.73)$$

and the directional covariant derivative is

$$\nabla_X \equiv X^\mu\nabla_\mu. \qquad (3.74)$$

Note that the third term in (3.72) vanishes when X and Y are coordinate basis since

$$[\partial_\mu, \partial_\nu] = 0. \tag{3.75}$$

In the case of metric compatible connection the Riemann tensor vanishes if the metric is constant everywhere on the manifold,

$$\begin{aligned} g_{\mu\nu} &= \text{const} \Rightarrow \\ \partial_\sigma g_{\mu\nu} &= 0 \Rightarrow \\ \Gamma^\rho_{\mu\nu} &= 0 \Rightarrow \\ \partial_\sigma \Gamma^\rho_{\mu\nu} &= 0 \Rightarrow \\ R^\rho{}_{\sigma\mu\nu} &= 0. \end{aligned} \tag{3.76}$$

One can also show that the opposite is true, i.e.

$$R^\rho{}_{\sigma\mu\nu} = 0 \quad \Rightarrow \quad g_{\mu\nu} = \text{const}. \tag{3.77}$$

Consider normal coordinates, i.e.

$$g_{\mu\nu} = \eta_{\mu\nu} \tag{3.78}$$

at some point p and with basis vectors $\hat{e}_{(\mu)}$ whose dot product is

$$g_{\sigma\rho}(p)\hat{e}^\sigma_{(\mu)}(p)\hat{e}^\rho_{(\nu)}(p) = \eta_{\mu\nu}. \tag{3.79}$$

This set of basis vectors can be parallel transported to some other point q. Since the Riemann tensor is zero the result of the transport does not depend on the trajectory and the dot product remains unchanged,

$$g_{\sigma\rho}(q)\hat{e}^\sigma_{(\mu)}(q)\hat{e}^\rho_{(\nu)}(q) = \eta_{\mu\nu}. \tag{3.80}$$

Moreover the commutator of the parallel transported basis vectors vanishes for torsion-free connections,

$$[\hat{e}_{(\mu)}, \hat{e}_{(\nu)}] = \nabla_{\hat{e}_{(\mu)}}\hat{e}_{(\nu)} - \nabla_{\hat{e}_{(\nu)}}\hat{e}_{(\mu)} - T(\hat{e}_{(\mu)}, \hat{e}_{(\nu)}) = 0. \tag{3.81}$$

Then according to the Forbenius's Theorem one can always find coordinates x^μ such that

$$\hat{e}_{(\mu)} = \frac{\partial}{\partial x^\mu} \tag{3.82}$$

in which the metric would have the form of $\eta_{\mu\nu}$.

Not all of the components of the Riemann tensor are independent and one can show that the components must obey,

$$\begin{aligned} R_{\rho\sigma\mu\nu} + R_{\sigma\rho\mu\nu} &= 0 && (3.83)\\ R_{\rho\sigma\mu\nu} - R_{\mu\nu\rho\sigma} &= 0 && (3.84)\\ R_{\rho\sigma\mu\nu} + R_{\rho\mu\nu\sigma} + R_{\rho\nu\sigma\mu} &= 0. && (3.85) \end{aligned}$$

which reduces the number of independent components to only 20 out of $4^4 = 256$. (The number is 0, 1 and 6 in 1, 2 and 3 dimensions respectively.) In addition algebraic identities there is a differential **Bianchi identity**,

$$\nabla_\lambda R_{\rho\sigma\mu\nu} + \nabla_\rho R_{\sigma\lambda\mu\nu} + \nabla_\sigma R_{\lambda\rho\mu\nu} = 0 \tag{3.86}$$

which is closely related to the Jacobi identity since it is nothing but

$$[[\nabla_\lambda, \nabla_\rho], \nabla_\sigma] + [[\nabla_\rho, \nabla_\sigma], \nabla_\lambda] + [[\nabla_\sigma, \nabla_\lambda], \nabla_\rho] = 0. \tag{3.87}$$

Some other useful tensor that can be formed from the Riemann tensor are:

- Ricci tensor:
  $$R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu} \tag{3.88}$$
  which is symmetric for Christoffel connections.
- Ricci scalar:
  $$R = R^\mu{}_\mu \tag{3.89}$$
  which can be shown (using Bianchi identity) to obey
  $$\nabla^\mu R_{\rho\mu} = \frac{1}{2}\nabla_\rho R. \tag{3.90}$$
- Einstein tensor:
  $$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} \tag{3.91}$$
  which is also symmetric for Christoffel connections and can be shown to obey
  $$\nabla^\mu G_{\mu\nu} = 0. \tag{3.92}$$
- Weyl tensor:
  $$C_{\rho\sigma\mu\nu} = R_{\rho\sigma\mu\nu} - \frac{1}{2}\left(g_{\rho[\mu} R_{\nu]\sigma} - g_{\sigma[\mu} R_{\nu]\rho}\right) + \frac{1}{3} R\, g_{\rho[\mu} g_{\nu]\sigma} \tag{3.93}$$
  which is essentially the Riemann tensor with all of the contractions removed (i.e. with zero traces) while maintaining its symmetries.

