General Relativity and Gravitational Waves

Session 2: General Coordinates

2.1 Overview of this Session

As we saw in the last session, general relativity tells us that gravity results from curved spacetime. We have also seen how to describe the flat spacetime of special relativity using cartesian spatial coordinates and a time coordinate defined by synchronized clocks in an inertial frame. But in curved spacetimes, we cannot use cartesian coordinates. Moreover, since our eventual goal is to calculate how matter curves the spacetime around it, we often do not know the spacetime's geometry *a priori*, and therefore do not know what kind of coordinate system we can use at all!

Our goal in this session is to develop mathematical techniques for writing physical equations in a way that is completely independent of the coordinate system we actually end up using. This will generalize the principle of relativity in such a way that we can write equations that apply even when we have no idea what the underlying geometry of spacetime is.

An overview of the sections in this session follows:

2.2 Definition of a Coordinate Basis. This section will describe a particular way to define bases for a coordinate system that make it easy to generalize last session's tensor mathematics.

2.3 Tensors in a Coordinate Basis. In this section, we will see how to generalize the mathematics of tensors to handle arbitrary coordinate bases. We will discover, however, that the simple gradient ∂_μ of a tensor no longer transforms like a covector.

2.4 The Tensor Gradient. In this section, we will see how to generate a tensor gradient that reduces to the ordinary gradient in cartesian coordinates but but generally transforms like a tensor.

2.5 The Geodesic Equation. The tensor gradient will allow us to define the geodesic as the "straightest possible worldline." We will check this definition in a case where we know what the geodesics are.

2.6 Schwarzschild Geodesics. The Schwarzschild metric is a solution to the Einstein equation in the empty space surrounding a spherical star. This section will present the metric without proof and use the geodesic equation to derive equations of motion in such a spacetime.

2.7 Locally Orthonormal and Locally Inertial Frames. At any location in spacetime, we can define locally orthonormal and locally inertial reference frames that represent the kind of laboratories that we might set up in a given spacetime. This section will discuss how we can set up such coordinate systems and how to calculate quantities that would be measured in such a system.

2.2 Definition of a Coordinate Basis.

A coordinate system is simply an organized scheme for attaching numbers (coordinates) to points in space and/or events in spacetime. The normal approach we have used in special relativity (involving using an orthonormal coordinate lattice and synchronizing clocks using light-flashes) is one way, but by no means the only way, to attach coordinates to events, and this method does not yield self-consistent results in general spacetimes. The only assumptions that we will make here about our coordinate systems are that (1) our spacetime is not so pathologically curved that we cannot treat a sufficiently small patch of it as if it were *not* curved, and (2) our coordinates vary smoothly so that neighboring points have nearly the same coordinates.

To make things simpler, I will use an example of strange coordinates applied to a flat two-dimensional (2D) space. The methods we develop for handling arbitrary coordinates in such a context will end up working just as well for curved spaces in any number of dimensions. I am also not going to emphasize rigor in what follows, but on developing your intuition, which I hope you will find more valuable.

Now, the differential distance ds between two infinitesimally separated points in a two-dimensional space (or the spacetime separation between events in a spacetime) is a coordinate-independent quantity, because we can measure it directly with a ruler (or maybe a clock in spacetime) without having to define a coordinate system at all. The way that we connect arbitrary coordinates to physical reality is by linking the spacetime separation between two infinitesimally-separated points (or events) to their coordinate separations.

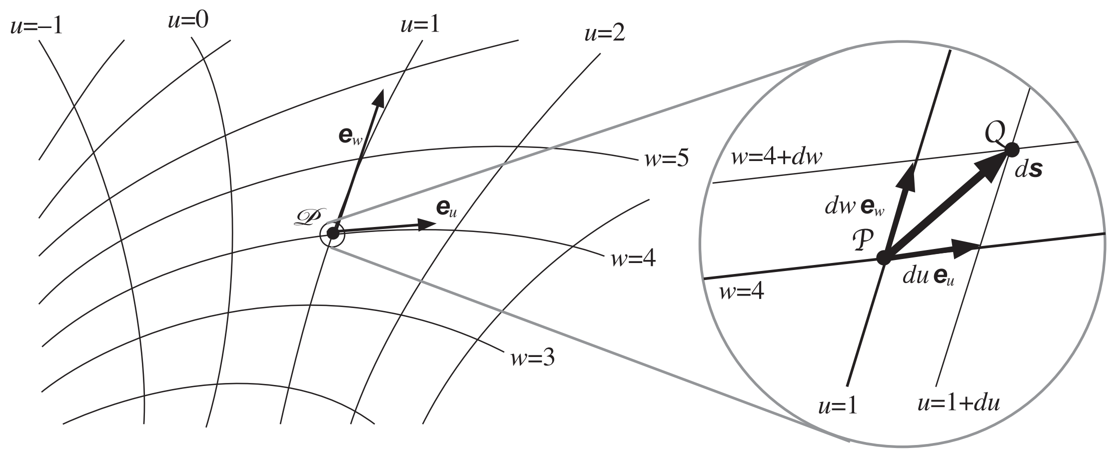

Figure 1: This drawing shows an arbitrary coordinate system in a possibly curved space, a point $\mathcal{P}$, the basis vectors $\boldsymbol{e}_u$, $\boldsymbol{e}_w$ at that point, and a close-up view of how we define the basis vectors so that an infinitesimal displacement $d\boldsymbol{s}$ is the vector sum of the basis vectors times simply du and dw respectively.

So consider a two-dimensional flat space. In such a space, a **cartesian** xy coordinate system is one in which the distance ds between two infinitesimally separated points is everywhere $ds^2 = dx^2 + dy^2$. A curvilinear coordinate system is one where this pythagorean relationship is not true. How can we connect the coordinate-independent distance between two points with their coordinate separations in such a case?

Consider arbitrary coordinates u, w for a 2D space. When using index notation, we will interpret dx^u and dx^w as being equivalent to du and dw, respectively, and we will assume that Greek indices have two possible values u and w. I will also represent abstract vectors in this 2D space with the same bold-face notation as we used for four-vectors last time. This will keep the notation fixed when we generalize to 4D spacetimes.

Now, no matter how our u, w coordinate system is defined, at each point $\mathcal{P}$ in the space, we can define a pair of basis vectors $\boldsymbol{e}_u$, $\boldsymbol{e}_w$ such that

1. $\boldsymbol{e}_u$ points tangent to the w = constant curve toward increasing values of u.

2. $\boldsymbol{e}_w$ points tangent to the u = constant curve toward increasing values of w.

3. We define their lengths so that the displacement vector $d\boldsymbol{s}$ between a point $\mathcal{P}$ at coordinates u, w and an infinitesimally separated neighboring point $\mathcal{Q}$ at coordinates $u + du, w + dw$ can be written

$$d\boldsymbol{s} = du\,\boldsymbol{e}_u + dw\,\boldsymbol{e}_w = dx^\mu \boldsymbol{e}_\mu \qquad \text{(See figure 1.)} \tag{2.1}$$

(Note: the lower index on $\boldsymbol{e}_\mu$ tells us *which basis vector*, not which component, we are talking about.)

Now, this only works for *differential* separations, because the directions of the basis vectors change as we move significant distances. Moreover, vector addition as I have illustrated it in Figure 1 is really only defined in a flat space. In a curved space, the separation between $\mathcal{P}$ and $\mathcal{Q}$ must be so small that we can treat the space as flat, just as we treat a city map as flat on the curved surface of the earth. (Technically, we are doing this vector addition in a flat space that is tangent to the curved space at point $\mathcal{P}$.)

Now, if we define basis vectors this way, then du and dw become the components of $d\boldsymbol{s}$ in that basis and we call the set of basis vectors $\boldsymbol{e}_u$, $\boldsymbol{e}_w$ a **coordinate basis.** A coordinate basis is generally different than the cartesian coordinate basis vectors $\boldsymbol{e}_x$, $\boldsymbol{e}_y$ in that (1) $\boldsymbol{e}_u$ and $\boldsymbol{e}_w$ may not be perpendicular, (2) $\boldsymbol{e}_u$ and $\boldsymbol{e}_w$ may not have unit length, and (3) $\boldsymbol{e}_u$ and/or $\boldsymbol{e}_w$ may change in magnitude and/or direction as we move from point to point. Now, a coordinate basis is not the *only* way to define a set of basis vectors or a coordinate system. For example, standard polar coordinates in two-dimensional flat space and spherical coordinates in a three-dimensional flat space do *not* use a coordinate basis, because the basis vectors in those coordinate systems always have unit magnitude. Using a coordinate basis makes *components* of the displacement vector $d\boldsymbol{s}$ simpler (at the expense of complexity in the basis vectors), something that turns out to be very valuable.

Once we have defined a coordinate basis for coordinates u and w, then we *define* the components of a four-vector $\boldsymbol{A}$ at any point $\mathcal{P}$ to be A^u, A^w such that

$$\boldsymbol{A} = A^u \boldsymbol{e}_u + A^w \boldsymbol{e}_w = A^\mu \boldsymbol{e}_\mu \tag{2.2}$$

This ensures that the components of $d\boldsymbol{s}$ and $\boldsymbol{A}$ transform alike when we change basis systems.

