General coordinates

February 19, 2013

In order to describe curved spaces, it is necessary to assign a set of labels – coordinates – to each point. There are infinitely many ways to do this, and we need to write our physical laws in a way that doesn't depend on which choice we make. While we often use certain coordinate systems, e.g., Cartesian, polar, spherical, in flat space, we need a more general formulation.

1 General coordinate transformations

1.1 From Lorentz transformations to general coordinate transformations

We have defined (Lorentz) vectors to be objects which transform in the same way as the coordinates,

$$\bar{v}^{\alpha} = \Lambda^{\alpha}_{\beta} v^{\beta}$$

where Λ^{α}_{β} may be any Lorentz transformation. Suppose instead, we have a fully general change of coordinates,

$$y^{\alpha} = y^{\alpha}\left(x^{\beta}\right)$$

Then the differential of the new coordinates changes by

$$dy^{\alpha} = \frac{\partial y^{\alpha}}{\partial x^{\beta}} dx^{\beta}$$

Notice the similarity between this and the Lorentz transformation law. The differentials transform linearly and homogeneously, with the transformation matrix being the Jacobian matrix,

$$J^{\alpha}{}_{\beta} = \frac{\partial y^{\alpha}}{\partial x^{\beta}}$$

This transformation is invertible provided the determinant of $J^{\alpha}{}_{\beta}$, called the Jacobian, is nonzero.

There is another object with a similar transformation law. Consider a parameterized curve, $x^{\alpha}(\lambda)$. The derivative,

$$t^{\alpha} = \frac{dx^{\alpha}}{d\lambda}$$

is tangent to this curve. If we change coordinates as before, we can write the curve in the new coordinates, $y^{\alpha}(\lambda) = y^{\alpha}\left(x^{\beta}(\lambda)\right)$, and we may use the chain rule to write

$$\begin{aligned} t^{\alpha} &= \frac{\partial x^{\alpha}}{\partial y^{\beta}} \frac{dy^{\beta}}{d\lambda} \\ &= \frac{\partial x^{\alpha}}{\partial y^{\beta}} \bar{t}^{\beta} \end{aligned}$$

where $\bar{t}^{\beta} = \frac{dy^{\beta}}{d\lambda}$ is the tangent expressed in terms of the y^{β} coordinates.

Below, we make these notions precise by defining the vector space of differentials and the vector space of tangents, but we will wait until we can make the definitions in a way that works for curved spaces as well as flat ones.

1.2 Derivatives

There is an important difference between the Lorentz transformations and general coordinate transformations. For Lorentz transformations, we were able to differentiate vectors in the usual way to get other vectors. Thus, if we wanted the derivatives of a vector v^α, we could simply compute them all, and have a type-$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ tensor,

$$T^\alpha_{\ \beta} = \frac{\partial v^\alpha}{\partial x^\beta}$$

Then $T^\alpha_{\ \beta}$ is a tensor, because in any other Lorentz frame,

$$\begin{aligned}
\bar{T}^\alpha_{\ \beta} &= \Lambda^\nu_{\ \beta}\frac{\partial}{\partial x^\nu}\left(\Lambda^\alpha_{\ \mu}v^\mu\right) \\
&= \Lambda^\nu_{\ \beta}\Lambda^\alpha_{\ \mu}\frac{\partial}{\partial x^\nu}v^\mu \\
&= \Lambda^\nu_{\ \beta}\Lambda^\alpha_{\ \mu}T^\mu_{\ \nu}
\end{aligned}$$

This happens only because $\Lambda^\alpha_{\ \beta}$ is constant. However, if a vector v^α transforms with a change of coordinates as

$$\bar{v}^\alpha = \frac{\partial y^\alpha}{\partial x^\beta}v^\beta$$

its derivative is not a tensor. Instead,

$$\begin{aligned}
\frac{\partial}{\partial y^\beta}\bar{v}^\alpha &= \frac{\partial}{\partial y^\beta}\left(\frac{\partial y^\alpha}{\partial x^\mu}v^\mu\right) \\
&= \left(\frac{\partial x^\nu}{\partial y^\beta}\frac{\partial}{\partial x^\nu}\right)\left(\frac{\partial y^\alpha}{\partial x^\mu}v^\mu\right) \\
&= \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial}{\partial x^\nu}v^\mu + \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial^2 y^\alpha}{\partial x^\nu \partial x^\mu}v^\mu \\
&= \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial}{\partial x^\nu}v^\mu + \frac{\partial x^\nu}{\partial y^\beta}\left(\frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial x^\mu}{\partial y^\sigma}\right)\frac{\partial^2 y^\sigma}{\partial x^\nu \partial x^\rho}v^\rho \\
&= \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu}\left(\frac{\partial}{\partial x^\nu}v^\mu + \frac{\partial x^\mu}{\partial y^\sigma}\frac{\partial^2 y^\sigma}{\partial x^\nu \partial x^\rho}v^\rho\right)
\end{aligned}$$

so the change is inhomogeneous.

We define the *covariant derivative* in such a way as to correct this problem and produce a tensor. The idea is to add another term to the partial derivative, and let the extra term change in just the right way to cancel the extra, inhomogeneous part. Define

$$D_\beta v^\alpha = \partial_\beta v^\alpha + v^\mu \Gamma^\alpha_{\ \mu\beta}$$

and require $\Gamma^\beta_{\ \mu\alpha}$ to transform so that $D_\alpha v^\beta$ transforms as a tensor when we change coordinates. That is, we require the covariance condition,

$$\bar{D}_\beta \bar{v}^\alpha = \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu}D_\nu v^\mu$$

where

$$\begin{aligned}
\bar{D}_\beta \bar{v}^\alpha &= \frac{\partial}{\partial y^\beta}\bar{v}^\alpha + \bar{v}^\mu \bar{\Gamma}^\alpha_{\ \mu\beta} \\
\bar{v}^\alpha &= \frac{\partial y^\alpha}{\partial x^\mu}v^\mu
\end{aligned}$$

The symbol $\Gamma^\alpha_{\ \mu\beta}$ is called the *connection.*

Substituting into covariance condition,

$$\begin{aligned}
\bar{D}_\beta \bar{v}^\alpha &= \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu} D_\nu v^\mu \\
\frac{\partial x^\nu}{\partial y^\beta}\frac{\partial}{\partial x^\nu}\left(\frac{\partial y^\alpha}{\partial x^\mu} v^\mu\right) + \left(\frac{\partial y^\nu}{\partial x^\mu} v^\mu\right)\bar{\Gamma}^\alpha_{\ \nu\beta} &= \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu}\left(\frac{\partial v^\mu}{\partial x^\nu} + v^\sigma \Gamma^\mu_{\ \sigma\nu}\right) \\
\frac{\partial x^\nu}{\partial y^\alpha}\frac{\partial y^\beta}{\partial x^\mu}\left(\frac{\partial}{\partial x^\nu} v^\mu + \frac{\partial x^\mu}{\partial y^\sigma}\frac{\partial^2 y^\sigma}{\partial x^\nu \partial x^\rho} v^\rho\right) + \left(\frac{\partial y^\nu}{\partial x^\mu} v^\mu\right)\bar{\Gamma}^\alpha_{\ \nu\beta} &= \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu}\left(\frac{\partial v^\mu}{\partial x^\nu} + v^\sigma \Gamma^\mu_{\ \sigma\nu}\right) \\
\frac{\partial x^\nu}{\partial y^\alpha}\frac{\partial y^\beta}{\partial x^\mu}\left(\frac{\partial x^\mu}{\partial y^\sigma}\frac{\partial^2 y^\sigma}{\partial x^\nu \partial x^\rho} v^\rho\right) + \frac{\partial y^\nu}{\partial x^\rho} v^\rho \bar{\Gamma}^\alpha_{\ \nu\beta} &= \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu} v^\rho \Gamma^\mu_{\ \rho\nu}
\end{aligned}$$