Consider an example of two dimensional sphere with metric

$$ds^2 = a^2\left(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2\right) \tag{3.94}$$

whose non-zero connection coefficients are

$$\begin{aligned} \Gamma^\theta_{\phi\phi} &= -\sin\theta\cos\theta \\ \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} &= \cot\theta. \end{aligned} \tag{3.95}$$

Then

$$\begin{aligned} R^\theta{}_{\phi\theta\phi} &= \partial_\theta\Gamma^\theta_{\phi\phi} - \partial_\phi\Gamma^\theta_{\theta\phi} + \Gamma^\theta_{\theta\lambda}\Gamma^\lambda_{\phi\phi} - \Gamma^\theta_{\phi\lambda}\Gamma^\lambda_{\theta\phi} = \\ &= (\sin^2\theta - \cos^2\theta) - (-\sin\theta\cos\theta\cot\theta) = \sin^2\theta \end{aligned} \tag{3.96}$$

and

$$R_{\theta\phi\theta\phi} = a^2\sin^2\theta. \tag{3.97}$$

It follows that

$$\begin{aligned} R_{\theta\theta} &= g^{\phi\phi}R_{\phi\theta\phi\theta} = 1 \\ R_{\theta\phi} &= R_{\phi\theta} = 0 \\ R_{\phi\phi} &= g^{\theta\theta}R_{\theta\phi\theta\phi} = \sin^2\theta \end{aligned} \tag{3.98}$$

and

$$R = g^{\theta\theta}R_{\theta\theta} + g^{\phi\phi}R_{\phi\phi} = 2a^{-2}. \tag{3.99}$$

Note that the Ricci scalar is positive as it should be for a positively curved space such as the two-sphere, but for negatively curved spaces such as a saddle it would be negative. The considered two-sphere is an example of a **maximally symmetric** space generically defined by

$$R_{\rho\sigma\mu\nu} = a^{-2}\left(g_{\rho\mu}g_{\sigma\nu} - g_{\rho\nu}g_{\sigma\mu}\right). \tag{3.100}$$

Consider a one parameter family of geodesics $\gamma_s(t)$ on a manifold M such that for each $s \in \mathbb{R}$, γ_s is parametrized by the affine parameter t. For a non-crossing family of geodesics (which may or may not be the case) we can choose the affine parameters in such a way that t and s form a system of coordinates on a two-dimensional surfaces embedded in M, i.e.

$$x^\mu(s,t) \in M. \tag{3.101}$$

There are two vector fields "tangent vector",

$$T^\mu = \frac{\partial x^\mu}{\partial t} \tag{3.102}$$

and "deviation vector"

$$S^\mu = \frac{\partial x^\mu}{\partial s} \tag{3.103}$$

whose commutator vanishes,

$$[S, T] = 0. \tag{3.104}$$

This means that they can be used as basis vectors for a coordinate system of our two-dimensional surface. Then the quantities defined as

$$V^\mu = (\nabla_T S)^\mu = T^\rho \nabla_\rho S^\mu \tag{3.105}$$

and

$$a^\mu = (\nabla_T V)^\mu = T^\rho \nabla_\rho V^\mu \tag{3.106}$$

we can call the "relative velocity" and "relative acceleration" of the geodesics.