Now, $d\boldsymbol{s} \cdot d\boldsymbol{s}$ is the square of the physical distance between the endpoints of the displacement $d\boldsymbol{s}$:

$$\begin{aligned} ds^2 &= d\boldsymbol{s} \cdot d\boldsymbol{s} = (du\, \boldsymbol{e}_u + dw\, \boldsymbol{e}_w) \cdot (du\, \boldsymbol{e}_u + dw\, \boldsymbol{e}_w) \\ &= du^2 \boldsymbol{e}_u \cdot \boldsymbol{e}_u + du\, dw\, \boldsymbol{e}_u \cdot \boldsymbol{e}_w + dw\, du\, \boldsymbol{e}_w \cdot \boldsymbol{e}_u + dw^2\, \boldsymbol{e}_w \cdot \boldsymbol{e}_w \\ &= dx^\mu dx^\nu\, \boldsymbol{e}_\mu \cdot \boldsymbol{e}_\nu \equiv g_{\mu\nu} dx^\mu dx^\nu \end{aligned} \tag{2.3}$$

The set of four components $g_{\mu\nu} = \boldsymbol{e}_\mu \cdot \boldsymbol{e}_\nu$ represents the **metric tensor** of our two-dimensional coordinate basis. Note that because both the magnitudes and directions of the basis vectors depend on position, the value of $g_{\mu\nu}$ also generally depends on position. But since the dot-product of vectors is commutative, this definition implies that the metric is *always symmetric:* $g_{\alpha\beta} = g_{\beta\alpha}$.

We can easily generalize this to four dimensional spacetime. In spacetime, $g_{\mu\nu} = \boldsymbol{e}_\mu \cdot \boldsymbol{e}_\nu$ (where the indices now range over four values) represents the generalization of the metric tensor $\eta_{\mu\nu}$ introduced in the last session. In *flat* spacetime, we can always find a cartesian coordinate basis where the basis vectors are everywhere orthogonal ($\boldsymbol{e}_\mu \cdot \boldsymbol{e}_\nu = 0$ for $\mu \neq \nu$), and have unit magnitude ($\boldsymbol{e}_\mu \cdot \boldsymbol{e}_\nu = \pm 1$ when $\mu = \nu$, with -1 indicating a time coordinate) at all events and still have the components of the differential four-displacement $d\boldsymbol{s}$ be simply dt, dx, dy, and dz, but this is not generally possible in curved spacetime. The quantity $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ generalizes this idea for an *arbitrary* coordinate basis in an *arbitrary* spacetime. The quantity ds represents the spacetime separation if the events at ends of the differential displacement have a spacelike separation, and $d\tau = \sqrt{-ds^2}$ is the infinitesimal spacetime interval (which is the same as the proper time in the infinitesimal limit) between the events if they have a timelike separation. The point is that the metric connects the arbitrary coordinates with the physical distances or intervals in the physical universe behind that coordinate system. The symmetric metric tensor in spacetime has 10 independent components (the 4 diagonal components and half of the 12 off-diagonal components).

2.2.1 Exercise: The Metric for Spherical Coordinates

Consider θ-ϕ coordinates on the surface of a sphere of radius R, where curves of constant θ and ϕ are lines of latitude and longitude, respectively (but assume that $\theta = 0$ at the north pole, as is normal in physics, rather than at the equator). Note that these curves are perpendicular everywhere but the poles. By considering what the formula for ds^2 between infinitesimally-separated points must be in terms of $d\theta$ and $d\phi$, find the metric components $g_{\theta\theta}, g_{\theta\phi}, g_{\phi\theta}$, and $g_{\phi\phi}$ as a function of position for a coordinate basis based on the coordinates θ and ϕ. Also, what are the lengths of the $\boldsymbol{e}_\theta$ and $\boldsymbol{e}_\phi$ basis vectors as a function of position?

2.3 Tensors in a Coordinate Basis

Again, for the moment, let's go back to considering a two-dimensional, possibly curved surface in space. Consider a general transformation between our original coordinates u, w to new coordinates $p(u,w)$ and $q(u,w)$. The chain rule for partial derivatives implies that infinitesimal changes in the new coordinates are related to changes in the old coordinates as follows:

$$dp = \frac{\partial p}{\partial u} du + \frac{\partial p}{\partial w} dw \quad \text{and} \quad dq = \frac{\partial q}{\partial u} du + \frac{\partial q}{\partial w} dw \tag{2.4}$$

If we consider the p, q coordinates the primed coordinate system and u, w coordinates the unprimed system, then we can write this rule compactly as

$$dx'^{\,\mu} = \frac{\partial x'^{\,\mu}}{\partial x^\nu} dx^\nu \tag{2.5}$$

(with an implicit sum over the ν subscript: we consider a superscript *in the denominator of a partial derivative* to equivalent to a subscript).

Now note that in a coordinate basis, the values of dp and dq are the *actual components* of the infinitesimal displacement vector $d\boldsymbol{s}$ in the primed system and du and dw are the same in the unprimed system. Since by definition, the components of an arbitrary vector $\boldsymbol{A}$ transform in the same ways as the components of the displacement vector, The transformation law for the components of $\boldsymbol{A}$ is

$$A'^{\,\mu} = \frac{\partial x'^{\,\mu}}{\partial x^\nu} A^\nu \tag{2.6}$$

This is true only *if we are using a coordinate basis*, something that we will simply assume from now on.

Note in that context that equation 2.6 looks just like the transformation law for the components of a four-vector in spacetime with $\partial x'^{\nu}/\partial x^{\mu}$ replacing $\Lambda^{\mu}{}_{\nu}$. Indeed, if you take partial derivatives of the Lorentz transformation functions $t'(t,x,y,z)$, $x'(t,x,y,z)$, $y'(t,x,y,z)$, and $z'(t,x,y,z)$, you will see that in fact

$$\frac{\partial x'^{\,\mu}}{\partial x^{\nu}} = \Lambda^{\mu}{}_{\nu} \tag{2.7}$$

for that particular coordinate transformation. We see, therefore, that expressing the transformation coefficients in terms of partial derivatives is consistent with but generalizes the transformation between cartesian coordinates in inertial frames that we considered earlier.

Now, basic partial differential calculus implies that

$$\frac{\partial x'^{\,\mu}}{\partial x^{\nu}}\frac{\partial x^{\nu}}{\partial x'^{\,\alpha}} = \delta^{\mu}{}_{\alpha} \tag{2.8}$$

For example, if we write this out for our $u, w \;\to\; p, q$ transformations, this says that

$$\begin{aligned} \frac{\partial p}{\partial u}\frac{\partial u}{\partial p} + \frac{\partial p}{\partial w}\frac{\partial w}{\partial p} &= \frac{dp}{dp} = 1, & \qquad \frac{\partial p}{\partial u}\frac{\partial u}{\partial q} + \frac{\partial p}{\partial w}\frac{\partial w}{\partial q} &= \frac{dp}{dq} = 0, \\ \frac{\partial q}{\partial u}\frac{\partial u}{\partial p} + \frac{\partial q}{\partial w}\frac{\partial w}{\partial p} &= \frac{dq}{dp} = 0, & \qquad \frac{\partial q}{\partial u}\frac{\partial u}{\partial q} + \frac{\partial q}{\partial w}\frac{\partial w}{\partial q} &= \frac{dq}{dq} = 1 \end{aligned} \tag{2.9}$$

Equation 2.8 basically says that $\partial x'^{\,\mu}/\partial x^{\nu}$ and $\partial x^{\mu}/\partial x'^{\,\nu}$ represent inverse transformations, analogous to $\Lambda^{\mu}{}_{\nu}$ and $(\Lambda^{-1})^{\mu}{}_{\nu}$ in flat spacetime.

Now we can easily state generalized transformation laws for arbitrary tensors: an nth-rank tensor's components transform: we define an ***n*th-rank tensor** $T^{\alpha\cdots}{}_{\beta\cdots}{}^{\gamma\cdots}$ to be an n-index object (with 2^n components in a 2D space and 4^n components in spacetime) that transforms according to

$$T'^{\alpha\cdots}{}_{\beta\cdots}{}^{\gamma\cdots} = \frac{\partial x'^{\,\alpha}}{\partial x^{\mu}}\cdots\frac{\partial x^{\nu}}{\partial x'^{\,\beta}}\cdots\frac{\partial x'^{\,\gamma}}{\partial x^{\sigma}}\cdots T^{\mu\cdots}{}_{\nu\cdots}{}^{\sigma\cdots} \tag{2.10}$$

that is, a partial-derivative factor with the primed coordinate in the numerator for every upper (superscript) index and one with the primed coordinate in the denominator for every lower (subscript) index.