This must hold for every vector, v^ρ, so

$$\begin{aligned}
\frac{\partial x^\nu}{\partial y^\alpha}\frac{\partial y^\beta}{\partial x^\mu}\frac{\partial x^\mu}{\partial y^\sigma}\frac{\partial^2 y^\sigma}{\partial x^\nu \partial x^\rho} + \frac{\partial y^\nu}{\partial x^\rho}\bar{\Gamma}^\alpha_{\ \nu\beta} &= \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu}\Gamma^\mu_{\ \rho\nu} \\
\frac{\partial y^\nu}{\partial x^\rho}\bar{\Gamma}^\alpha_{\ \nu\beta} &= \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu}\Gamma^\mu_{\ \rho\nu} - \frac{\partial x^\nu}{\partial y^\alpha}\frac{\partial y^\beta}{\partial x^\mu}\frac{\partial x^\mu}{\partial y^\sigma}\frac{\partial^2 y^\sigma}{\partial x^\nu \partial x^\rho} \\
\frac{\partial x^\rho}{\partial y^\lambda}\frac{\partial y^\nu}{\partial x^\rho}\bar{\Gamma}^\alpha_{\ \nu\beta} &= \frac{\partial x^\rho}{\partial y^\lambda}\frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu}\Gamma^\mu_{\ \rho\nu} - \frac{\partial x^\rho}{\partial y^\lambda}\frac{\partial x^\nu}{\partial y^\alpha}\frac{\partial y^\beta}{\partial x^\mu}\frac{\partial x^\mu}{\partial y^\sigma}\frac{\partial^2 y^\sigma}{\partial x^\nu \partial x^\rho} \\
\bar{\Gamma}^\alpha_{\ \lambda\beta} &= \frac{\partial x^\rho}{\partial y^\lambda}\frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu}\left(\Gamma^\mu_{\ \rho\nu} - \frac{\partial x^\mu}{\partial y^\sigma}\frac{\partial^2 y^\sigma}{\partial x^\nu \partial x^\rho}\right)
\end{aligned}$$

Fortunately, we do not need to use this formula very often.

The covariant derivative satisfies the basic properties of any *derivation*: it is linear and Leibnitz. Linearity is easy to see, but to understand the Leibnitz (product rule) property, we need to consider how the covariant derivative acts on general tensors. We determine this by *requiring* the product rule and applying it to a product of vectors. Thus, we demand that

$$\begin{aligned}
D_\mu\left(u^\alpha v^\beta\right) &= \left(D_\mu u^\alpha\right) v^\beta + u^\alpha D_\mu v^\beta \\
&= \left(\partial_\mu u^\alpha + u^\nu \Gamma^\alpha_{\ \nu\mu}\right) v^\beta + u^\alpha\left(\partial_\mu v^\beta + v^\nu \Gamma^\beta_{\ \nu\mu}\right) \\
&= \left(\partial_\mu u^\alpha\right) v^\beta + u^\alpha \partial_\mu v^\beta + \left(u^\nu v^\beta\right)\Gamma^\alpha_{\ \nu\mu} + \left(u^\alpha v^\nu\right)\Gamma^\beta_{\ \nu\mu} \\
&= \partial_\mu\left(u^\alpha v^\beta\right) + \left(u^\nu v^\beta\right)\Gamma^\alpha_{\ \nu\mu} + \left(u^\alpha v^\nu\right)\Gamma^\beta_{\ \nu\mu}
\end{aligned}$$

If we define a type-$\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ tensor $T^{\alpha\beta} = u^\alpha v^\beta$ then we see from the calculation above that its covariant derivative must be

$$D_\mu T^{\alpha\beta} = \partial_\mu T^{\alpha\beta} + T^{\nu\beta}\Gamma^\alpha_{\ \nu\mu} + T^{\alpha\nu}\Gamma^\beta_{\ \nu\mu}$$

Using linearity, we can sum outer products of pairs of vectors to produce a general rank-2 tensor. It is easy to see the pattern here: each index of $T^{\alpha\beta}$ needs to be contracted with a copy of the connection. A rank-3 tensor will have three terms containing $\Gamma^\beta_{\ \nu\mu}$ and so on for higher rank tensors.

Exercise: Prove that $D_\mu\left(au^\alpha + bv^\alpha\right) = aD_\mu u^\alpha + bD_\mu v^\alpha$ for arbitrary constants a, b and arbitrary vectors u^α, v^α.

1.3 Example: polar coordinates

There is an easier way to do this, but this example shows that the construction so far actually works.

Consider what happens when we change from Cartesian to polar coordinates. We know that in Cartesian coordinates in flat space, the covariant derivative is the same as the usual partial derivative,

$$D_i v^j = \partial_i v^j$$

If we change to polar coordinates, the components of the vector v^i change to

$$\begin{aligned} v^r &= \frac{\partial r}{\partial x} v^x + \frac{\partial r}{\partial y} v^y \\ v^\varphi &= \frac{\partial \varphi}{\partial x} v^x + \frac{\partial \varphi}{\partial y} v^y \end{aligned}$$

Since

$$\begin{aligned} r &= \sqrt{x^2 + y^2} \\ \tan\varphi &= \frac{y}{x} \end{aligned}$$

we have:

$$\begin{aligned} \frac{\partial r}{\partial x} &= \frac{x}{r} \\ \frac{\partial r}{\partial y} &= \frac{y}{r} \end{aligned}$$

for the r derivatives. For φ we take the differential,

$$\begin{aligned} d(\tan\varphi) &= d\left(\frac{y}{x}\right) \\ \frac{1}{\cos^2\varphi} d\varphi &= \frac{1}{x} dy - \frac{y}{x^2} dx \end{aligned}$$

Because $\frac{1}{\cos^2\varphi} = 1 + \tan^2\varphi = 1 + \frac{y^2}{x^2}$, this becomes

$$\begin{aligned} \frac{\partial \varphi}{\partial x} dx + \frac{\partial \varphi}{\partial y} dy &= d\varphi \\ &= \cos^2\varphi \left(\frac{1}{x} dy - \frac{y}{x^2} dx\right) \\ &= \frac{1}{1 + \frac{y^2}{x^2}} \left(\frac{x dy - y dx}{x^2}\right) \\ &= \frac{x dy - y dx}{x^2 + y^2} \end{aligned}$$

and therefore

$$\begin{aligned} \frac{\partial \varphi}{\partial x} &= -\frac{y}{x^2 + y^2} \\ &= -\frac{y}{r^2} \\ \frac{\partial \varphi}{\partial y} &= \frac{x}{r^2} \end{aligned}$$

The components of the vector are therefore

$$\begin{aligned} v^r &= \frac{\partial r}{\partial x} v^x + \frac{\partial r}{\partial y} v^y \\ &= \frac{x v^x}{r} + \frac{y v^y}{r} \\ &= v^x \cos\varphi + v^y \sin\varphi \end{aligned}$$

$$
\begin{aligned}
v^\varphi &= \frac{\partial\varphi}{\partial x}v^x + \frac{\partial\varphi}{\partial y}v^y \\
&= -\frac{y}{r^2}v^x + \frac{x}{r^2}v^y \\
&= -\frac{v^x}{r}\sin\varphi + \frac{v^y}{r}\cos\varphi
\end{aligned}
$$