For a connection with vanishing torsion equation (3.104) implies

$$S^\rho \nabla_\rho T^\mu = T^\rho \nabla_\rho S^\mu$$

and the "relative acceleration" is

$$\begin{aligned}
a^\mu &= T^\rho \nabla_\rho \left(T^\sigma \nabla_\sigma S^\mu\right) \\
&= T^\rho \nabla_\rho \left(S^\sigma \nabla_\sigma T^\mu\right) \\
&= \left(T^\rho \nabla_\rho S^\sigma\right) \nabla_\sigma T^\mu + T^\rho S^\sigma \left(\nabla_\rho \nabla_\sigma T^\mu\right) \\
&= \left(S^\rho \nabla_\rho T^\sigma\right) \nabla_\sigma T^\mu + T^\rho S^\sigma \left(\nabla_\sigma \nabla_\rho T^\mu + R^\mu{}_{\nu\rho\sigma} T^\nu\right) \\
&= \left(S^\rho \nabla_\rho T^\sigma\right) \nabla_\sigma T^\mu + S^\sigma \nabla_\sigma \left(T^\rho \nabla_\rho T^\mu\right) - \left(S^\sigma \nabla_\sigma T^\rho\right)\left(\nabla_\rho T^\mu\right) + R^\mu{}_{\nu\rho\sigma} T^\nu T^\rho S^\sigma \\
&= S^\sigma \nabla_\sigma \left(T^\rho \nabla_\rho T^\mu\right) + R^\mu{}_{\nu\rho\sigma} T^\nu T^\rho S^\sigma \\
&= R^\mu{}_{\nu\rho\sigma} T^\nu T^\rho S^\sigma
\end{aligned} \tag{3.107}$$

where in the last line we used that T^ρ is a tangent vector to geodesic and thus

$$T^\rho \nabla_\rho T^\mu = 0. \tag{3.108}$$

The resulting equation is known as the **geodesic deviation equation**,

$$a^\mu = \frac{D^2}{dt^2} S^\mu = R^\mu{}_{\nu\rho\sigma} T^\nu T^\rho S^\sigma \tag{3.109}$$

which can be interpreted as the **gravitational tidal force** due to curvature $R^\mu{}_{\nu\rho\sigma}$.

3.5 Pullback and pushforward

Consider a map

$$\phi : M \to N \tag{3.110}$$

where the manifolds M and N with respective coordinates x^μ and y^α might have different dimensions. Then one can think of ϕ as a function which maps function on N,

$$f : N \to \mathbb{R} \tag{3.111}$$

to functions on M,

$$\phi_* f : M \to \mathbb{R}, \tag{3.112}$$

where $\phi_* f$ defined as a composite map

$$\phi_* f \equiv f \circ \phi. \tag{3.113}$$

Such map is called the **pullback** since it is defined by pulling back the action of function f from manifold N to manifold M (i.e. in the direction opposite to the way ϕ was defined).

Clearly, it is not be possible to define a pushforward of a function, but it is possible to define a pushforward of a vector V which, as we know, is nothing but a map from functions on manifolds to real numbers. Then the **pushforward** of the vector field defined on manifold M, i.e.

$$V : \mathcal{F}(M) \to \mathbb{R}. \tag{3.114}$$

is a vector field on manifold N, i.e.

$$\phi^* V : \mathcal{F}(N) \to \mathbb{R} \tag{3.115}$$

defined by the action of vector V on the pullback of functions, i.e.

$$(\phi^* V)(f) \equiv V(\phi_* f). \tag{3.116}$$

In coordinate basis, the pushforward of a vector can be written as

$$\begin{aligned} (\phi^* V)^\alpha \partial_\alpha f &= V^\mu \partial_\mu (\phi_* f) \\ &= V^\mu \partial_\mu (f \circ \phi) \\ &= V^\mu \frac{\partial y^\alpha}{\partial x^\mu} \partial_\alpha f \end{aligned} \tag{3.117}$$

or

$$(\phi^* V)^\alpha = V^\mu \frac{\partial y^\alpha}{\partial x^\mu}. \tag{3.118}$$

Then we can defined the **pushforward matrix operator** as

$$(\phi^*)^\alpha{}_\mu = \frac{\partial y^\alpha}{\partial x^\mu} \tag{3.119}$$

so that

$$(\phi^* V)^\alpha = (\phi^*)^\alpha{}_\mu V^\mu. \tag{3.120}$$

Note that the pushforward matrix reduces to the general coordinate transformations matrix for $M = N$.

