Lecture V: Vectors and vector calculus in curved spacetime

Christopher M. Hirata
Caltech M/C 350-17, Pasadena CA 91125, USA[*]
(Dated: October 12, 2011)

I. OVERVIEW

Thus far we have studied mathematics and physics in flat spacetime extensively. It is time now for a mathematical digression: how do we do geometry and vector calculus in curved spacetime? In some ways, this is analogous to geometry on the surface of a sphere (e.g. the Earth's surface), and we will use this example frequently. However, there are some differences. Most importantly, the unit sphere S^2 is embedded in $\mathbb{R}^3$, and in spherical trigonometry the geometry of $\mathbb{R}^3$ is inherited by S^2. In general relativity, 4-dimensional curved spacetime simply *is* – it is not embedded in any flat higher-dimensional space. Therefore we will need a new set of tools to speak meaningfully of vectors and their derivatives.

The recommended reading for this lecture is:

- MTW §8.1–8.6. This presents an overview of the subject of mathematics in curved spacetime. The lecture will go into slightly more detail than the book, but does neither aspires to mathematical rigor. [If you want the latter, try reading Chapters 2 & 3 of the GR book by Wald.]

Note that we will cover the "Track 1" material in the book first; selected examples of "Track 2" material will be introduced later in the course.

II. MANIFOLDS

We have an intuitive notion of what a curved surface is – but our first step will be to sharpen this definition, and in particular to make sense of the idea without regard to an externally defined flat space. A hint is provided by the familiar way in which we refer to points on S^2 – by their "coordinates," the colatitude θ and longitude ϕ. So one might begin by trying to characterize a curved surface $\mathcal{M}$ by its coordinate system: a point in the manifold corresponds via a *coordinate system* (or *chart*) ψ to a point $(x^1, ...x^n) \in \mathbb{R}^n$. The number of coordinates needed is the *dimension* n of the surface.

The example of S^2 immediately exhibits two problems with this idea. First is that the mapping $\psi : S^2 \rightarrow \mathbb{R}^2$ is multiple-valued. One may describe Pasadena as having colatitude 0.98 (radians) and longitude -2.06; but it also has longitude $-2.06 + 2\pi = 4.22$. So we will restrict the range of a coordinate system ψ to make it single valued, just as we restrict the range of the arcsine "function."

The second problem is more interesting. No coordinate system can cover the entire sphere without coordinate singularities (such as the North and South poles in spherical coordinates). So we will do what was done in your childhood world atlas – form an *atlas*, which is a collection of well-behaved charts $\{\psi_A\}$ that cover the entire surface. The sphere S^2 requires at least 2 charts, but one can have as many as one wants (even infinitely many! – but we won't do that in this class). The requirements we will impose in order for $\mathcal{M}$ to be called a *differentiable manifold* are:

- The domains of the chart cover the entire manifold: $\bigcup_A \mathcal{O}_A = \mathcal{M}$.
- Each chart is a one-to-one onto mapping from some subset of the manifold $\mathcal{O}_A \subseteq \mathcal{M}$ to an open subset $\mathcal{U}_A \subseteq \mathbb{R}^n$. [The "open" requirement is imposed so that the domains must overlap. This way we can't build a cube by using the closed square faces to form 6 charts.]
- The charts are consistent with each other in the following sense: consider the overlap region $\mathcal{O}_A \cap \mathcal{O}_B$ between two charts ψ_A and ψ_B. Then the charts map this to two overlap regions $\mathcal{W}_A = \psi_A[\mathcal{O}_A \cap \mathcal{O}_B]$ and $\mathcal{W}_B = \psi_B[\mathcal{O}_A \cap \mathcal{O}_B]$. There is then a one-to-one onto function $\psi_A \circ \psi_B : \mathcal{W}_B \rightarrow \mathcal{W}_A$ that maps the overlap region from chart B to chart A. We require that this function be smooth – i.e. infinitely many times differentiable, or C^∞.