Now find the connection. We will learn a much easier way to compute this later. Moreover, once we have them, we can use them to compute many properties of the geometry and all covariant derivatives in that geometry. For now, however, we use the general transformation,

$$
\bar{\Gamma}^{\alpha}_{\ \lambda\beta} = \frac{\partial x^\rho}{\partial y^\lambda}\frac{\partial x^\nu}{\partial y^\alpha}\frac{\partial y^\beta}{\partial x^\mu}\left(\Gamma^{\mu}_{\ \rho\nu} - \frac{\partial x^\mu}{\partial y^\sigma}\frac{\partial^2 y^\sigma}{\partial x^\nu \partial x^\rho}\right)
$$

but since the original coordinate system is Cartesian, the original connection vanishes, $\Gamma^{\mu}_{\ \rho\nu} = 0$. Therefore,

$$
\begin{aligned}
\bar{\Gamma}^{\alpha}_{\ \lambda\beta} &= -\frac{\partial x^\rho}{\partial y^\lambda}\frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial x^\mu}{\partial y^\sigma}\frac{\partial^2 y^\sigma}{\partial x^\nu \partial x^\rho} \\
&= -\frac{\partial x^\rho}{\partial y^\lambda}\frac{\partial x^\nu}{\partial y^\beta}\delta^\alpha_\sigma\frac{\partial^2 y^\sigma}{\partial x^\nu \partial x^\rho} \\
&= -\frac{\partial x^\rho}{\partial y^\lambda}\frac{\partial x^\nu}{\partial y^\beta}\frac{\partial^2 y^\alpha}{\partial x^\nu \partial x^\rho}
\end{aligned}
$$

Notice that $\bar{\Gamma}^{\alpha}_{\ \lambda\beta} = \bar{\Gamma}^{\alpha}_{\ \beta\lambda}$ because the transformation depends on $\frac{\partial^2 y^\alpha}{\partial x^\nu \partial x^\rho}$. This is always the case: if the connection is symmetric in one coordinate system, $\Gamma^{\mu}_{\ \alpha\beta} = \Gamma^{\mu}_{\ \beta\alpha}$, it remains symmetric in all coordinate systems. In two dimensions, there are therefore 6 different terms to compute: $\Gamma^{r}_{\ rr}, \Gamma^{r}_{\ r\varphi}, \Gamma^{\varphi}_{\ rr}, \Gamma^{r}_{\ \varphi\varphi}, \Gamma^{\varphi}_{\ r\varphi}, \Gamma^{\varphi}_{\ \varphi\varphi}$. Substituting,

$$
\begin{aligned}
\bar{\Gamma}^{r}_{\ rr} &= -\frac{\partial x^\rho}{\partial r}\frac{\partial x^\nu}{\partial r}\frac{\partial^2 r}{\partial x^\nu \partial x^\rho} \\
&= -\frac{\partial x}{\partial r}\frac{\partial x}{\partial r}\frac{\partial^2 r}{\partial x^2} - 2\frac{\partial x}{\partial r}\frac{\partial y}{\partial r}\frac{\partial^2 r}{\partial x \partial y} - \frac{\partial y}{\partial r}\frac{\partial y}{\partial r}\frac{\partial^2 r}{\partial y^2} \\
&= -\cos^2\varphi\frac{y^2}{r^3} + 2\cos\varphi\sin\varphi\frac{xy}{r^3} - \sin^2\varphi\frac{x^2}{r^3} \\
&= -\cos^2\varphi\sin^2\varphi\left(\frac{1}{r} - \frac{2}{r} - \frac{1}{r}\right) \\
&= 0
\end{aligned}
$$

where we have used

$$
\begin{aligned}
\frac{\partial^2 r}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{x}{r}\right) \\
&= \frac{1}{r} - \frac{x^2}{r^3} \\
&= \frac{y^2}{r^3} \\
\frac{\partial^2 r}{\partial y^2} &= \frac{x^2}{r^3} \\
\frac{\partial^2 r}{\partial y \partial x} &= \frac{\partial}{\partial y}\left(\frac{x}{r}\right) \\
&= -\frac{xy}{r^3}
\end{aligned}
$$

The other components proceed similarly, with the only nonzero cases being $\Gamma^{r}{}_{\varphi\varphi}, \Gamma^{\varphi}{}_{r\varphi}$. There is an easy way to see this, once we find the relationship between the connection and the metric. For now, we compute these two directly:

$$
\begin{aligned}
\Gamma^{r}{}_{\varphi\varphi} &= -\frac{\partial x^{\rho}}{\partial\varphi}\frac{\partial x^{\nu}}{\partial\varphi}\frac{\partial^2 r}{\partial x^{\nu}\partial x^{\rho}} \\
&= -\frac{\partial x}{\partial\varphi}\frac{\partial x}{\partial\varphi}\frac{\partial^2 r}{\partial x^2} - 2\frac{\partial x}{\partial\varphi}\frac{\partial y}{\partial\varphi}\frac{\partial^2 r}{\partial x\partial y} - \frac{\partial y}{\partial\varphi}\frac{\partial y}{\partial\varphi}\frac{\partial^2 r}{\partial y^2} \\
&= -\sin^2\varphi\frac{y^2}{r^3} - 2\cos\varphi\sin\varphi\frac{xy}{r^3} - \cos^2\varphi\frac{x^2}{r^3} \\
&= -r\left(\sin^4\varphi + 2\cos^2\varphi\sin^2\varphi + \cos^4\varphi\right) \\
&= -r\left(\sin^2\varphi + \cos^2\varphi\right)^2 \\
&= -r
\end{aligned}
$$

and

$$
\begin{aligned}
\Gamma^{\varphi}{}_{r\varphi} &= -\frac{\partial x^{\rho}}{\partial r}\frac{\partial x^{\nu}}{\partial\varphi}\frac{\partial^2 \varphi}{\partial x^{\nu}\partial x^{\rho}} \\
&= -\frac{\partial x}{\partial r}\frac{\partial x}{\partial\varphi}\frac{\partial^2\varphi}{\partial x^2} - \frac{\partial x}{\partial r}\frac{\partial y}{\partial\varphi}\frac{\partial^2\varphi}{\partial x\partial y} - \frac{\partial y}{\partial r}\frac{\partial x}{\partial\varphi}\frac{\partial^2\varphi}{\partial y\partial x} + \frac{\partial y}{\partial r}\frac{\partial y}{\partial\varphi}\frac{\partial^2\varphi}{\partial y^2} \\
&= -\frac{\partial x}{\partial r}\frac{\partial x}{\partial\varphi}\frac{2xy}{r^4} + \frac{\partial x}{\partial r}\frac{\partial y}{\partial\varphi}\left(\frac{x^2-y^2}{r^4}\right) + \frac{\partial y}{\partial r}\frac{\partial x}{\partial\varphi}\left(\frac{x^2-y^2}{r^4}\right) + \frac{\partial y}{\partial r}\frac{\partial y}{\partial\varphi}\frac{2xy}{r^4} \\
&= r\cos\varphi\sin\varphi\frac{2xy}{r^4} + r\cos^2\varphi\left(\frac{x^2-y^2}{r^4}\right) - r\sin^2\varphi\left(\frac{x^2-y^2}{r^4}\right) + r\sin\varphi\cos\varphi\frac{2xy}{r^4} \\
&= \frac{1}{r}\left(2\cos^2\varphi\sin^2\varphi + \cos^2\varphi\left(\cos^2\varphi - \sin^2\varphi\right) - \sin^2\varphi\left(\cos^2\varphi - \sin^2\varphi\right) + 2\sin^2\varphi\cos^2\varphi\right) \\
&= \frac{1}{r}\left(\cos^4\varphi + \sin^4\varphi + 2\sin^2\varphi\cos^2\varphi\right) \\
&= \frac{1}{r}
\end{aligned}
$$

where we have used

$$
\begin{aligned}
\frac{\partial^2\varphi}{\partial x^2} &= -\frac{\partial}{\partial x}\left(\frac{y}{r^2}\right) \\
&= \frac{2xy}{r^4} \\
\frac{\partial^2\varphi}{\partial y^2} &= -\frac{2xy}{r^4} \\
\frac{\partial^2\varphi}{\partial y\partial x} &= -\left(\frac{1}{r^2} - \frac{2y^2}{r^4}\right) \\
&= -\frac{x^2-y^2}{r^4}
\end{aligned}
$$