In particular, we can prove that the metric correctly transforms as a tensor with two lower indices as follows. The coordinate-independence of the spacetime separation implies that

$$\begin{aligned} g'_{\mu\nu}dx'^{\,\mu}dx'^{\,\nu} &= g_{\alpha\beta}dx^{\alpha}dx^{\beta} \\ &= g_{\alpha\beta}\left(\frac{\partial x^{\alpha}}{\partial x'^{\,\gamma}}dx'^{\,\alpha}\right)\left(\frac{\partial x^{\beta}}{\partial x'^{\,\sigma}}dx^{\sigma}\right) \\ &= \left(\frac{\partial x^{\alpha}}{\partial x'^{\,\gamma}}\frac{\partial x^{\beta}}{\partial x'^{\,\sigma}}g_{\alpha\beta}\right)dx'^{\,\gamma}dx'^{\,\sigma} \\ &= \left(\frac{\partial x^{\alpha}}{\partial x'^{\,\mu}}\frac{\partial x^{\beta}}{\partial x'^{\,\nu}}g_{\alpha\beta}\right)dx'^{\,\mu}dx'^{\,\nu} \\ \Rightarrow \quad 0 &= \left(\frac{\partial x^{\alpha}}{\partial x'^{\,\mu}}\frac{\partial x^{\beta}}{\partial x'^{\,\nu}}g_{\alpha\beta} - g'_{\mu\nu}\right)dx'^{\,\mu}dx'^{\,\nu} \end{aligned} \tag{2.11}$$

Now, we can't just divide both sides by $dx'^{\,\mu}dx'^{\,\nu}$ because a sum can be zero even if the individual terms in the sum are not. But this relation must be true for *arbitrary* differential displacements. So if I choose the displacement to be entirely in the p direction ($dw = 0$), then the only nonzero term in the sum is the one where $\mu = \nu = p$, so the quantity in parentheses must be 0 when $\mu = \nu = p$. In a similar way, I can choose displacement components to show that all of the other components of that quantity must be zero as well:

$$0 = \left(\frac{\partial x^{\alpha}}{\partial x'^{\,\mu}}\frac{\partial x^{\beta}}{\partial x'^{\,\nu}}g_{\alpha\beta} - g'_{\mu\nu}\right) \quad \Rightarrow \quad g'_{\mu\nu} = \frac{\partial x'^{\,\alpha}}{\partial x^{\mu}}\frac{\partial x^{\nu}}{\partial x'^{\,\beta}}g_{\alpha\beta} \tag{2.12}$$

which is the correct transformation law for a tensor. Using similar methods, you can show that the Kronecker delta is still a tensor, that $g^{\mu\nu}$ defined such that $g^{\mu\nu}g_{\mu\alpha} = \delta^{\nu}{}_{\alpha}$ (the matrix inverse of the metric tensor), is still a tensor, that multiplying by $g_{\mu\nu}$ and $g^{\mu\nu}$ and summing over μ still lowers or raises a tensor index, that contracting over an upper and lower index of a tensor quantity still yields a tensor with rank $n-2$, etc.

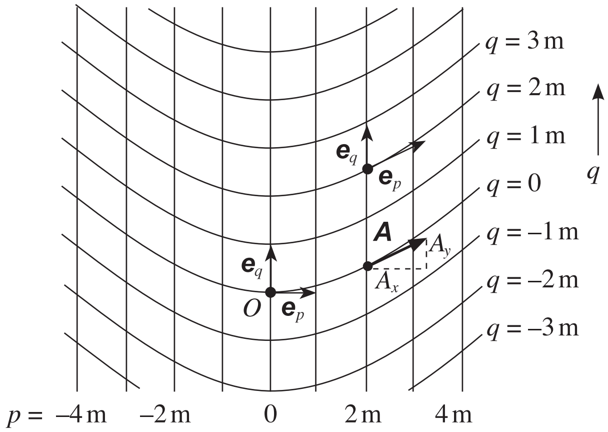

Figure 2: Parabolic coordinates for a flat two-dimensional plane. Adapted from Moore, *A General Relativity Workbook,* University Science Books, 2013, p. 60.)

The gradient of a *scalar* Φ also still yields a tensor (covector): by the chain rule

$$\partial'_\mu \Phi \equiv \frac{\partial \Phi}{\partial x'^{\,\mu}} = \frac{\partial x^\nu}{\partial x'^{\,\mu}} \frac{\partial \Phi}{\partial x^\nu} = \frac{\partial x^\nu}{\partial x'^{\,\mu}} (\partial_\nu \Phi) \tag{2.13}$$

which is the correct transformation law for a tensor with one lower index. However (unlike in the Lorentz transformation case), the gradient of a vector (or any larger-rank tensor) is *not* a tensor:

$$\begin{aligned} \partial'_\mu A^\alpha \equiv \frac{\partial A^\alpha}{\partial x'^{\,\mu}} &= \frac{\partial}{\partial x'^{\,\mu}} \left(\frac{\partial x'^{\,\alpha}}{\partial x^\nu} A^\nu \right) = \frac{\partial x^\beta}{\partial x'^{\,\mu}} \frac{\partial}{\partial x^\beta} \left(\frac{\partial x'^{\,\alpha}}{\partial x^\nu} A^\nu \right) \\ &= \frac{\partial x^\beta}{\partial x'^{\,\mu}} \left(\frac{\partial^2 x'^{\,\alpha}}{\partial x^\beta \partial x^\nu} A^\nu \right) + \frac{\partial x^\beta}{\partial x'^{\,\mu}} \frac{\partial x'^{\,\alpha}}{\partial x^\nu} \left(\frac{\partial A^\nu}{\partial x^\beta} \right) \end{aligned} \tag{2.14}$$

The last term looks like the transformation law for a tensor quantity with one upper and one lower index, but the first term in the last line does not. This term did not arise in the Lorentz transformation case because the Lorentz transformation coefficients are constant, implying that the double partial derivatives are all zero.

This is a serious problem, because many physics equations involve calculating derivatives of quantities that will generalize to vectors or tensors of higher rank. We will address this problem in the next section.

2.3.1 Exercise: Parabolic coordinates.

Consider the parabolic coordinate system p, q shown in figure 2. The transformation functions from ordinary cartesian x, y coordinates are (with c a constant having units of inverse meters):

$$p(x, y) = x \quad \text{and} \quad q(x, y) = y - cx^2 \tag{2.15}$$

(a) Show that the inverse transformation functions are

$$x(p, q) = p \quad \text{and} \quad y(p, q) = cp^2 + q \tag{2.16}$$

(b) The eight partial derivatives $\partial x'^{\,\mu}/\partial x^\nu$ and $\partial x^\mu/\partial x'^{\,\nu}$ in this case are

$$\frac{\partial p}{\partial x} = 1, \quad \frac{\partial p}{\partial y} = 0, \quad \frac{\partial q}{\partial x} = -2cx, \quad \frac{\partial q}{\partial y} = 1 \tag{2.17}$$

$$\frac{\partial x}{\partial p} = 1, \quad \frac{\partial x}{\partial q} = 0, \quad \frac{\partial y}{\partial p} = 2cp, \quad \frac{\partial y}{\partial q} = 1 \tag{2.18}$$

The metric tensor components for cartesian coordinates in space are $g_{xx} = g_{yy} = 1, g_{xy} = g_{yx} = 0$. Use the general tensor transformation rule to show that the metric tensor for p, q coordinates is

$$g'_{\mu\nu} = \begin{bmatrix} 1 + 4c^2p^2 & 2cp \\ 2cp & 1 \end{bmatrix} \tag{2.19}$$

(c) Let a vector $\boldsymbol{A}$ have components $A^p = 1, A^q = 0$ in the p, q coordinate system. Find this vector's components in the x, y coordinate system (as a function of x and y). But show that $A^2 = \boldsymbol{A} \cdot \boldsymbol{A}$ has the same value in both coordinate systems at every position.

2.4 The Tensor Gradient

To see more clearly why the simple gradient of a vector field is not a tensor consider the simple case of a constant vector field $\boldsymbol{A}(x^\mu)$ in a flat two-dimensional space. In flat space at least, we can define "constant" in a coordinate-independent way by saying that at all points, the vector points in the same direction and has the same magnitude. Therefore the physical gradient of such a field should be zero in any coordinate system (because the vector does not change as we change positions). In a cartesian coordinate system, the components A^μ of such a field are constant, so $\partial_\alpha A^\mu = 0$ as expected. But in a curvilinear coordinate system, the components of even a truly constant vector field may *not* be constant, because the basis vectors used to define the components change as one goes from point to point. We need to find a way to correct $\partial_\alpha A^\mu$ in such a coordinate system so as to remove the part due to changes in the basis vectors if we hope to find the true change in the vector function $\boldsymbol{A}(x^\mu)$.

Tensor Gradient of a Vector. Let us define a set of coefficients $\Gamma^\nu_{\nu\alpha}$ (which we call **Christoffel symbols**) at a given point or event $\mathcal{P}$ such that

$$\frac{\partial \boldsymbol{e}_\alpha}{\partial x^\mu} \equiv \Gamma^\nu_{\mu\alpha} \boldsymbol{e}_\nu \tag{2.20}$$

The numerator here is the differential change in the basis vector $\boldsymbol{e}_\alpha$ as we move from point $\mathcal{P}$ to a point a differential displacement dx^μ along a curve where the other coordinates are constant, divided by that differential displacement dx^μ (where α and μ have some specific values here). Since the change $\partial \boldsymbol{e}_\alpha$ is a vector (an arrow with a certain magnitude pointing in a certain direction), we can write that change as a sum over the basis vectors $\boldsymbol{e}_\nu$ evaluated at point $\mathcal{P}$: this is the point of the sum over ν. In a two-dimensional space, there are 8 such coefficients; in four-dimensional spacetime, there are 64 such coefficients.