Similarly to how it is not possible to define a pushforward of a function it is not possible to define a pullback of a vector, but it is possible to define a pullback of a one-form which is nothing but a map from vectors to real numbers. Then the the pullback of a one-form ω on manifold N, is a one-form $\phi_*\omega$ on manifold M, defined by how it acts on pushforward vectors, i.e.

$$(\phi_*\omega)(V) = \omega(\phi^* V). \tag{3.121}$$

Then we can also define the **pullback matrix operator** as

$$(\phi_*)_\mu{}^\alpha = \frac{\partial y^\alpha}{\partial x^\mu} \tag{3.122}$$

so that

$$(\phi_*\omega)_\mu = (\phi_*)_\mu{}^\alpha \omega_\alpha. \tag{3.123}$$

The pushforward and pullback can respectively be generalized to arbitrary $(k,0)$ and $(0,l)$ tensors,

$$(\phi^* S)\left(\omega^{(1)}, ..., \omega^{(k)}\right) = S\left(\phi_*\omega^{(1)}, ..., \phi_*\omega^{(k)}\right), \tag{3.124}$$

$$(\phi_* T)\left(V^{(1)}, ..., V^{(l)}\right) = T\left(\phi^* V^{(1)}, ..., \phi^* V^{(l)}\right) \tag{3.125}$$

or in indices notation

$$(\phi^* S)^{\alpha_1 ... \alpha_k} = \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} ... \frac{\partial y^{\alpha_k}}{\partial x^{\mu_k}} S^{\mu_1 ... \mu_k} \tag{3.126}$$

$$(\phi_* T)_{\mu_1 ... \mu_l} = \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} ... \frac{\partial y^{\alpha_l}}{\partial x^{\mu_l}} T_{\alpha_1 ... \alpha_l}. \tag{3.127}$$

The pushforward and pullback of other tensors is defined only if the map ϕ is invertible and then

$$(\phi^* S)\left(\omega^{(1)}, ..., \omega^{(k)}, V^{(1)}, ..., V^{(l)}\right) = S\left(\phi_*\omega^{(1)}, ..., \phi_*\omega^{(k)}, \left(\phi^{-1}\right)^* V^{(1)}, ..., \left(\phi^{-1}\right)^* V^{(l)}\right) \tag{3.128}$$

or

$$(\phi_* T)\left(\omega^{(1)}, ..., \omega^{(k)}, V^{(1)}, ..., V^{(l)}\right) = T\left(\left(\phi^{-1}\right)_* \omega^{(1)}, ..., \left(\phi^{-1}\right)_* \omega^{(k)}, \phi^* V^{(1)}, ..., \phi^* V^{(l)}\right). \tag{3.129}$$

Clearly the pullback of ϕ is nothing but a pushforward of ϕ^{-1}.

Consider an example of a two-sphere S^2 with coordinates (θ, φ) embedded into a three dimensional Euclidean space $\mathbb{R}^3$ with coordinates (x, y, z) such that there is a natural map

$$\phi : S^2 \to \mathbb{R}^3 \tag{3.130}$$

defined by

$$\phi(\theta, \varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta). \tag{3.131}$$

Then the metric on $\mathbb{R}^3$ as a $(0, 2)$ tensor can be pullback to S^2. Then the **pullback matrix operator**

$$\phi_* = \begin{pmatrix} \cos\theta\cos\varphi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\theta\sin\varphi & \sin\theta\cos\phi & 0 \end{pmatrix} \tag{3.132}$$

can be used to determine the pullback of the metric on $\mathbb{R}^3$, i.e.

$$g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.133}$$

to the metric on S^2, i.e.

$$\begin{aligned} \phi_* g &= \begin{pmatrix} \cos\theta\cos\varphi & \cos\theta\sin\varphi & -\sin\theta \\ -\sin\theta\sin\varphi & \sin\theta\cos\varphi & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta\cos\varphi & -\sin\theta\sin\varphi \\ \cos\theta\sin\varphi & \sin\theta\cos\varphi \\ -\sin\theta & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta\cos\varphi & \cos\theta\sin\varphi & -\sin\theta \\ -\sin\theta\sin\varphi & \sin\theta\cos\varphi & 0 \end{pmatrix} \begin{pmatrix} \cos\theta\cos\varphi & -\sin\theta\sin\varphi \\ \cos\theta\sin\varphi & \sin\theta\cos\varphi \\ -\sin\theta & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix} \end{aligned} \tag{3.134}$$

As a result we will obtain a $(0, 2)$ tensor on S^2 which is also called the **induced metric**, $(\phi_* g)_{\mu\nu}$.