We may now take the covariant derivative of the vector v^i,

$$D_k v^i = \partial_k v^i + v^j \Gamma^{i}{}_{jk}$$

There are four components,

$$D_r v^r = \partial_r v^r + v^r\Gamma^{r}{}_{rr} + v^{\varphi}\Gamma^{r}{}_{\varphi r}$$

$$
\begin{aligned}
&= \partial_r v^r \\
&= \partial_r \left(v^x \cos\varphi + v^y \sin\varphi\right) \\
&= \frac{\partial v^x}{\partial r}\cos\varphi + \frac{\partial v^y}{\partial r}\sin\varphi
\end{aligned}
$$

and

$$
\begin{aligned}
D_r v^\varphi &= \partial_r v^\varphi + v^r \Gamma^\varphi_{\ rr} + v^\varphi \Gamma^\varphi_{\ \varphi r} \\
&= \partial_r v^\varphi + \frac{1}{r} v^\varphi \\
&= \frac{\partial}{\partial r}\left(-\frac{v^x}{r}\sin\varphi + \frac{v^y}{r}\cos\varphi\right) + \frac{1}{r}\left(-\frac{v^x}{r}\sin\varphi + \frac{v^y}{r}\cos\varphi\right) \\
&= -\left(-\frac{v^x}{r^2} + \frac{1}{r}\frac{\partial v^x}{\partial r}\right)\sin\varphi + \left(-\frac{v^y}{r^2} + \frac{1}{r}\frac{\partial v^y}{\partial r}\right)\cos\varphi - \frac{v^x}{r^2}\sin\varphi + \frac{v^y}{r^2}\cos\varphi \\
&= -\frac{\sin\varphi}{r}\frac{\partial v^x}{\partial r} + \frac{\cos\varphi}{r}\frac{\partial v^y}{\partial r}
\end{aligned}
$$

and

$$
\begin{aligned}
D_\varphi v^r &= \partial_\varphi v^r + v^r \Gamma^r_{\ r\varphi} + v^\varphi \Gamma^r_{\ \varphi\varphi} \\
&= \partial_\varphi v^r + r v^\varphi \\
&= \frac{\partial}{\partial \varphi}\left(v^x \cos\varphi + v^y \sin\varphi\right) - r\left(-\frac{v^x}{r}\sin\varphi + \frac{v^y}{r}\cos\varphi\right) \\
&= -v^x \sin\varphi + \frac{\partial v^x}{\partial \varphi}\cos\varphi + v^y\cos\varphi + \frac{\partial v^y}{\partial \varphi}\sin\varphi + v^x \sin\varphi - v^y\cos\varphi \\
&= \frac{\partial v^x}{\partial \varphi}\cos\varphi + \frac{\partial v^y}{\partial \varphi}\sin\varphi
\end{aligned}
$$

and finally,

$$
\begin{aligned}
D_\varphi v^\varphi &= \partial_\varphi v^\varphi + v^r \Gamma^\varphi_{\ r\varphi} + v^\varphi \Gamma^\varphi_{\ \varphi\varphi} \\
&= \partial_\varphi v^\varphi + v^r \Gamma^\varphi_{\ r\varphi} \\
&= \partial_\varphi v^\varphi + \frac{1}{r} v^r \\
&= \partial_\varphi\left(-\frac{v^x}{r}\sin\varphi + \frac{v^y}{r}\cos\varphi\right) + \frac{1}{r}\left(v^x\cos\varphi + v^y\sin\varphi\right) \\
&= -\frac{1}{r}\frac{\partial v^x}{\partial \varphi}\sin\varphi + \frac{1}{r}\frac{\partial v^y}{\partial \varphi}\cos\varphi - \frac{v^x}{r}\cos\varphi - \frac{v^y}{r}\sin\varphi + \frac{1}{r}v^x\cos\varphi + \frac{1}{r}v^y\sin\varphi \\
&= -\frac{1}{r}\frac{\partial v^x}{\partial \varphi}\sin\varphi + \frac{1}{r}\frac{\partial v^y}{\partial \varphi}\cos\varphi
\end{aligned}
$$

2 Covariant derivatives

2.1 The Laplacian

Even in flat space, the covariant derivative is useful. For example, suppose we want to compute the Laplacian in spherical coordinates. We know that the Laplacian may be written as the divergence of the gradient,

$$\nabla^2 f = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} f$$

In general coordinates, the divergence of a vector must be a proper contraction,

$$D_\alpha v^\alpha$$

while the gradient of a function is a form,

$$D_\alpha f = \partial_\alpha f$$

No connection is required for this because the function is a scalar – its value at a point does not change if we change coordinates. To combine these, we need the metric to change the gradient to a vector,

$$D^\alpha f = g^{\alpha\beta} D_\beta f$$

and now it becomes possible to write the Laplacian,

$$D_\alpha D^\alpha f = D_\alpha \left(g^{\alpha\beta} D_\beta f \right)$$

In our example of polar coordinates, the metric is

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$$

and its inverse is

$$g^{ij} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{r^2} \end{pmatrix}$$

We have also found the components of the connection,

$$\Gamma^r_{\ \varphi\varphi} = -r$$

and

$$\Gamma^\varphi_{\ r\varphi} = \frac{1}{r}$$

The Laplacian is therefore,

$$\begin{aligned} D_\alpha D^\alpha f &= D_\alpha \left(g^{\alpha\beta} D_\beta f \right) \\ &= \partial_\alpha \left(g^{\alpha\beta} D_\beta f \right) + \left(g^{\mu\beta} D_\beta f \right) \Gamma^\alpha_{\ \mu\alpha} \\ &= + \left(g^{\mu\beta} D_\beta f \right) \Gamma^\alpha_{\ \mu\alpha} \end{aligned}$$

The first term expands as

$$\begin{aligned} \partial_\alpha \left(g^{\alpha\beta} D_\beta f \right) &= \partial_r \left(g^{r\beta} D_\beta f \right) + \partial_\varphi \left(g^{\varphi\beta} D_\beta f \right) \\ &= \partial_r \left(g^{rr} D_r f + g^{r\varphi} D_\varphi f \right) + \partial_\varphi \left(g^{\varphi r} D_r f + g^{\varphi\varphi} D_\varphi f \right) \\ &= \partial_r \left(\partial_r f + 0 \cdot D_\varphi f \right) + \partial_\varphi \left(0 \cdot \partial_r f + \frac{1}{r^2} \partial_\varphi f \right) \\ &= \frac{\partial^2}{\partial r^2} f + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} f \end{aligned}$$

which we know to be correct.