Now consider a vector field $\boldsymbol{A}$ that is a function of position. The product rule implies that the *true* amount $d\boldsymbol{A}$ that $\boldsymbol{A}$ changes when we move an arbitrary infinitesimal displacement $d\boldsymbol{s}$ (whose components are some specific set of components dx^α) is

$$d\boldsymbol{A} = d(A^\mu \boldsymbol{e}_\mu) = \left(\frac{\partial A^\mu}{dx^\sigma} dx^\sigma\right) \boldsymbol{e}_\mu + A^\mu \frac{\partial \boldsymbol{e}_\mu}{\partial x^\alpha} dx^\alpha \tag{2.21}$$

One can use equation 2.20 and rename some indices to rewrite this as

$$d\boldsymbol{A} = \left[\frac{\partial A^\mu}{\partial x^\alpha} + \Gamma^\mu_{\alpha\nu} A^\nu\right] \boldsymbol{e}_\mu dx^\alpha \equiv (\nabla_\alpha A^\mu) \boldsymbol{e}_\mu dx^\alpha \tag{2.22}$$

The quantities $(\nabla_\alpha A^\mu) dx^\alpha$ for different values of μ are by definition components (in whatever coordinate system we are using) of the *true change* in $\boldsymbol{A}$ for that differential displacement. Therefore, we define

$$\nabla_\alpha A^\mu \equiv \frac{\partial A^\mu}{\partial x^\alpha} + \Gamma^\mu_{\alpha\nu} A^\nu \tag{2.23}$$

to be the **tensor gradient** of the vector field $\boldsymbol{A}$: the term involving the Christoffel symbols corrects the partial derivative of the field for the variations in the basis vectors. The tensor gradient $\nabla_\alpha A^\mu$ must be a (second-rank) tensor, because we see that equation 2.22 implies that $(\nabla_\alpha A^\mu) dx^\alpha$ yields the components of a vector. However, neither the partial derivative alone nor the Christoffel symbol alone are tensors: only the combination given by equation 2.23 is a tensor.

The Tensor Gradient of a Covector. To determine a covector's tensor gradient, we can take advantage of the fact that $A^\mu B_\mu$ is a *scalar*, and the tensor gradient of a scalar is the same as its ordinary gradient:

$$\begin{aligned}\partial_\alpha (A^\mu B_\mu) &= \frac{\partial A^\mu}{\partial x^\alpha} B_\mu + A^\mu \frac{\partial B_\mu}{\partial x^\alpha} \\ &= \nabla_\alpha (A^\mu B_\mu) = (\nabla_\alpha A^\mu) B_\mu + A^\mu (\nabla_\alpha B_\mu)\end{aligned} \tag{2.24}$$

Subtracting one side from the other and substituting what we know $\nabla_\alpha A^\mu$ to be yields

$$\begin{aligned}
0 &= \frac{\partial A^\mu}{\partial x^\alpha} B_\mu + A^\mu \frac{\partial B_\mu}{\partial x^\alpha} - (\nabla_\alpha A^\mu) B_\mu - A^\mu (\nabla_\alpha B_\mu) \\
&= \cancel{\frac{\partial A^\mu}{\partial x^\alpha} B_\mu} + A^\mu \frac{\partial B_\mu}{\partial x^\alpha} - \cancel{\frac{\partial A^\mu}{\partial x^\alpha} B_\mu} - (\Gamma^\mu_{\alpha\nu} A^\nu) B_\mu - A^\mu (\nabla_\alpha B_\mu) \\
&= A^\mu \left[\frac{\partial B_\mu}{\partial x^\alpha} - \Gamma^\nu_{\alpha\mu} B_\nu - \nabla_\alpha B_\mu \right]
\end{aligned} \tag{2.25}$$

where in the last step, I exchanged the μ, ν index names in the third term of the line above so that I could pull out the common factor of A^μ. Since this must be true for arbitrary four-vectors $\boldsymbol{A}$, the quantity in brackets must be zero for all index values μ, implying that

$$\nabla_\alpha B_\mu = \frac{\partial B_\mu}{\partial x^\alpha} - \Gamma^\nu_{\alpha\mu} B_\nu \tag{2.26}$$

The Tensor Gradient of a Tensor. The generalization to arbitrary tensors (e.g., $T^{\mu\nu}{}_\sigma$) is not hard:

$$\nabla_\alpha T^{\mu\nu}{}_\sigma = \frac{\partial T^{\mu\nu}{}_\sigma}{\partial x^\alpha} + \Gamma^\mu_{\alpha\beta} T^{\beta\nu}{}_\sigma + \Gamma^\mu_{\alpha\delta} T^{\mu\delta}{}_\sigma - \Gamma^\gamma_{\alpha\sigma} T^{\mu\nu}{}_\gamma \tag{2.27}$$

The general rule is that the tensor gradient of a tensor is the sum of its ordinary gradient + plus a positive Christoffel symbol (times the tensor) for each upper index (summing over that index and the second lower Christoffel symbol index) and a *negative* Christoffel symbol term (times the tensor) for each lower index (summing over that index and the upper Christoffel symbol index.

Calculating Christoffel Symbols. We will *assume* in what follows that the Christoffel symbols are symmetric in their lower indices (which amounts to assuming that spacetime is free of "torsion"[1]):

$$\Gamma^\alpha_{\mu\nu} = \Gamma^\alpha_{\nu\mu} \tag{2.28}$$

Given this assumption, one can calculate Christoffel symbols in terms of the metric tensor as follows (the proof is one of your homework problems):

$$\Gamma^\alpha_{\mu\nu} = \tfrac{1}{2} g^{\alpha\sigma} \left[\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu} \right] \tag{2.29}$$

This is easier to remember than one might think. A factor of the inverse metric generates the Christoffel symbol's superscript index. The final negative term has the symbol's lower indices as the indices of the metric, and one generates the two positive terms by rotating the last term's indices either right or left.

2.4.1 Exercise: The Tensor Gradient in Parabolic Coordinates.

Consider the parabolic coordinate system described in the previous exercise, where $p(x, y) = x$ and $q(x, y) = y - cx^2$. Consider a truly constant vector field in Cartesian coordinates defined so that $H^x = 1$ and $H^y = 0$.

(a) Find the components of $\boldsymbol{H}$ in the p, q coordinate system.

(b) In the previous exercise, we found the metric tensor for p, q coordinates to be

$$g'_{\mu\nu} = \begin{bmatrix} 1 + 4c^2p^2 & 2cp \\ 2cp & 1 \end{bmatrix} \tag{2.30}$$

Verify (by matrix multiplication) that the inverse metric is given by

$$g'^{\mu\nu} = \begin{bmatrix} 1 & -2cp \\ -2cp & 1 + 4c^2p^2 \end{bmatrix} \tag{2.31}$$

(c) The Christoffel symbols for p, q coordinates are

$$\Gamma^q_{pp} = 2c, \quad \text{all other } \Gamma^\alpha_{\mu\nu} = 0 \tag{2.32}$$

(This makes sense, because we can see from Figure 2 that only the $\boldsymbol{e}_p$ unit vector changes with position, and then only in the q direction as we vary p.) Use $\Gamma^\alpha_{\mu\nu} = \frac{1}{2} g^{\alpha\sigma} \left[\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu} \right]$ to verify that these are the correct values for Γ^q_{pp} and Γ^p_{pp}.

(d) Calculate all four components of $\nabla_\mu H^\nu$ in p, q coordinates. Are the results what you expect?

2.5 The Geodesic Equation.

The four-velocity $\boldsymbol{u} = d\boldsymbol{s}/d\tau$ at every event along a particle's worldline is a vector parallel to the differential displacement $d\boldsymbol{s}$ that the particle moves during a differential proper time $d\tau$ (as measured by a clock traveling with the particle). The four-velocity is therefore always tangent to the particle's worldline. Now a geodesic is by definition "the straightest possible" path at every point. A mathematical way to define such a path is to say that at all events along the particle's worldline, the particle's four-velocity is constant: $d\boldsymbol{u}/d\tau = 0$.

In a curved space or spacetime, what we mean by this is that the vector $\boldsymbol{u}$ does not change direction during an infinitesimal step of proper time $d\tau$, as we would determine in a cartesian coordinate system set up in a patch of area around the particle that is small enough to be considered locally flat. For example, we might map a great circle on the earth's surface by laying out a cartesian coordinate system on a city-sized patch, draw a straight line in that coordinate system a couple of kilometers long, then set up a new coordinate system centered at the endpoint, draw a new straight line segment, and repeat.

In an arbitrary coordinate system in a curved or flat space or spacetime, we can calculate this path mathematically as follows:

$$0 = \frac{d\boldsymbol{u}}{d\tau} = \frac{d}{d\tau}(u^\mu \boldsymbol{e_\mu}) = \frac{du^u}{d\tau}\boldsymbol{e}_\mu + u^\mu \frac{d\boldsymbol{e}_\mu}{d\tau} \tag{2.33}$$

Now, $\boldsymbol{e}_\mu$ depends on τ because the particle's position in spacetime depends on τ and $\boldsymbol{e}_\mu$ depends on position. So using the chain rule and substituting $u^\mu \equiv dx^\mu/d\tau$ into the above, we find that

$$\begin{aligned} 0 &= \frac{d^2x^u}{d\tau^2}\boldsymbol{e}_\mu + \frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\frac{d\boldsymbol{e}_\mu}{dx^\nu} = \frac{d^2x^u}{d\tau^2}\boldsymbol{e}_\mu + \frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\Gamma^\alpha_{\mu\nu}\boldsymbol{e}_\alpha = \left[\frac{d^2x^u}{d\tau^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau}\right]\boldsymbol{e}_\mu \\ \Rightarrow \quad 0 &= \frac{d^2x^u}{d\tau^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} \end{aligned} \tag{2.34}$$

where in going to the last step on the first line, I renamed the bound μ, ν, α indices in the last term to α, β, μ, respectively. The last line follows because a vector will only be zero if each of its components are zero. The last equation is the **geodesic equation**: it represents four second-order differential equations that one can solve to yield the equations $x^\mu(\tau)$ that parametrically describe a geodesic in an arbitrary coordinate system in an arbitrary curved or flat space or spacetime. Note that in flat spacetime, where the metric $g_{\mu\nu} = \eta_{\mu\nu} =$ constant, all the Christoffel symbols (which involve derivatives of the metric) are zero, and the geodesic equation yields $dx^\mu/d\tau^2 = 0$, whose solutions are straight worldlines, as we would expect.