3.6 Killing vectors

Given a one parameter family of diffeomorphism

$$\phi_t : M \to M \tag{3.135}$$

such that

$$\phi_{s+t} = \phi_s \circ \phi_t \tag{3.136}$$

we can consider trajectory of a given point p under ϕ_t. These trajectories $x^\mu(t)$ can used to define vector fields by evaluating tangent vector at $t = 0$:

$$V^\mu(x) = \left[\frac{dx^\mu}{dt}\right]_{t=0}. \tag{3.137}$$

Similarly, we can define a one parameter family of diffeomorphism (or **integral curves**) from a given vector field (or **generator** of the diffeomorphism) by solving the differential equation

$$\frac{dx^\mu}{dt} = V^\mu. \tag{3.138}$$

Then for any tensor field we can ask how it changes as we travel along the **integral curves**, i.e.

$$\Delta_t T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}(p) = \phi_{t*}\left(T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}(p)\right) - T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}(p) \tag{3.139}$$

as well as define the so-called **Lie derivative**, i.e.

$$\mathcal{L}_V\left(T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}\right) = \lim_{t\to 0}\left(\frac{\Delta_t T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}}{t}\right) \tag{3.140}$$

which is a map from (k, l) to (k, l). The Lie derivative is linear

$$\mathcal{L}_V(aT + bS) = a\mathcal{L}_V T + b\mathcal{L}_V S \tag{3.141}$$

and obeys product rule

$$\mathcal{L}_V(T \otimes S) = (\mathcal{L}_V T) \otimes S + T \otimes (\mathcal{L}_V S). \tag{3.142}$$

Although the Lie derivative does not depend on the coordinate system it does not require specification of connections. It is easy to show that the Lie derivative reduces to an ordinary derivatives for scalars

$$\mathcal{L}_V f = V(f) = V^\mu \partial_\mu f \tag{3.143}$$

and to a commutator (also called the **Lie bracke**t) for vectors

$$\mathcal{L}_V U^\mu = [V, U]^\mu. \tag{3.144}$$

Then the Lie derivative of a one-form can be derived by considering the Lie derivative of a scalar

$$\begin{aligned}
\mathcal{L}_V\left(\omega_\mu U^\mu\right) &= V^\nu\partial_\nu\left(\omega_\mu U^\mu\right)\\
\left(\mathcal{L}_V\omega_\mu\right)U^\mu+\omega_\mu\left(\mathcal{L}_V U\right)^\mu &= V^\nu U^\mu\partial_\nu\omega_\mu+V^\nu\omega_\mu\partial_\nu U^\mu\\
\left(\mathcal{L}_V\omega_\mu\right)U^\mu+\omega_\mu\left(V^\nu\partial_\nu U^\mu-U^\nu\partial_\nu V^\mu\right) &= V^\nu U^\mu\partial_\nu\omega_\mu+V^\nu\omega_\mu\partial_\nu U^\mu\\
\left(\mathcal{L}_V\omega_\mu\right)U^\mu &= V^\nu U^\mu\partial_\nu\omega_\mu+\omega_\mu U^\nu\partial_\nu V^\mu\\
\mathcal{L}_V\omega_\mu &= V^\nu\partial_\nu\omega_\mu+\omega_\nu\partial_\mu V^\nu. \qquad (3.145)
\end{aligned}$$

For an arbitrary tensor the Lie derivative is given by

$$\mathcal{L}_V T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}=V^\sigma\partial_\sigma T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}-\left(\partial_\lambda V^{\mu_1}\right)T^{\lambda\mu_2\ldots\mu_k}{}_{\nu_1\ldots\nu_l}\ldots+\left(\partial_{\nu_1}V^\lambda\right)T^{\mu_1\ldots\mu_k}{}_{\lambda\nu_2\ldots\nu_l}\ldots. \qquad (3.146)$$

Although the above expression is coordinate independent one can rewrite it in a more covariant form

$$\mathcal{L}_V T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}=V^\sigma\nabla_\sigma T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}-\left(\nabla_\lambda V^{\mu_1}\right)T^{\lambda\mu_2\ldots\mu_k}{}_{\nu_1\ldots\nu_l}\ldots+\left(\nabla_{\nu_1}V^\lambda\right)T^{\mu_1\ldots\mu_k}{}_{\lambda\nu_2\ldots\nu_l}\ldots. \qquad (3.147)$$

A diffeomorphism ϕ is called a **symmetry** of some tensor T if it leaves the tensor invariant, i.e.

$$\phi_* T=\phi^* T=T. \qquad (3.148)$$

If the symmetry is generated by a vector field $V^\mu(x)$ then the symmetry implies that the Lie derivative vanishes

$$\mathcal{L}_V T=0. \qquad (3.149)$$

In that case one find a coordinate system in which the components of T are independent of one of the coordinates (pointing in the direction of the integral curves of V).