2.2 The connection at a point

We have shown that the connection transforms as

$$\bar{\Gamma}^\alpha_{\ \lambda\beta} = \frac{\partial x^\rho}{\partial y^\lambda} \frac{\partial x^\nu}{\partial y^\alpha} \frac{\partial y^\beta}{\partial x^\mu} \left(\Gamma^\mu_{\ \rho\nu} - \frac{\partial x^\mu}{\partial y^\sigma} \frac{\partial^2 y^\sigma}{\partial x^\nu \partial x^\rho} \right)$$

Now, we know that a manifold is a space which is R^n to lowest order in a sufficiently small neighborhood of any point $\mathcal{P}$. This means that, again in a sufficiently small neighborhood, we can choose Cartesian coordinates where the connection vanishes. There are deviations as we move away from $\mathcal{P}$, but the fact remains that there exists a coordinate system in which the connection vanishes at $\mathcal{P}$.

We can show this explicitly by asking for a coordinate transformation from a general $\Gamma^{\mu}_{\ \rho\nu}$ to a set of new coordinates where $\bar{\Gamma}^{\alpha}_{\ \lambda\beta}$ is simpler. Imposing this condition gives

$$\bar{\Gamma}^{\alpha}_{\ \lambda\beta} = \frac{\partial x^{\rho}}{\partial y^{\lambda}}\frac{\partial x^{\nu}}{\partial y^{\alpha}}\frac{\partial y^{\beta}}{\partial x^{\mu}}\left(\Gamma^{\mu}_{\ \rho\nu} - \frac{\partial x^{\mu}}{\partial y^{\sigma}}\frac{\partial^2 y^{\sigma}}{\partial x^{\nu}\partial x^{\rho}}\right)$$

and since the matrices

$$\begin{aligned} J^{\alpha}_{\ \ \beta} &= \frac{\partial y^{\alpha}}{\partial x^{\beta}} \\ \bar{J}^{\alpha}_{\ \ \beta} &= \frac{\partial x^{\alpha}}{\partial y^{\beta}} \end{aligned}$$

are invertible, we may eliminate the leading factors to get

$$\begin{aligned} \frac{\partial x^{\mu}}{\partial y^{\sigma}}\frac{\partial^2 y^{\sigma}}{\partial x^{\nu}\partial x^{\rho}} &= \Gamma^{\mu}_{\ \rho\nu} \\ \frac{\partial^2 y^{\alpha}}{\partial x^{\mu}\partial x^{\nu}} - \frac{\partial y^{\alpha}}{\partial x^{\beta}}\Gamma^{\beta}_{\ \nu\mu} &= 0 \end{aligned}$$

Now expand $y^{\alpha}, \bar{\Gamma}^{\alpha}_{\ \lambda\beta}$ and $\Gamma^{\beta}_{\ \nu\mu}$ around the point $\mathcal{P}$ (with coordinates x_0^{α})

$$\begin{aligned} y^{\alpha} &= x_0^{\alpha} + a^{\alpha}_{\ \mu}\left(x^{\mu} - x_0^{\mu}\right) + \frac{1}{2}b^{\alpha}_{\ \mu\nu}\left(x^{\mu} - x_0^{\mu}\right)\left(x^{\nu} - x_0^{\nu}\right) + \frac{1}{3!}c^{\alpha}_{\ \mu\nu\sigma}\left(x^{\mu} - x_0^{\mu}\right)\left(x^{\nu} - x_0^{\nu}\right)\left(x^{\sigma} - x_0^{\sigma}\right)\ldots \\ \Gamma^{\beta}_{\ \nu\mu} &= \Gamma^{\beta}_{\ \nu\mu}(\mathcal{P}) + \left(x^{\rho} - x_0^{\rho}\right)\frac{\partial}{\partial x^{\rho}}\Gamma^{\beta}_{\ \nu\mu}(\mathcal{P}) + \ldots \\ \bar{\Gamma}^{\beta}_{\ \nu\mu} &= \bar{\Gamma}^{\beta}_{\ \nu\mu}(\mathcal{P}) + \left(x^{\rho} - x_0^{\rho}\right)\frac{\partial}{\partial x^{\rho}}\bar{\Gamma}^{\beta}_{\ \nu\mu}(\mathcal{P}) + \ldots \end{aligned}$$

We may choose teh constant coefficients in the expansion of $y^{\alpha}\left(x^{\beta}\right)$ in any way we choose.

Writing out the lowest order terms,

$$\begin{aligned} b^{\alpha}_{\ \mu\nu} - a^{\alpha}_{\ \beta}\Gamma^{\beta}_{\ \nu\mu}(\mathcal{P}) &= \bar{\Gamma}^{\alpha}_{\ \nu\mu}(\mathcal{P}) \\ c^{\alpha}_{\ \mu\nu\sigma}\left(x^{\sigma} - x_0^{\sigma}\right) - b^{\alpha}_{\ \beta\sigma}\left(x^{\sigma} - x_0^{\sigma}\right)\Gamma^{\beta}_{\ \nu\mu}(\mathcal{P}) - a^{\alpha}_{\ \beta}\left(x^{\sigma} - x_0^{\sigma}\right)\frac{\partial}{\partial x^{\sigma}}\Gamma^{\beta}_{\ \nu\mu}(\mathcal{P}) &= \left(x^{\sigma} - x_0^{\sigma}\right)\frac{\partial}{\partial x^{\sigma}}\bar{\Gamma}^{\alpha}_{\ \nu\mu}(\mathcal{P}) \end{aligned}$$

Choose $b^{\alpha}_{\ \mu\nu} = a^{\alpha}_{\ \beta}\Gamma^{\beta}_{\ \nu\mu}(\mathcal{P})$. This makes the new connection vanish at $x^{\alpha} = x_0^{\alpha}$,

$$\bar{\Gamma}^{\beta}_{\ \nu\mu}(\mathcal{P}) = 0$$

At next order we have

$$c^{\alpha}_{\ \mu\nu\sigma}\left(x^{\sigma} - x_0^{\sigma}\right) - a^{\alpha}_{\ \rho}\Gamma^{\rho}_{\ \beta\sigma}(\mathcal{P})\Gamma^{\beta}_{\ \nu\mu}(\mathcal{P})\left(x^{\sigma} - x_0^{\sigma}\right) - a^{\alpha}_{\ \rho}\left(x^{\sigma} - x_0^{\sigma}\right)\frac{\partial}{\partial x^{\sigma}}\Gamma^{\rho}_{\ \nu\mu}(\mathcal{P}) = \left(x^{\sigma} - x_0^{\sigma}\right)\frac{\partial}{\partial x^{\sigma}}\bar{\Gamma}^{\alpha}_{\ \nu\mu}(\mathcal{P})$$

Since $a^{\alpha}_{\ \rho}$ must be invertible, we multiply by its inverse, $\bar{a}^{\lambda}_{\ \alpha}$, leaving

$$\bar{a}^{\lambda}_{\ \alpha}c^{\alpha}_{\ \mu\nu\sigma} - \Gamma^{\lambda}_{\ \beta\sigma}(\mathcal{P})\Gamma^{\beta}_{\ \nu\mu}(\mathcal{P}) - \frac{\partial}{\partial x^{\sigma}}\Gamma^{\lambda}_{\ \nu\mu}(\mathcal{P}) = \bar{a}^{\lambda}_{\ \alpha}\frac{\partial}{\partial x^{\sigma}}\bar{\Gamma}^{\alpha}_{\ \nu\mu}(\mathcal{P})$$

Our only free choice here is $\bar{a}^{\lambda}_{\ \alpha}c^{\alpha}_{\ \mu\nu\sigma}$, but this must be totally symmetric in the three lower indices. We therefore have only four independent choices for the lower three indices, and four for λ, giving a total of 16 constants. The remaining terms on the left include constants $\frac{\partial}{\partial x^{\sigma}}\Gamma^{\lambda}_{\ \nu\mu}(\mathcal{P})$. The symmetry on $\mu\nu$ means that $\Gamma^{\lambda}_{\ \nu\mu}(\mathcal{P})$ contains 40 constants, and the derivatives increase this to 160. Therefore, we do not have enough choice left to eliminate the first derivatives of the new connection, so $\bar{a}^{\lambda}_{\ \alpha}\frac{\partial}{\partial x^{\sigma}}\bar{\Gamma}^{\alpha}_{\ \nu\mu}(\mathcal{P})$ is in general nonzero.