The geodesic equation is one of the two core equations of general relativity. It expresses the first clause of Wheeler's aphorism: "Spacetime tells matter how to move." If we know the metric of spacetime in any coordinate system, we can use this equation to calculate the worldlines of free particles in that spacetime.

2.5.1 Exercise: Geodesics in Parabolic Coordinates.

Consider the parabolic coordinate system for two-dimensional flat space described in the previous two exercises, where $p(x, y) = x$ and $q(x, y) = y - cx^2$. In a two dimensional space like this, we parameterize paths using the arclength s along the path instead of the proper time, so the geodesic equation becomes

$$0 = \frac{d^2x^u}{ds^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} \tag{2.35}$$

and we look for solutions of the form $x^\mu(s)$. We found in the last exercise that the Christoffel symbols for this coordinate system all zero except for $\Gamma^q_{pp} = 2c$.

(a) Find the p and q components of the geodesic equation.

(b) The solution to the p-component equation is easy: $p = as$, where a is a constant of integration, if we define s to be zero where p is zero. Use this to show that the solution to the q-component equation is $q = -ca^2s^2 + bs + q_0$, where b and q_0 are constants of integration.

(c) The transformations back to cartesian coordinates are $x(p, q) = p$ and $y(p, q) = cp^2 + q$. Use these transformations to convert the solutions for $p(s)$ and $q(s)$ to solutions for $x(s)$ and $y(s)$. Argue that the resulting solutions are straight lines (*Hint:* Express y as a function of x.)

2.6 Schwarzschild Geodesics.

As we will see in a subsequent lecture, one can solve the Einstein equation in the vacuum around a spherical star to get the **Schwarzschild metric**:

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \frac{dr^2}{1 - 2GM/r} + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2 \tag{2.36}$$

(meaning that $g_{tt} = -(1 - 2GM/r)$, $g_{rr} = (1 - 2GM/r)^{-1}$, $g_{\theta\theta} = r^2$ and $g_{\phi\phi} = r^2 \sin^2\theta$ and the off-diagonal metric components are zero). This is by no means the only solution even for the vacuum spacetime around a spherical star, since we can redefine coordinates in completely arbitrary ways and so generate an infinite number coordinate systems that describe the same physical reality. But this coordinate system does have some advantages: it is independent of the time coordinate t (in a diagonal metric, the time coordinate is the coordinate whose metric component is negative) and this metric reduces to the metric for flat spacetime (in a spherical coordinate-basis) as $r \to \infty$. I should note that though I am using units where time and space are both measured in meters (so that $c = 1$), I am *not* using units where the universal gravitational constant $G = 1$ as well, as is often done. I am doing this so that connections to Newtonian mechanics will be clearer. But it is good to keep in mind that in this unit system GM has units of meters, and in fact $GM = 1477$ m for a star with the mass of the sun (and is 4.45 mm for the Earth).

It is not my purpose here to discuss the many fascinating features of this metric: it looks like others may talk about that, and the topic is somewhat tangential to the main line of my argument. But I want to introduce this metric so that we can explore the implications of the geodesic equation in a realistic situation (one that even applies to our daily lives!).

Consider a particle at least momentarily rest at some radial coordinate r. The spatial components of such a particle's four-velocity $\boldsymbol{u}$ are zero by definition ($dr/d\tau = d\theta/d\tau = d\phi/d\tau = 0$), but in an arbitrary coordinate system we cannot conclude that the four-velocity's time component $u^t = dt/d\tau$, as we could in special relativity. Rather, we must go back to the basic result for the magnitude of the four-velocity, which (by definition of the differential proper time) still says:

$$\boldsymbol{u} \cdot \boldsymbol{u} \equiv u^\mu g_{\mu\nu} u^\nu = \frac{dx^\mu}{d\tau} g_{\mu\nu} \frac{dx^\nu}{d\tau} = \frac{g_{\mu\nu} dx^\mu dx^\nu}{d\tau^2} = \frac{ds^2}{d\tau^2} = \frac{-d\tau^2}{d\tau^2} = -1 \tag{2.37}$$

For a particle at rest, this equation reduces to

$$-1 = u^\mu g_{\mu\nu} u^\nu = g_{tt}(u^t)^2 + 0 + 0 + 0 \quad \Rightarrow \quad u^t = \frac{dt}{d\tau} = \sqrt{\frac{1}{-g_{tt}}} = \frac{1}{\sqrt{1 - 2GM/r}} \tag{2.38}$$

This equation says that a clock attached to a particle at rest registers a proper time between events along its worldline that is smaller than the Schwarzschild coordinate time t between those events. This one consequence of the curvature of spacetime in the Schwarzschild geometry. We see that t, though it is a "time coordinate," is not the time that just any clock at any finite r would read, but rather a kind of global bookkeeper's time stamp. It only coincides with the time that a clock would directly read if the clock is at rest at infinity.

The r-component of the geodesic equation in this situation says that the particle's radial coordinate proper acceleration is given by

$$\frac{d^2 r}{d\tau^2} = -\Gamma^r_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = -\Gamma^r_{tt} u^t u^t + \text{zeros} \tag{2.39}$$

So to evaluate this radial acceleration, we only need to evaluate

$$\begin{aligned} \Gamma^r_{tt} &= \tfrac{1}{2} g^{r\alpha}(\partial_t g_{t\alpha} + \partial_t g_{\alpha t} - \partial_\alpha g_{tt}) = \tfrac{1}{2} g^{rr}(0 + 0 - \partial_r g_{tt}) \\ &= -\frac{1}{2}\left(1 - \frac{2GM}{r}\right) \frac{d}{dr}\left(-1 + \frac{2GM}{r}\right) = +\left(1 - \frac{2GM}{r}\right) \frac{GM}{r^2} \end{aligned} \tag{2.40}$$

Substituting this into the geodesic equation yields

$$\frac{d^2 r}{d\tau^2} = -\Gamma^r_{tt} u^t u^t = -\frac{GM}{r^2}\left(1 - \frac{2GM}{r}\right) \frac{1}{1 - 2GM/r} = -\frac{GM}{r^2} \tag{2.41}$$

This seems say us that the radial acceleration of a particle initially at rest (according to the particle's own clock time) is exactly what we would expect from Newtonian mechanics. However, this is a *bit* misleading, because radial distance a particle moves between two events as it falls radially is not *quite* the same as the difference dr in the radial coordinate r. In fact, according to the metric, the spatial distance between two events that occur at the same time but at points with a purely radial coordinate separation of dr is

$$ds = \sqrt{ds^2} = \sqrt{g_{\mu\nu}dx^\mu dx^\nu} = \sqrt{g_{rr}dr^2 + \text{zeros}} = \frac{dr}{\sqrt{1 - 2GM/r}} \tag{2.42}$$

However, as long as $r > 2GM$, this distinction is not important. For example, at the surface of the earth, where $r = 6380$ km and $2GM \approx 0.9$ cm, we'd have to keep track better than 8 decimal places in our measurements to notice that the Newtonian result for falling objects is not correct.

Still, there is an important lesson here that I am going to repeat over and over in the days to come. ***Coordinates by themselves have no physical meaning.*** Calling coordinates t, r, θ, ϕ may cause you to *think* that you know what they mean, but we could have called these coordinates anything. ***The only thing that gives coordinates physical meaning is the metric***, which is what anchors human coordinates to physical reality. One must always resort to the metric to learn what your coordinates actually mean.

This is a tricky idea, because we are so used to working with coordinates in flat spacetime (for example, spherical coordinates with an orthonormal basis) that have a reasonably direct intuitive meaning. But in general relativity, we have to give this up and use the metric *exclusively* to learn about what coordinates mean. But if you find this hard to grasp or remember, be comforted. Even Einstein struggled with this in the months leading up to his final November 1915 paper describing the Einstein equation. [2]

2.6.1 Exercise: Free-Fall from Infinity.

It turns out one solution to the full geodesic equation for a radially falling particle is $dr/d\tau = -\sqrt{2GM/r}$. At what value of r is the particle at rest? At what value of r is $dr/d\tau = 1$? Does that necessarily mean that it is traveling at the speed of light?

2.7 Locally Orthonormal and Locally Inertial Frames.

The problem with restricting ourselves to coordinate bases is that the kinds of coordinate systems that we use every day in the laboratory employ orthonormal basis vectors. If we want to know what a particle's energy or velocity or some other physical quantity would be "according to (some specified) observer," we are implicitly assuming that the observer is doing his or her measurements or calculations using a frame having a standard orthonormal basis. The purpose of this section is to explore a trick for transforming tensor quantities to such a frame.

The **local flatness theorem** is important in this context (and will be important to us later as well). A statement of the theorem follows.