One of the most important symmetries are of the metric

$$\phi_* g_{\mu\nu}=g_{\mu\nu} \qquad (3.150)$$

in which case the diffeomorphism ϕ is called an **isometry**. The vector field $V^\mu(x)$ is called a **Killing vector field** if it generates the isometries, i.e.

$$\mathcal{L}_V g_{\mu\nu}=0 \qquad (3.151)$$

or

$$\begin{aligned}
0 &= V^\sigma\nabla_\sigma g_{\mu\nu}+\left(\nabla_\mu V^\sigma\right)g_{\sigma\nu}+\left(\nabla_\nu V^\sigma\right)g_{\mu\sigma}\\
0 &= \nabla_\mu V_\nu+\nabla_\nu V_\mu\\
0 &= \nabla_{(\mu}V_{\nu)} \qquad (3.152)
\end{aligned}$$

which is the **Killing equations**. The maximal number of Killing vector field in n-dimentional manifold is the sum of n translations and $n(n-1)/2$ rotations (from any axis n to any other axis $n-1$ but without double-counting). The sum of rotations and translations is $n(n+1)/2$ or ten for the four-dimensional space-time. The space-time of general relativity is called **maximally symmetric** if its metric tensor has ten independent Killing vectors which is the same as the number of independent parameters in a symmetric $(0,2)$ tensor. For example,

$$ds^2 = \mathrm{d}x^2 + \mathrm{d}y^2 \tag{3.153}$$

has three Killing vectors (i.e. satisfy the Killing equations (3.152)) corresponding to two translation

$$X = (1,0) \tag{3.154}$$
$$Y = (0,1) \tag{3.155}$$

and one rotation

$$R^\mu = (-y, x) \tag{3.156}$$

Consider a free particle moving along a geodesic $x^\mu(\lambda)$ whose tangent vector is

$$U^\mu = \frac{dx^\mu}{d\lambda}. \tag{3.157}$$

Then for a Killing vector V^μ we have

$$\begin{aligned} U^\nu \nabla_\nu (V_\mu U^\mu) &= U^\nu U^\mu \nabla_\nu V_\mu + V_\mu (U^\nu \nabla_\nu U^\mu) \\ &= \frac{1}{2}(U^\nu U^\mu \nabla_\nu V_\mu + U^\mu U^\nu \nabla_\mu V_\nu) \\ &= 0. \end{aligned} \tag{3.158}$$

which implies that the quantity $V_\mu U^\mu$ is conserved along the trajectory of motion.

3.7 Tetrads

So far we were using the coordinate basis for the tangent space

$$\hat{e}_{(\mu)} = \partial_\mu \tag{3.159}$$

defined as partial derivatives with respect to coordinate functions and for dual space

$$\hat{\theta}^{(\mu)} = \mathrm{d}x^\mu \tag{3.160}$$

defined as gradients of the coordinate functions. Of course, we are free to use any basis we like and there is one particularly useful set of bias vectors, known as **tetrads** or **vierbein**, defined at each point by

$$g\left(\hat{e}_{(a)}, \hat{e}_{(b)}\right) = \eta_{ab}. \tag{3.161}$$

The coordinate basis vectors can be expressed in tetrad basis as

$$\hat{e}_{(\mu)} = e^a_\mu \hat{e}_{(a)} \tag{3.162}$$

where the e^a_μ matrix an invertible matrix whose inverse e^μ_a expresses tetrad basis vectors in coordinate basis

$$\hat{e}_{(a)} = e^\mu_a \hat{e}_{(\mu)}. \tag{3.163}$$

As usual (with the abuse of notations) the component e^a_μ and e^μ_a are often called the tetrads and inverse tetrads respectively. Clearly, they must satisfy

$$e^\mu_a e^a_\nu = \delta^\mu_\nu \tag{3.164}$$

or

$$e^a_\mu e^\mu_b = \delta^a_b, \tag{3.165}$$

and (3.161) can be rewritten in the indices notation as

$$g_{\mu\nu} e^\mu_a e^\nu_b = \eta_{ab} \tag{3.166}$$

or

$$g_{\mu\nu} = e^a_\mu e^b_\nu \eta_{ab} \tag{3.167}$$

which is why tetrads are sometimes called the "square root" of the metric. After defining the orthonormal basis for vectors we can define an orthonormal basis for one-forms by setting

$$\hat{\theta}^{(a)}\left(\hat{e}_{(b)}\right) = \delta^a_b. \tag{3.168}$$

Then one can show that using the very same inverse tetrads e^μ_a and tetrads e^a_μ we can respectively express