2.3 The covariant derivative of the metric

The covariant derivative is required to satisfy the product rule for differentiation. Therefore, if we differentiate the inner product of two vectors, we have

$$D_\alpha \left(g_{\mu\nu} u^\mu v^\nu\right) = u^\mu v^\nu D_\alpha g_{\mu\nu} + g_{\mu\nu} v^\nu D_\alpha u^\mu + g_{\mu\nu} u^\mu D_\alpha v^\nu$$

and the question immediately arises: what is the covariant derivative of the metric, $D_\alpha g_{\mu\nu}$?

In flat space and Cartesian coordinates, the metric is the identiy matrix, $g_{ij} = \delta_{ij} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$, and its derivatives are zero,

$$\partial_i g_{jk} = 0$$

This relationship is identical to

$$D_i g_{jk} = 0$$

in Cartesian coordinates, but the second form is true independent of coordinate system. Therefore, in flat space we expect the covariant derivative of the metric to vanish.

We have a similar result in curved space because curved manifolds are Euclidean in a neighborhood of a point. In general, the covariant derivative of the metric is

$$D_\alpha g_{\mu\nu} = \partial_\alpha g_{\mu\nu} - g_{\beta\nu} \Gamma^\beta_{\ \mu\alpha} - g_{\mu\beta} \Gamma^\beta_{\ \nu\alpha}$$

As we showed above, at any point $\mathcal{P}$, there exists a change of coordinates

$$y^\alpha = x_0^\alpha + a^\alpha_{\ \mu} \left(x^\mu - x_0^\mu\right) + \frac{1}{2} b^\alpha_{\ \mu\nu} \left(x^\mu - x_0^\mu\right)\left(x^\nu - x_0^\nu\right) + \frac{1}{3!} c^\alpha_{\ \mu\nu\sigma} \left(x^\mu - x_0^\mu\right)\left(x^\nu - x_0^\nu\right)\left(x^\sigma - x_0^\sigma\right) \ldots$$

where we may choose the coefficients $b^\alpha_{\ \mu\nu}$ so that the connection vanishes at $\mathcal{P}$. This choice leaves the linear coefficients, $a^\alpha_{\ \mu}$, free. In these new coordinates, the new metric at $\mathcal{P}$ is given by

$$\begin{aligned} g_{\mu\nu}(y) &= \frac{\partial x^\alpha}{\partial y^\mu} \frac{\partial x^\beta}{\partial y^\nu} g_{\alpha\beta} \\ g_{\mu\nu}(\mathcal{P}) &= \bar{a}^\alpha_{\ \mu} \bar{a}^\beta_{\ \nu} g_{\alpha\beta}(\mathcal{P}) \end{aligned}$$

and we may use the matrix $\bar{a}^\alpha_{\ \mu}$ to transform the original $g_{\alpha\beta}$ to orthonormal form, $\eta_{\alpha\beta}$. Then

$$\begin{aligned} D_\alpha g_{\mu\nu}(y) &= \partial_\alpha \eta_{\mu\nu} - g_{\beta\nu} \Gamma^\beta_{\ \mu\alpha}(\mathcal{P}) - g_{\mu\beta} \Gamma^\beta_{\ \nu\alpha}(\mathcal{P}) \\ &= 0 \end{aligned}$$

Since this is a tensor equation, it must hold in every coordinate system, and we conclude

$$D_\alpha g_{\mu\nu} = 0$$

regardless of our choice of chart.

2.4 Connection in terms of the metric

The vanishing of the covariant derivative of the metric allows us to find an expression for the connection in terms of the metric. Expanding the covariant derivative,

$$\begin{aligned} D_\alpha g_{\mu\nu} &= \partial_\alpha g_{\mu\nu} - g_{\beta\nu} \Gamma^\beta_{\ \mu\alpha} - g_{\mu\beta} \Gamma^\beta_{\ \nu\alpha} \\ &= \partial_\alpha g_{\mu\nu} - \Gamma_{\nu\mu\alpha} - \Gamma_{\mu\nu\alpha} \end{aligned}$$

where we define $\Gamma_{\nu\mu\alpha} \equiv g_{\beta\nu}\Gamma^{\beta}_{\ \mu\alpha}$. Equating to zero and moving the connection terms to the left, we have

$$\Gamma_{\mu\alpha\nu} + \Gamma_{\alpha\mu\nu} = \partial_\nu g_{\alpha\mu}$$

This relationship holds no matter what we name the indices, as long as we maintain the correspondence of free indices across the whole equation. We may therefore equally well write

$$\begin{aligned} \Gamma_{\alpha\nu\mu} + \Gamma_{\nu\alpha\mu} &= \partial_\mu g_{\nu\alpha} \\ \Gamma_{\nu\mu\alpha} + \Gamma_{\mu\nu\alpha} &= \partial_\alpha g_{\mu\nu} \end{aligned}$$

For any choice of α, μ, ν all three of these must hold. Combining the three equations, we add the first two and subtract the third,

$$\begin{aligned} \Gamma_{\mu\alpha\nu} + \Gamma_{\alpha\mu\nu} + \Gamma_{\alpha\nu\mu} + \Gamma_{\nu\alpha\mu} - \Gamma_{\nu\mu\alpha} - \Gamma_{\mu\nu\alpha} &= \partial_\nu g_{\alpha\mu} + \partial_\mu g_{\nu\alpha} - \partial_\alpha g_{\mu\nu} \\ (\Gamma_{\alpha\mu\nu} + \Gamma_{\alpha\nu\mu}) + (\Gamma_{\nu\alpha\mu} - \Gamma_{\nu\mu\alpha}) + (\Gamma_{\mu\alpha\nu} - \Gamma_{\mu\nu\alpha}) &= g_{\alpha\mu,\nu} + g_{\nu\alpha,\mu} - g_{\mu\nu,\alpha} \end{aligned}$$

Now, because the connection is symmetric on the last two indices, $\Gamma_{\alpha\mu\nu} = \Gamma_{\alpha\nu\mu}$, the left side simplifies to

$$\begin{aligned} 2\Gamma_{\alpha\mu\nu} &= g_{\alpha\mu,\nu} + g_{\nu\alpha,\mu} - g_{\mu\nu,\alpha} \\ g^{\beta\alpha}\Gamma_{\alpha\mu\nu} &= \frac{1}{2} g^{\beta\alpha} \left(g_{\alpha\mu,\nu} + g_{\nu\alpha,\mu} - g_{\mu\nu,\alpha} \right) \end{aligned}$$

and therefore,

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

This immensely simplifies computing the connection. For example, in polar coordinates the metric is