> At any given event $\mathcal{P}$ in spacetime, our freedom to choose coordinates allows us to construct a coordinate system where (1) $g_{\mu\nu}(\mathcal{P}) = \eta_{\mu\nu}$ (with six degrees of freedom left over that correspond to arbitrary rotations and Lorentz boosts), (2) all 40 of the independent values of $\partial_\alpha g_{\mu\nu}|_\mathcal{P} = 0$, and (3) all but 20 of the 100 independent values of $\partial_\alpha\partial_\beta g_{\mu\nu}|_\mathcal{P}$ equal to zero.

The proof of this theorem is too involved to go into here: if you are interested, look at pp. 207-209 in my textbook.[3] But the result is very important. It tells us that at every event in spacetime, we can (in principle) construct a reference frame that at least locally (if we don't get too far from event $\mathcal{P}$) behaves like laboratories we are used to.

We call a reference frame where $g_{\mu\nu}(\mathcal{P}) = \eta_{\mu\nu}$ a **locally orthonormal frame (LOF)**, because $\eta_{\mu\nu} = g_{\mu\nu} \equiv \boldsymbol{e}_\mu \cdot \boldsymbol{e}_\nu$ means that the basis vectors for this coordinate system are orthogonal and have unit length, at least at event $\mathcal{P}$. If our coordinate system only satisfies this first criterion but not the second (meaning that we don't bother to set the first derivatives of the metric equal to zero), then at least some Christoffel symbols will not be zero, and geodesics will not be straight lines in our coordinate system. This would be analogous to a laboratory coordinate system at rest on the surface of the Earth. We can definitely set up such a coordinate system (we do it all the time), but geodesics (the worldlines of freely falling particles) are not straight lines in such a coordinate system.

We call a LOF that also satisfies the second criterion (all the first derivatives of the metric are zero) a **locally inertial frame (LIF)**. In this frame, the Christoffel symbols are all zero, and the worldlines of geodesics are (locally) straight lines. Near the earth's surface, such a frame would be a freely-falling frame, such as the reference frame in a falling elevator or the International Space Station. In such a frame, we would need to look at geodesics over fairly large distances (or for fairly long times) to see the tidal effects that distinguish a freely-falling frame from a globally inertial frame in deep space. But no matter how good our coordinate system is, we will eventually see such effects, because we cannot set all of the *second* metric derivatives to zero (and thus first derivatives of the Christoffel symbols) to zero. This means that as we move far enough away from $\mathcal{P}$ in space and/or time, we will eventually see the geodesics begin to converge or diverge as the Christoffel symbols begin to become appreciably nonzero.

But how can we transform quantities expressed in Schwarzschild coordinates into values we would measure in either a LOF or LIF? In principle, if we knew the coordinate transformation equations from Schwarzschild coordinates to the new coordinates, then we could use the normal tensor transformation rule. But it is usually very difficult to derive those transformation equations.

Fortunately, there is another way. Suppose we want to know the components of a four-vector $\boldsymbol{A}$ in a certain LOF, and suppose we know the orthonormal basis vectors $\boldsymbol{e}_\mu$ of the LOF, which we will take to be the primed frame. Since in *any* coordinate basis $\boldsymbol{A} = A^\nu \boldsymbol{e}'_\nu$ and $g'_{\mu\nu} \equiv \boldsymbol{e}'_\mu \cdot \boldsymbol{e}'_\nu$, we have

$$\boldsymbol{e}'_\mu \cdot \boldsymbol{A} = \boldsymbol{e}'_\mu \cdot \boldsymbol{e}'_\nu A'^{\,\nu} = g'_{\mu\nu} A'^{\,\nu} = A'_\mu \tag{2.43}$$

In our particular LOF, $g'_{\mu\nu} = \eta_{\mu\nu}$ by definition, so we can say that

$$\eta^{\alpha\mu}(\boldsymbol{e}'_\mu \cdot \boldsymbol{A}) = \eta^{\alpha\mu} A'_\mu = A'^{\,\alpha} \tag{2.44}$$

So, we can evaluate the components of the vector $\boldsymbol{A}$ in any LOF if we can evaluate the dot product of $\boldsymbol{A}$ with the LOF's basis vectors $\boldsymbol{e}'_\mu$.

The clever trick is that since the dot product must have a coordinate-independent value, we can evaluate the dot product in whatever global coordinate system (such as the Schwarzschild coordinate system) that describes our spacetime. All that we need to is to evaluate the Schwarzschild components of both the LOF's basis vectors $\boldsymbol{e}'_\mu$ and the four-vector $\boldsymbol{A}$ of interest.

Consider the following application. Suppose we have a particle that is falling from rest at infinity in Schwarzschild spacetime, and we would like to know the particle's ordinary velocity according to an observer at rest at a given Schwarzschild coordinate r. In order to do this calculation, we must first find the Schwarzschild coordinates of the observer's LOF basis vectors, which we will call $\boldsymbol{o}_T, \boldsymbol{o}_x, \boldsymbol{o}_y$, and $\boldsymbol{o}_z$, corresponding to the observer's time T and spatial x, y, z coordinates, respectively. (Note that the observer's time coordinate T is *not* the Schwarzschild time coordinate t: this is why I am giving them different names. In what follows, we will assume that A^T, A^x, A^y, A^z refer to the components of $\boldsymbol{A}$ evaluated in the LOF, and $A^t, A^r, A^\theta, A^\phi$ to its components in the Schwarzschild system. This will avoid all the primes.)

We begin by noting that the observer's four-velocity $\boldsymbol{u}_{\rm obs}$ is appropriately normalized for a time basis vector ($\boldsymbol{u}_{\rm obs} \cdot \boldsymbol{u}_{obs} = -1$). Also, the observer is not moving in his or her reference frame, so this vector points purely in the observer's time direction. Therefore, $\boldsymbol{o}_T$ must be $\boldsymbol{u}_{\rm obs}$. What are the Schwarzschild components of this four-vector? These components are $[u^t, u^r, u^\theta, u^\phi]_{\rm obs} = [dt/d\tau,\ dr/d\tau,\ d\theta/d\tau,\ d\phi/d\tau]_{\rm obs}$ by definition. If the observer is at rest relative to the Schwarzschild coordinates, then we know that $dr = d\theta = d\phi = 0 \quad \Rightarrow \quad u^r_{\rm obs} = u^\theta_{\rm obs} = u^\phi_{\rm obs} = 0$. We can then compute $u^t_{\rm obs}$ using $\boldsymbol{u}_{\rm obs} \cdot \boldsymbol{u}_{obs} = -1$:

$$-1 = g_{\mu\nu} u^\mu_{\rm obs} u^\nu_{\rm obs} = g_{tt}(u^t_{\rm obs})^2 \quad \Rightarrow \quad u^t_{\rm obs} = \frac{1}{\sqrt{-g_{tt}}} = \frac{1}{\sqrt{1 - 2GM/r}} \tag{2.45}$$

We know that the Schwarzschild basis vectors are orthogonal at every point (because the Schwarzschild metric is diagonal), so we can conveniently choose the observer's spatial basis vectors to align with the Schwarzchild basis vectors. Let's choose $\boldsymbol{o}_z$ parallel to the Schwarzschild $+\boldsymbol{e}_r$ (that is, "upward"), $\boldsymbol{o}_x$ parallel to $+\boldsymbol{e}_\phi$, and $\boldsymbol{o}_y$ to be parallel to $-\boldsymbol{e}_\theta$ (the minus sign is necessary to make the observer's coordinate system right-handed, as you can check with your fingers). This means that the only nonzero Schwarzschild components of $\boldsymbol{o}_x, \boldsymbol{o}_y$ and $\boldsymbol{o}_z$ are their ϕ, θ, and r components, respectively. The only thing we then need to do is ensure that the observer's basis vectors normalized. For the $\boldsymbol{o}_z$ vector, we have

$$1 = \boldsymbol{o}_z \cdot \boldsymbol{o}_z = g_{\mu\nu}(\boldsymbol{o}_z)^\mu(\boldsymbol{o}_z)^\nu = g_{rr}(\boldsymbol{o}_z)^r(\boldsymbol{o}_z)^r \quad \Rightarrow \quad (\boldsymbol{o}_z)^r = \frac{1}{\sqrt{g_{rr}}} = \sqrt{1 - 2GM/r} \tag{2.46}$$

and all other Schwarzschild components of the basis vectors are zero. One can find the components of the other basis vectors similarly.

Now, as we saw in exercise 2.6.1, a radially falling particle yields $u^r = dr/d\tau = -\sqrt{2GM/r}$ if we assume the particle started at rest at infinity. Then $-1 = \boldsymbol{u} \cdot \boldsymbol{u}$ requires that

$$
\begin{aligned}
-1 = g_{tt}(u^t)^2 + g_{rr}(u^r)^2 &= -\left(1 - \frac{2GM}{r}\right)(u^t)^2 + \left(\frac{1}{1-2GM/r}\right)\frac{2GM}{r} \\
\Rightarrow -\left(1 - \cancel{\frac{2GM}{r}}\right) = -\left(1 - \frac{2GM}{r}\right)^2 (u^t)^2 + \cancel{\frac{2GM}{r}} \quad &\Rightarrow \quad u^t = \frac{1}{1-2GM/r} \qquad (2.47)
\end{aligned}
$$

in this case. This gives us all of the Schwarzschild components of the particle's four-velocity.