$$\hat{\theta}^{(\mu)} = e^\mu_a \hat{\theta}^{(a)} \tag{3.169}$$

and

$$\hat{\theta}^{(a)} = e^a_\mu \hat{\theta}^{(\mu)}. \tag{3.170}$$

Of course, no only the basis vectors but any vectors expressed in terms of the coordinate basis can be re-expressed in terms of tetrad basis. In other words

$$V = V^a \hat{e}_{(a)} = V^\mu \hat{e}_{(\mu)} \tag{3.171}$$

implies

$$V^a = e^a_\mu V^\mu. \tag{3.172}$$

Similarly for tensors

$$V^a_{\ b} = e^a_\mu V^\mu_{\ b} = e^\nu_b V^a_{\ \nu} = e^a_\mu e^\nu_b V^\mu_{\ \nu} \tag{3.173}$$

and one example of such transformation we had already seen before for metric tensor (3.166).

For a non-coordinate tetrad basis we can illustrate how the change of tetrad basis would change components of tensors even without changing coordinates. If we go from one tetrad basis to another the orthonormality must be preserved and thus the metric η_{ab} must remain flat. Of course the group of such transformation is known as Lorentz group and the transformation matrices are known as Lorentz matrices. In contrast to the global Lorentz transformations in special relativity we are now free to have different changes of basis at different points (**local Lorentz transformations**), i.e.

$$\hat{e}_{(a)} \to \hat{e}_{(a')} = \Lambda_{a'}^{\ a}(x) \hat{e}_{(a)} \tag{3.174}$$

where $\Lambda_{a'}^{\ a}(x)$ (actually an inverse Lorentz transformation) is a function of position such that

$$\Lambda_{a'}^{\ a}(x) \Lambda_{b'}^{\ b}(x) \eta_{ab} = \eta_{a'b'} \tag{3.175}$$

at every point. But since the local Lorentz transformations only transform basis, we can still transform the coordinates using the general coordinate transformations. So even if the tensor is expressed in mixed basis, we know exactly how to transform it

$$T^{a'\mu'}_{\ \ \ \ b'\nu'} = \Lambda^{a'}_{\ \ a} \left(\frac{\partial x^{\mu'}}{\partial x^\mu}\right) \Lambda_{b'}^{\ \ b} \left(\frac{\partial x^{\nu'}}{\partial x^\nu}\right) T^{a\mu}_{\ \ \ b\nu}. \tag{3.176}$$

Things become a little bit more complicated with covariant derivatives where the connection coefficients $\Gamma^\lambda_{\mu\nu}$ are replaced with the so-called spin connections $\omega_\mu{}^a{}_b$. For example, the covariant derivative of a (1,1) tensors components in tetrads basis is given by

$$\nabla_\mu X^a_{\ b} = \partial_\mu X^a_{\ b} + \omega_\mu{}^a{}_c X^c_{\ b} - \omega_\mu{}^c{}_b X^a_{\ c}. \tag{3.177}$$

Of course in mixed basis one can get terms with connection coefficients and spin coefficients in the same expression, e.g.

$$\nabla_\mu X^{a\nu} = \partial_\mu X^{a\nu} + {\omega_\mu}^a{}_c X^{c\nu} + \Gamma^\nu_{\mu\lambda} X^{a\lambda}. \tag{3.178}$$

But since the same object (e.g. covariant derivative of vector) can be written in coordinate basis

$$\nabla X = \left(\nabla_\mu X^\lambda\right) \mathrm{d}x^\mu \otimes \partial_\lambda = \left(\partial_\mu X^\lambda + \Gamma^\lambda_{\mu\nu} X^\nu\right) \mathrm{d}x^\mu \otimes \partial_\lambda \tag{3.179}$$

as well as in tetrads basis

$$\begin{aligned} \nabla X &= \left(\nabla_\mu X^a\right) \mathrm{d}x^\mu \otimes \hat{e}_{(a)} \\ &= \left(\partial_\mu X^a + {\omega_\mu}^a{}_b X^b\right) \mathrm{d}x^\mu \otimes \hat{e}_{(a)} \\ &= \left(\partial_\mu \left(e^a_\nu X^\nu\right) + {\omega_\mu}^a{}_b \left(e^b_\nu X^\nu\right)\right) \mathrm{d}x^\mu \otimes \left(e^\lambda_a \partial_\lambda\right) \\ &= \left(e^a_\nu \partial_\mu X^\nu + \partial_\mu e^a_\nu X^\nu + {\omega_\mu}^a{}_b e^b_\nu X^\nu\right) \mathrm{d}x^\mu \otimes \left(e^\lambda_a \partial_\lambda\right) \\ &= \left(\partial_\mu X^\lambda + e^\lambda_a \partial_\mu e^a_\nu X^\nu + {\omega_\mu}^a{}_b e^\lambda_a e^b_\nu X^\nu\right) \mathrm{d}x^\mu \otimes \partial_\lambda \end{aligned} \tag{3.180}$$