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$$

so the only component with nonvanishing derivative is $g_{\varphi\varphi} = r^2$. This means that $g_{\alpha\mu,\nu} + g_{\nu\alpha,\mu} - g_{\mu\nu,\alpha}$ will vanish unless two of the indices are φ and one is r. There are only three ways to make this happen:

$$\begin{aligned} \Gamma_{r\varphi\varphi} &= \frac{1}{2} \left(g_{r\varphi,\varphi} + g_{r\varphi,\varphi} - g_{\varphi\varphi,r} \right) \\ &= \frac{1}{2} \left(-\partial_r r^2 \right) \\ &= -r \\ \Gamma_{\varphi r\varphi} = \Gamma_{\varphi\varphi r} &= \frac{1}{2} \left(g_{\varphi r,\varphi} + g_{\varphi\varphi,r} - g_{r\varphi,\varphi} \right) \\ &= \frac{1}{2} \left(g_{\varphi\varphi,r} \right) \\ &= r \end{aligned}$$

Using the inverse

$$g^{ij} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{r^2} \end{pmatrix}$$

we find the same result we struggled to get before,

$$\begin{aligned} \Gamma^{r}_{\ \varphi\varphi} &= g^{rr}\Gamma_{r\varphi\varphi} + g^{r\varphi}\Gamma_{\varphi\varphi\varphi} \\ &= -r \\ \Gamma^{\varphi}_{\ r\varphi} = \Gamma^{\varphi}_{\ \varphi r} &= g^{\varphi\varphi}\Gamma_{\varphi r\varphi} \\ &= \frac{1}{r} \end{aligned}$$

2.5 Conditions on the metric and curvature at a point

Now that we know that, if $D_\alpha g_{\mu\nu} = 0$, then the connection can be built out of first derivatives of the metric, we can prove the results above in a more direct way. Consider a Taylor series for a coordinate transformation of the metric. Expand the old coordinates in terms of the new, and the old and new metrics,

$$\begin{aligned}
x^\alpha(y) &= x_0^\alpha + a^\alpha_{\ \mu}\left(y^\mu - y_0^\mu\right) + \frac{1}{2} b^\alpha_{\ \mu\nu}\left(y^\mu - y_0^\mu\right)\left(y^\nu - y_0^\nu\right) + \frac{1}{3!} c^\alpha_{\ \mu\nu\sigma}\left(y^\mu - y_0^\mu\right)\left(y^\nu - y_0^\nu\right)\left(y^\sigma - y_0^\sigma\right)\dots \\
g_{\alpha\beta}(x) &= g_{\alpha\beta}(\mathcal{P}) + \partial_\mu g_{\alpha\beta}(\mathcal{P})\left(y^\mu - y_0^\mu\right) + \frac{1}{2}\partial_\nu\partial_\mu g_{\alpha\beta}(\mathcal{P})\left(y^\mu - y_0^\mu\right)\left(y^\nu - y_0^\nu\right) + \dots \\
\tilde{g}_{\alpha\beta}(y) &= \tilde{g}_{\alpha\beta}(\mathcal{P}) + \partial_\mu \tilde{g}_{\alpha\beta}(\mathcal{P})\left(y^\mu - y_0^\mu\right) + \frac{1}{2}\partial_\nu\partial_\mu \tilde{g}_{\alpha\beta}(\mathcal{P})\left(y^\mu - y_0^\mu\right)\left(y^\nu - y_0^\nu\right) + \dots
\end{aligned}$$

where the coefficients in y^α are symmetric,

$$\begin{aligned}
b^\alpha_{\ \mu\nu} &= b^\alpha_{\ \nu\mu} \\
c^\alpha_{\ \mu\nu\sigma} &= c^\alpha_{\ (\mu\nu\sigma)}
\end{aligned}$$

The coordinate transformation matrix is

$$\frac{\partial x^\alpha}{\partial y^\beta} = a^\alpha_{\ \beta} + b^\alpha_{\ \beta\mu}\left(y^\mu - y_0^\mu\right) + \frac{1}{2!} c^\alpha_{\ \beta\nu\sigma}\left(y^\nu - y_0^\nu\right)\left(y^\sigma - y_0^\sigma\right)\dots$$

Then, expanding the new metric in terms of the old, we have

$$\begin{aligned}
\tilde{g}_{\alpha\beta}(y) &= g_{\alpha\beta}(x)\frac{\partial x^\alpha}{\partial y^\beta}\frac{\partial x^\alpha}{\partial y^\beta} \\
\tilde{g}_{\mu\nu}(\mathcal{P}) + \partial_\rho \tilde{g}_{\mu\nu}(\mathcal{P})\left(y^\rho - y_0^\rho\right) & \\
+\frac{1}{2}\partial_\rho\partial_\sigma \tilde{g}_{\mu\nu}(\mathcal{P})\left(y^\rho - y_0^\rho\right)\left(y^\sigma - y_0^\sigma\right) + \dots &= \left(g_{\alpha\beta}(\mathcal{P}) + \partial_\rho g_{\alpha\beta}(\mathcal{P})\left(y^\rho - y_0^\rho\right) + \frac{1}{2}\partial_\sigma\partial_\rho g_{\alpha\beta}(\mathcal{P})\left(y^\rho - y_0^\rho\right)\left(y^\sigma - y_0^\sigma\right) + \dots\right) \\
&\quad \times\left(a^\alpha_{\ \mu} + b^\alpha_{\ \mu\rho}\left(y^\rho - y_0^\rho\right) + \frac{1}{2!} c^\alpha_{\ \mu\rho\sigma}\left(y^\rho - y_0^\rho\right)\left(y^\sigma - y_0^\sigma\right)\dots\right) \\
&\quad \times\left(a^\beta_{\ \nu} + b^\beta_{\ \nu\lambda}\left(y^\lambda - y_0^\lambda\right) + \frac{1}{2!} c^\alpha_{\ \beta\lambda\tau}\left(y^\lambda - y_0^\lambda\right)\left(y^\tau - y_0^\tau\right)\dots\right)
\end{aligned}$$

Collect all zeroth order terms,

$$\tilde{g}_{\mu\nu}(\mathcal{P}) = g_{\alpha\beta}(\mathcal{P})\, a^\alpha_{\ \mu} a^\beta_{\ \nu}$$

and first order terms,

$$\begin{aligned}
\partial_\rho \tilde{g}_{\mu\nu}(\mathcal{P})\left(y^\rho - y_0^\rho\right) &= g_{\alpha\beta}(\mathcal{P})\left(a^\alpha_{\ \mu} b^\beta_{\ \nu\lambda}\left(y^\lambda - y_0^\lambda\right) + b^\alpha_{\ \mu\rho}\left(y^\rho - y_0^\rho\right) a^\beta_{\ \nu}\right) + \partial_\rho g_{\alpha\beta}(\mathcal{P})\left(y^\rho - y_0^\rho\right) a^\alpha_{\ \mu} a^\beta_{\ \nu} \\
\partial_\rho \tilde{g}_{\mu\nu}(\mathcal{P}) &= g_{\alpha\beta}(\mathcal{P})\left(a^\alpha_{\ \mu} b^\beta_{\ \nu\rho} + b^\alpha_{\ \mu\rho} a^\beta_{\ \nu}\right) + \partial_\rho g_{\alpha\beta}(\mathcal{P})\, a^\alpha_{\ \mu} a^\beta_{\ \nu}
\end{aligned}$$