So, to summarize, the Schwarzschild coordinates of all the vectors of interest are:

$$[\,(\boldsymbol{o}_T)^t, (\boldsymbol{o}_T)^r, (\boldsymbol{o}_T)^\theta, (\boldsymbol{o}_T)^\phi\,] = \left[\frac{1}{\sqrt{1-2GM/r}},\ 0,\ 0,\ 0\right] \qquad (2.48a)$$

$$[\,(\boldsymbol{o}_z)^t, (\boldsymbol{o}_z)^r, (\boldsymbol{o}_z)^\theta, (\boldsymbol{o}_z)^\phi\,] = \left[\ 0, \sqrt{1-2GM/r},\ 0,\ 0\ \right] \qquad (2.48b)$$

$$[\,u^t, u^r, u^\theta, u^\phi\,] = \left[\frac{1}{1-2GM/r},\ -\frac{2GM}{r},\ 0,\ 0\right] \qquad (2.48c)$$

We can now calculate the components of the particle's four-velocity in the observer's frame by calculating the necessary dot products *in the global Schwarzschild coordinate system*:

$$
\begin{aligned}
u^T &= \eta^{T\mu}\boldsymbol{o}_\mu \cdot \boldsymbol{u} = \eta^{TT}\boldsymbol{o}_T \cdot \boldsymbol{u} = (-1)(\boldsymbol{o}_T)^\alpha g_{\alpha\beta}u^\beta = -(\boldsymbol{o}_T)^t g_{tt}u^t \\
&= +\frac{1}{\sqrt{1-2GM/r}}\left(1 - \frac{2GM}{r}\right)\frac{1}{1-2GM/r} = \frac{1}{\sqrt{1-2GM/r}} \qquad (2.49a)
\end{aligned}
$$

$$
\begin{aligned}
u^z &= \eta^{z\mu}\boldsymbol{o}_\mu \cdot \boldsymbol{u} = \eta^{zz}\boldsymbol{o}_z \cdot \boldsymbol{u} = (+1)(\boldsymbol{o}_z)^\alpha g_{\alpha\beta}u^\beta = +(\boldsymbol{o}_z)^r g_{rr}u^r \\
&= -\sqrt{\frac{2GM}{r}}\frac{1}{1-2GM/r}\sqrt{1-2GM/r} = \frac{-2GM/r}{\sqrt{1-2GM/r}} \qquad (2.49b)
\end{aligned}
$$

Finally, we can evaluate the components of the particle's ordinary speed in the observer's frame as follows:

$$v_x = \frac{u^x}{u^T} = 0, \quad v_y = \frac{u^y}{u^T} = 0, \quad v_z = \frac{u^z}{u^T} = \frac{-2GM/r}{\sqrt{1-2GM/r}}\sqrt{1-2GM/r} = -\frac{2GM}{r} \qquad (2.50)$$

The particle's measured speed approaches that of light as the observer's radial coordinate $r \to 2GM$. (An observer will not measure a speed larger than that of light at a position $r < 2GM$ because an observer cannot be at rest when $r < 2GM$.)

So the general approach to calculating quantities in an observer's LOF or LIF is as follows:

1. Use the observer's four-velocity $\boldsymbol{u}_{\text{obs}}$ as the the observer's time-directed basis vector $\boldsymbol{o}_T$.
2. Construct a set of spatial basis vectors $\boldsymbol{o}_x, \boldsymbol{o}_y, \boldsymbol{o}_z$ such that $\boldsymbol{o}_\mu \cdot \boldsymbol{o}_\nu = \eta_{\mu\nu}$.
3. Find the components of $\boldsymbol{o}_\mu$ in whatever global coordinate system describes spacetime on the large scale.
4. Also determine the components of the four-vector $\boldsymbol{A}$ of interest in that global coordinate system.
5. The components of the four-vector in the observer's system are $A'^\mu = \eta^{\mu\nu}\boldsymbol{o}_\nu \cdot \boldsymbol{A}$, where one evaluates the dot product $\boldsymbol{o}_\mu \cdot \boldsymbol{A} = g_{\alpha\beta}(\boldsymbol{o}_\mu)^\alpha A^\beta$ in the *global* coordinate system.

One can generalize this scheme to tensor quantities as well (see the homework). This is a very powerful scheme for translating abstract tensor quantities expressed in arcane coordinates into quantities that one can more easily interpret physically.

Appendix: More on Schwarzschild Geodesics

In this appendix, I will discuss a more extensive application of the geodesic equation in Schwarzschild spacetime that I included in the original version of the talk but removed from the material I actually discussed in the session for length reasons.

To do more with the geodesic equation, we need to calculate more of the Christoffel symbols, which is simply an example of applying equation 2.29 repeatedly until you have evaluated all 40 that might be independent. There are various labor-saving ways to do this (including using computer tools), but a straightforward and low-tech way to do this is to use something that I call the **Diagonal Metric Worksheet**. This worksheet assumes a diagonal metric of the form $ds^2 = -Ax^0 + Bx^1 + Cx^2 + Dx^3$ (where the numbers are placeholders for coordinate names, not exponents) and that the metric components A, B, C, D can depend on any or all of the coordinates. The worksheet then simply lists all of the Christoffel symbols in terms of the metric components and their derivatives (the latter written in a shorthand form where $A_0 \equiv \partial A/\partial x^0$ and so on. On a copy of the worksheet, one merely writes above each term what the Christoffel symbol evaluates to for whatever the functions for A, B, C, D actually are in a given particular circumstance.

In the Schwarzschild case, we identify $x^0 = t, x^1 = r, x^2 = \theta, x^3 = \phi$, and $A = 1 - 2GM/r, B = (1 - 2GM/r)^{-1}, C = r^2$, and $D = r^2 \sin^2\theta$. First note that no metric component depends on t or ϕ, and only D depends on θ, so a lot of the derivative terms that might appear in general will be zero. So we look on the Diagonal Metric Worksheet, and see, for example that $\Gamma^0_{00} = \Gamma^t_{tt} = \frac{1}{2A}A_0 = 0$, because A does not depend on $x^0 = t$. The next listing on the worksheet tells us that

$$\Gamma^0_{01} = \Gamma^0_{10} = \tfrac{1}{2A}A_1 \quad \Rightarrow \quad \Gamma^t_{tr} = \Gamma^t_{rt} = \frac{1}{2(1-2GM/r)}\frac{\partial}{\partial r}\left(1 - \frac{2GM}{r}\right) = \frac{GM}{r^2(1-2GM/r)} \tag{2.51}$$

and the remaining entries on that line are zero because A does not depend on θ or ϕ. One can go through the worksheet in a similar way to show that the complete set of nonzero Christoffel symbols for the Schwarzschild coordinate system are (in addition to the pair evaluated above):

$$\begin{aligned} &\Gamma^r_{tt} = \Gamma^r_{tt} = \frac{GM}{r^2}\left(1 - \frac{2GM}{r}\right), && \Gamma^r_{rr} = \frac{-GM}{r^2(1-2GM/r)}, \\ &\Gamma^r_{\theta\theta} = -(r - 2GM), && \Gamma^r_{\phi\phi} = -(r-2GM)\sin^2\theta \end{aligned} \tag{2.52}$$

$$\Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{r}, \quad \Gamma^\theta_{\phi\phi} = -\cos\theta\sin\theta \tag{2.53}$$

$$\Gamma^\phi_{r\phi} = \Gamma^\phi_{\phi r} = \frac{1}{r}, \quad \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} - \cot\theta \tag{2.54}$$

Now let's consider a particle in orbit around an object whose exterior geometry is described by the Schwarzschild metric. Since the star and its surrounding spacetime are spherically symmetric, we can take any plane through the star's center to be the equatorial ($\theta = \pi/2$) plane, and if the particle's initial velocity lies in that plane then it must stay on that plane by symmetry. So without loss of generality, we can consider only orbits in the equatorial plane, where $sin^2\theta = 1$ and $d\theta/d\tau = 0$. Now, note under these circumstances that the ϕ component of the geodesic equation implies that

$$0 = \frac{d^2\phi}{d\tau^2} + \Gamma^\phi_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = \frac{d^2\phi}{d\tau^2} + 2\Gamma^\phi_{r\phi}\frac{dr}{d\tau}\frac{d\phi}{d\tau} + 2\Gamma^\phi_{\theta\phi}\frac{d\theta}{d\tau}\frac{d\phi}{d\tau} \tag{2.55}$$

since only these Christoffel symbols with ϕ in the upper position are nonzero. But the last term is zero because $d\theta/d\tau = 0$, and substituting in the value of the remaining Christoffel symbol yields

$$0 = \frac{d^2\phi}{d\tau^2} + \frac{2}{r}\frac{dr}{d\tau}\frac{d\phi}{d\tau} = \frac{1}{r^2}\frac{d}{d\tau}\left(r^2\frac{d\phi}{d\tau}\right) \quad \Rightarrow \quad r^2\frac{d\phi}{d\tau} = \text{constant} \equiv \ell \tag{2.56}$$

This expresses conservation of angular momentum, though the difference between Newtonian time and proper time and the difference between the radial coordinate and radial distance means that we are redefining "angular momentum" somewhat from the Newtonian definition. But this is indeed the conserved quantity that follows from spherical symmetry according to Noether's theorem.