we can express the connection coefficients in terms of spin coefficients

$$\Gamma^\lambda_{\mu\nu} = e^\lambda_a \partial_\mu e^a_\nu + {\omega_\mu}^a{}_b e^\lambda_a e^b_\nu \tag{3.181}$$

and

$${\omega_\mu}^a{}_b = e^a_\nu e^\lambda_b \Gamma^\nu_{\mu\lambda} - e^\lambda_b \partial_\mu e^a_\lambda \tag{3.182}$$

which is equivalent to

$$\nabla_\mu e^a_\nu = 0. \tag{3.183}$$

Just like the connection coefficients the spin connections are not legitimate tensors, but the lower Greek index does transform as a one-form. The transformation law for other indices is given by

$${\omega_{\mu'}}^{a'}{}_{b'} = \Lambda^{a'}{}_a \Lambda_{b'}{}^b {\omega_\mu}^a{}_b - \Lambda_{b'}{}^c \partial_\mu \Lambda^{a'}{}_{c.} \tag{3.184}$$

which can be derived by demanding that covariant derivatives transform as tensors.

It is sometime useful to view objects with mixed indices as tensor-valued (described by Latin indices) differential from (described by Greek indices). For example, ${A_{\mu\nu}}^a{}_b$ which in antisymmetric in μ and ν can be thought of as (1,1) tensor valued differential two-form, or ${X_\mu}^a$ can be thought of as a vector valued one form. Then one can show that the exterior derivative

$${(\mathrm{d}X)_{\mu\nu}}^a = \partial_\mu {X_\nu}^a - \partial_\nu {X_\mu}^a \tag{3.185}$$

does not transform as a tensor, but

$$(\mathrm{d}X)_{\mu\nu}{}^{a} + (\omega \wedge X)_{\mu\nu}{}^{a} = \partial_\mu X_\nu{}^a - \partial_\nu X_\mu{}^a + \omega_\mu{}^a{}_b X_\nu{}^b - \omega_\nu{}^a{}_b X_\mu{}^b \quad (3.186)$$

is a legitimate tensor. Then one can view torsion as a vector valued two-form $T_{\mu\nu}{}^a$ and Riemann curvature tensor as a (1,1) valued two-form $R^a{}_{b\mu\nu}$. By suppressing indices of the differential forms the torsions and curvature tensors we can written as

$$T^a = \mathrm{d}e^a + \omega^a{}_b \wedge e^b \quad (3.187)$$

and

$$R^a{}_b = \mathrm{d}\omega^a{}_b + \omega^a{}_c \wedge \omega^c{}_b \quad (3.188)$$

known as **Maurer-Cartan structure equations**. Similarly the Bianchi identities (3.85) and (3.86) can be written respectively as

$$\mathrm{d}T^a + \omega^a{}_b \wedge T^b = R^a{}_b \wedge e^b \quad (3.189)$$

and

$$\mathrm{d}R^a{}_b + \omega^a{}_c \wedge R^c{}_b - R^a{}_c \wedge \omega^c{}_b = 0. \quad (3.190)$$

The metric compatibility (i.e. $\nabla_\mu g_{\nu\lambda} = 0$)

$$0 = \nabla_\mu \eta_{ab} = \partial_\mu \eta_{ab} - \omega_\mu{}^c{}_a \eta_{cb} - \omega_\mu{}^c{}_b \eta_{ac} = -\omega_{\mu ba} - \omega_{\mu ab} \quad (3.191)$$

implies the antisymmetric of the spin connection

$$\omega_{\mu ab} = -\omega_{\mu ba}. \quad (3.192)$$

For Christoffel connection the torsion (3.187) vanishes and thus

$$\mathrm{d}e^a = -\omega^a{}_b \wedge e^b = \omega^{ab} \wedge e_b \quad (3.193)$$

The two equations can be used to solve for the spin connections in terms of tetrads.