Finally, collect the second order terms,

$$\begin{aligned}
\frac{1}{2}\partial_\rho\partial_\sigma \tilde{g}_{\mu\nu}(\mathcal{P})\left(y^\rho - y_0^\rho\right)\left(y^\sigma - y_0^\sigma\right) &= g_{\alpha\beta}(\mathcal{P})\left(\frac{1}{2!} a^\alpha_{\ \mu} c^\beta_{\ \nu\lambda\tau}\left(y^\lambda - y_0^\lambda\right)\left(y^\tau - y_0^\tau\right)\right. \\
&\quad \left. + b^\alpha_{\ \mu\rho}\left(y^\rho - y_0^\rho\right) b^\beta_{\ \nu\lambda}\left(y^\lambda - y_0^\lambda\right) + \frac{1}{2!} c^\alpha_{\ \mu\rho\sigma} a^\beta_{\ \nu}\left(y^\rho - y_0^\rho\right)\left(y^\sigma - y_0^\sigma\right)\right) \\
&\quad + \partial_\rho g_{\alpha\beta}(\mathcal{P})\left(y^\rho - y_0^\rho\right)\left(a^\alpha_{\ \mu} b^\beta_{\ \nu\lambda}\left(y^\lambda - y_0^\lambda\right) + a^\beta_{\ \nu} b^\alpha_{\ \mu\rho}\left(y^\rho - y_0^\rho\right)\right) \\
&\quad + \frac{1}{2}\partial_\sigma\partial_\rho g_{\alpha\beta}(\mathcal{P})\left(y^\rho - y_0^\rho\right)\left(y^\sigma - y_0^\sigma\right) a^\alpha_{\ \mu} a^\beta_{\ \nu}
\end{aligned}$$

so that

$$\begin{aligned}\frac{1}{2}\partial_\rho\partial_\sigma\tilde{g}_{\mu\nu}(\mathcal{P}) &= g_{\alpha\beta}(\mathcal{P})\left(\frac{1}{2!}a^\alpha_{\ \mu}c^\beta_{\ \nu\rho\sigma} + b^\alpha_{\ \mu\rho}b^\beta_{\ \nu\sigma} + \frac{1}{2!}c^\alpha_{\ \mu\rho\sigma}a^\beta_{\ \nu}\right)\\ &\quad +\partial_\rho g_{\alpha\beta}(\mathcal{P})\left(a^\alpha_{\ \mu}b^\beta_{\ \nu\sigma} + a^\beta_{\ \nu}b^\alpha_{\ \mu\sigma}\right)\\ &\quad +\frac{1}{2}a^\alpha_{\ \mu}a^\beta_{\ \nu}\partial_\sigma\partial_\rho g_{\alpha\beta}(\mathcal{P})\end{aligned}$$

Now we ask how simple we can make the expansion of $\tilde{g}_{\mu\nu}$. First, the transformation

$$\tilde{g}_{\mu\nu}(\mathcal{P}) = g_{\alpha\beta}(\mathcal{P})\, a^\alpha_{\ \mu}a^\beta_{\ \nu}$$

where $a^\alpha_{\ \mu}$ is arbitrary is a general linear transformation of the symmetric matrix $g_{\alpha\beta}(\mathcal{P})$, and this is sufficient to make $\tilde{g}_{\mu\nu}(\mathcal{P})$ orthonormal,

$$\tilde{g}_{\mu\nu}(\mathcal{P}) = \eta_{\mu\nu}$$

This determines the coefficients $a^\alpha_{\ \mu}$ up to a Lorentz transformation. For the next order, we have

$$\partial_\rho\tilde{g}_{\mu\nu}(\mathcal{P}) = g_{\alpha\beta}(\mathcal{P})\left(a^\alpha_{\ \mu}b^\beta_{\ \nu\rho} + b^\alpha_{\ \mu\rho}a^\beta_{\ \nu}\right) + \partial_\rho g_{\alpha\beta}(\mathcal{P})\, a^\alpha_{\ \mu}a^\beta_{\ \nu}$$

and the $b^\alpha_{\ \mu\rho}$ have just enough freedom to make the right side vanish. To see this, define

$$B_{\mu\nu\rho} \equiv g_{\alpha\beta}(\mathcal{P})\, a^\alpha_{\ \mu}b^\beta_{\ \nu\rho}$$

and notice that since $b^\beta_{\ \nu\rho}$ is arbitrary, so is $B_{\mu\nu\rho}$. Then the right side is

$$B_{\mu\nu\rho} + B_{\nu\mu\rho} + \partial_\rho g_{\alpha\beta}(\mathcal{P})\, a^\alpha_{\ \mu}a^\beta_{\ \nu}$$

The sum

$$B_{\mu\nu\rho} + B_{\nu\mu\rho}$$

is symmetric on $\mu\nu$, but otherwise arbitrary, while the symmetry of the metric makes $\partial_\rho g_{\alpha\beta}(\mathcal{P})\, a^\alpha_{\ \mu}a^\beta_{\ \nu}$ symmetric on $\mu\nu$ as well. Therefore, we may choose $B_{\mu\nu\rho}$ to make the right side vanish, so that the new metric has vanishing derivatives at $\mathcal{P}$,

$$\partial_\rho\tilde{g}_{\mu\nu}(\mathcal{P}) = 0$$

Now look at the second order equation. Defining $C_{\mu\nu\rho\sigma} = g_{\alpha\beta}(\mathcal{P})\, a^\alpha_{\ \mu}c^\beta_{\ \nu\rho\sigma}$, the right side is

$$\frac{1}{2!}\left(C_{\mu\nu\rho\sigma} + C_{\nu\mu\rho\sigma}\right) + \partial_\rho g_{\alpha\beta}(\mathcal{P})\left(a^\alpha_{\ \mu}b^\beta_{\ \nu\sigma} + a^\beta_{\ \nu}b^\alpha_{\ \mu\sigma}\right) + b^\alpha_{\ \mu\rho}b^\beta_{\ \nu\sigma} + \frac{1}{2}a^\alpha_{\ \mu}a^\beta_{\ \nu}\partial_\sigma\partial_\rho g_{\alpha\beta}(\mathcal{P})$$

The last term contains $\partial_\sigma\partial_\rho g_{\alpha\beta}(\mathcal{P})$. This is symmetric on $\rho\sigma$ and on $\alpha\beta$ so the 10 components of the metric each have 10 derivatives, giving 100 degrees of freedom. However, $C_{\mu\nu\rho\sigma}$ is built from $c^\beta_{\ \nu\rho\sigma}$ which is totally symmetric on $\nu\rho\sigma$. This means that there are $\frac{4\cdot5\cdot6}{1\cdot2\cdot3} = 20$ independent ways to choose these three indices, for each of the 4 values of β, giving only 80 degrees of freedom. In general, we cannot make this side vanish by any choice of $c^\beta_{\ \nu\rho\sigma}$, and the second derivatives of the metric remain.

In conclusion, our coordinate choice can make the metric orthonormal and the first derivatives of the metric vanish at $\mathcal{P}$,

$$\begin{aligned}\tilde{g}_{\mu\nu}(\mathcal{P}) &= \eta_{\mu\nu}\\ \partial_\rho\tilde{g}_{\mu\nu}(\mathcal{P}) &= 0\end{aligned}$$

The second condition is equivalent to the vanishing of the Christoffel connection, since

$$\begin{aligned}\Gamma^\beta_{\ \mu\nu}(\mathcal{P}) &= \frac{1}{2}\eta^{\beta\alpha}\left(g_{\alpha\mu,\nu}(\mathcal{P}) + g_{\alpha\nu,\mu}(\mathcal{P}) - g_{\mu\nu,\alpha}(\mathcal{P})\right)\\ &= 0\end{aligned}$$

This reproduces the conclusions of the previous sections.