The metric is likewise independent of t, so we would also expect a conserved quantity corresponding to energy. Indeed, if we look at the time component of the geodesic equation in this case, we see that

$$\begin{aligned} 0 &= \frac{d^2t}{d\tau^2} + \Gamma^t_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = \frac{d^2t}{d\tau^2} + 2\Gamma^t_{rt}\frac{dr}{d\tau}\frac{dt}{d\tau} = \frac{d^2t}{d\tau^2} + \frac{2GM}{r^2}\left(1 - \frac{2GM}{r}\right)\frac{dr}{d\tau}\frac{dt}{d\tau} \\ &= \frac{d}{d\tau}\left[\left(1 - \frac{2GM}{r}\right)\frac{dt}{d\tau}\right] \quad \Rightarrow \quad \left(1 - \frac{2GM}{r}\right)\frac{dt}{d\tau} = \text{constant} \equiv e \end{aligned} \tag{2.57}$$

At $r = \infty$, where the Schwarzschild metric becomes the metric of flat spacetime, this becomes simply $e = dt/d\tau$, and (since t really does correspond to time measured by clocks at rest at infinity), this is the equivalent to $e = p^t/m = E/m = 1/\sqrt{1-v^2}$, that is, the particle's relativistic energy per unit mass. The equation above therefore states that this quantity is conserved during an orbiting particle's motion.

Since we are considering only motion in the equatorial plane, we already know that $d\theta/d\tau = 0$, so the θ component of the geodesic equation will tell us nothing useful. We can in principle solve the equation's r component for $dr/d\tau$, but it turns out to be equivalent and quicker to get $dr/d\tau$ from the requirement that $\boldsymbol{u} \cdot \boldsymbol{u} = -1$. The geodesic equation must preserve this magnitude, and once we have solved the other components of the geodesic equation, the remaining component must enforce that restriction. Since the equation $\boldsymbol{u} \cdot \boldsymbol{u} = -1$ involves only first derivatives of $x^\mu(\tau)$, it should be the integral of what the remaining geodesic component requires, so it saves us doing an integral. The remaining components of the particle's four-velocity (remembering that $\sin\theta = 1$ in the equatorial plane) must therefore obey the constraint

$$\begin{aligned}
-1 = g_{\mu\nu}u^\mu u^\nu &= -\left(1 - \frac{2GM}{r}\right)\left(\frac{dt}{d\tau}\right)^2 + \frac{1}{1-2GM/r}\left(\frac{dr}{d\tau}\right)^2 + r^2\left(\frac{d\phi}{d\tau}\right)^2 \\
&= -\left(1 - \frac{2GM}{r}\right)\frac{e^2}{(1-2GM/r)^2} + \frac{1}{1-2GM/r}\left(\frac{dr}{d\tau}\right)^2 + r^2\left(\frac{\ell}{r^2}\right)^2 \\
\Rightarrow \quad -1 + \frac{2GM}{r} &= -e^2 + \left(\frac{dr}{d\tau}\right)^2 + \left(1 - \frac{2GM}{r}\right)\frac{\ell^2}{r^2} \\
\Rightarrow \quad \tfrac{1}{2}(e^2 - 1) &= \frac{1}{2}\left(\frac{dr}{d\tau}\right)^2 - \frac{GM}{r} + \frac{\ell^2}{2r^2} - \frac{GM\ell^2}{r^3}
\end{aligned} \tag{2.58}$$

We recognize in this equation a conservation-of-energy-like equation for the radial kinetic energy per unit mass $\frac{1}{2}(dr/d\tau)^2$, where we have used conservation of angular momentum to absorb the kinetic energy associated with $d\phi/d\tau$ into a r-dependent pseudo-potential term. The equivalent Newtonian equation would be

$$\frac{E}{m} = \frac{1}{2}\left(\frac{dr}{dt}\right)^2 - \frac{GM}{r} + \frac{\ell}{2r^2} \tag{2.59}$$

So (ignoring the subtle differences between the meanings of the r-coordinate and ℓ in the two equations and the distinction between Newtonian time t and proper time τ), the new thing that equation 2.58 adds is the final term. This term (which grows in importance as r becomes small), creates the distinctive features of Einsteinian gravity relative to Newtonian gravity, including the precession of the perihelion, the fact that all orbits are unstable for $r < 6GM$ and so on.

We see that the geodesic equation (in combination with an appropriate metric) does lead to the behavior that we expect of gravity from Newtonian mechanics (with some subtle adjustments). Moreover, the fact that the adjustments can be verified experimentally (for example, by measuring the numerical value of Mercury's perihelion precession rate) lends strong credence to the value of the geodesic hypothesis.

Homework Problems

2.1 What is the metric for an r, θ polar-coordinate system whose basis vectors point in the same directions as those in the usual orthonormal polar coordinates, but whose magnitudes are appropriate for a coordinate basis? What are the magnitudes of the basis vectors in this case?

2.2 We can derive equation 2.29 as follows. Calculate the partial derivative of the metric and use the definition of the metric $g_{\mu\nu} = \boldsymbol{e}_\mu \cdot \boldsymbol{e}_\nu$ and equation 2.20 that defines the Christoffel symbol:

$$\frac{\partial g_{\mu\nu}}{\partial x^\alpha} = \frac{\partial(\boldsymbol{e}\mu \cdot \boldsymbol{e}_\nu)}{\partial x^\alpha} = \frac{\partial \boldsymbol{e}_\mu}{\partial x^\alpha} \cdot \boldsymbol{e}_\nu + \boldsymbol{e}_\mu \cdot \frac{\partial \boldsymbol{e}_\nu}{\partial x^\alpha} = \Gamma^\gamma_{\mu\alpha}\boldsymbol{e}_\gamma \cdot \boldsymbol{e}_\nu + \Gamma^\beta_{\nu\alpha}\boldsymbol{e}_\beta \cdot \boldsymbol{e}_\mu = \Gamma^\beta_{\mu\alpha}g_{\beta\nu} + \Gamma^\gamma_{\nu\alpha}g_{\gamma\mu} \tag{2.60}$$

The following equations are equivalent (I have simply renamed the 3 free lower indices cyclically:

$$\frac{\partial g_{\alpha\mu}}{\partial x^\nu} = \Gamma^\beta_{\alpha\nu}g_{\beta\mu} + \Gamma^\gamma_{\mu\nu}g_{\gamma\alpha} \quad \text{and} \quad \frac{\partial g_{\nu\alpha}}{\partial x^\mu} = \Gamma^\beta_{\nu\mu}g_{\beta\alpha} + \Gamma^\gamma_{\alpha\mu}g_{\gamma\nu} \tag{2.61}$$

Add two of these equations and subtract the third and take advantage of the symmetry of the metric and the symmetry of the Christoffel symbol's lower indices to simplify the result. Then multiply both sides by $\frac{1}{2}$ and the inverse metric, and contract over one upper index of that inverse metric to get equation 2.29.

2.3 Calculate the Christoffel symbols for θ, ϕ coordinates on the surface of a two-dimensional sphere of radius R, where the metric is $g_{\theta\theta} = 1,\ g_{\phi\phi} = R^2 \sin^2\theta,\ g_{\theta\phi} = g_{\phi\theta} = 0$.

2.4 What is the four-velocity of a particle dropped from rest at radius r_0 in Schwarzschild spacetime, as a function of proper time τ, after it is dropped? (Note: Requires material in the appendix.)

2.5 We can evaluate the LOF components of a general tensor as follows.

(a) Consider a covector B. Note that in the global coordinate system we can find its vector components as follows: $B^\alpha = g^{\alpha\beta} B_\beta$. Then use the equation for the transformation for vector components to argue that $B'_\mu = (\boldsymbol{o}_\mu)^\nu B_\nu$.

(b) Now consider a second-rank tensor $\boldsymbol{T}$ that is the tensor product of $\boldsymbol{A}$ and $\boldsymbol{B}$. Show that its components in the LOF must be given by

$$T'^{\mu}{}_{\nu} = \eta^{\mu\alpha} T^{\gamma}{}_{\sigma}\, g_{\gamma\rho} (\boldsymbol{o}_\alpha)^\rho (\boldsymbol{o}_\mu)^\sigma \tag{2.62}$$

(c) Let's assume (plausibly) that this applies to such a second-rank tensor even if it isn't formed of the product of a four-vector and a covector. Using the same general pattern, write down the transformation rule for the fourth-rank tensor $M^{\alpha}{}_{\beta}{}^{\mu}{}_{\nu}$.

2.6 Find the orthonormal basis vectors $\boldsymbol{o}_T, \boldsymbol{o}_x, \boldsymbol{o}_y$ and $\boldsymbol{o}_z$ for a LIF that is freely falling from rest at infinity in Schwarzschild spacetime. (*Hint:* To make it orthogonal to the $\boldsymbol{o}_T$ vector in this case, $\boldsymbol{o}_z$ will have to have a t component as well as an r component. You can determine these components by requiring that $\boldsymbol{o}_T \cdot \boldsymbol{o}_z = 0$ and $\boldsymbol{o}_z \cdot \boldsymbol{o}_z = 1$. The other spatial vectors will not require such a component.)

Notes

[1]The symmetry of the Christoffel symbol cannot be proved from the stated fundamental principles of general relativity: it is a hidden assumption of the theory. However, the kinds of curved spacetimes that result from embedding a surface in a Euclidean spacetime of higher dimension are torsion-free (see Hobson, et al. *General Relativity*, Cambridge, 2006, p. 65). Various theorists have explored the consequences of relaxing that assumption, but there is no consensus about how spacetime torsion might be related the sources of the gravitational field or anything else. In any case, a theory with spacetime torsion is not general relativity.

[2]Pais, *Subtle is the Lord*, Oxford, 1982, pp. 250-256.

[3]Moore, A General Relativity Workbook, University Science Books, 2013. The argument here is actually not an iron-clad "proof," but a strong plausibility argument.