[*]Electronic address: `chirata@tapir.caltech.edu`

An advantage of this idea is that it is quite clear what is meant by introducing a new coordinate system and changing coordinates: one simply builds a new chart, consistent with the pre-existing ones. The charts also naturally define the topology of the manifold (if it wasn't defined already).

We will consider in this class only manifolds that require a finite number of charts. We will further impose the *Hausdorff condition.* Mathematically, this says: given any two distinct points p and q, one can find two disjoint open sets A and B with $p \in A$ and $q \in B$. The physical meaning of this is that there are no "zippers" or "bifurcations" in spacetime – you can't be running along and suddenly discover that spacetime splits into two sheets.

A manifold is *orientable* if you can choose a set of coordinates everywhere with the same orientation (i.e. the Jacobian of the coordinate transformation $\psi_A \circ \psi_B$ has positive determinant). [Example: S^2.] Otherwise it is non-orientable. [Example: the Möbius strip.]

III. VECTORS, DOT PRODUCTS, AND BASES

A. Generalities

We have previously described a vector as a possible velocity for a particle ("$d\mathcal{P}/d\lambda$"). The existence of a coordinate system allows us to describe a velocity in component language as

$$v^\mu = \frac{dx^\mu}{d\lambda}. \tag{1}$$

The *coordinate basis vectors* $\boldsymbol{e}_\alpha$ are the velocities associated with a particle whose coordinates are fixed except for x^α, which increases at unit rate:

$$x^\beta(\lambda) = x^\beta(\mathcal{P}) + \lambda\delta^\beta_\alpha. \tag{2}$$

As in special relativity, one may define the directional derivative of a scalar f associated with a vector,

$$\nabla_{\boldsymbol{v}} f = \frac{df}{d\lambda}, \tag{3}$$

and the coordinate basis vectors are associated with the directional derivatives

$$\nabla_{\boldsymbol{e}_\alpha} = \frac{\partial}{\partial x^\alpha}. \tag{4}$$

[Note: the mathematical literature often uses the directional derivative as the rigorous definition of a vector, so you will see $\boldsymbol{e}_\alpha \equiv \partial/\partial x^\alpha$.]

This all looks the same as in special relativity, but beware! The coordinates are **not** components of a vector, **despite** having upper indices. (They are up since this is more appropriate than putting them down: infinitesimal displacements dx^α are vectors.) To see this, let's consider what happens when we do a general coordinate transformation (i.e. switch to a different chart) – as usual we will use primed and unprimed labels to describe the two systems. The components of velocity change according to

$$v^{\alpha'} = \frac{dx^{\alpha'}}{d\lambda} = \frac{\partial x^{\alpha'}}{\partial x^\beta}\frac{dx^\beta}{d\lambda} = \frac{\partial x^{\alpha'}}{\partial x^\beta}v^\beta. \tag{5}$$

Therefore, changing the coordinate system introduces a change of basis for the vectors with transformation matrix given by the Jacobian:

$$[\mathbf{L}^{-1}]^{\alpha'}{}_\beta = \frac{\partial x^{\alpha'}}{\partial x^\beta} \qquad \leftrightarrow \qquad L^\beta{}_{\alpha'} = \frac{\partial x^\beta}{\partial x^{\alpha'}}. \tag{6}$$

The coordinates do not in general obey a relation such as Eq. (5):

$$x^{\alpha'} \neq \frac{\partial x^{\alpha'}}{\partial x^\beta}x^\beta. \tag{7}$$

With the vectors so defined, we may also define the other objects used in flat spacetime – 1-forms, basis 1-forms, tensors, tensor components, and so on, and their transformation to different coordinate systems via Eq. (6). Notably absent from this list is differentiation, but the algebra operations are at least straightforward.

B. Dot products

Of special interest is the metric tensor (or dot product operation) $g_{\alpha\beta}$. This is represented as some $n \times n$ symmetric matrix with entries $\boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\beta$ in the coordinate basis. We have not yet defined the dot product – indeed, its existence is not part of the notion of a differentiable manifold – but if we define one then we have a *Riemannian manifold*. In the study of Riemannian manifolds, the dot product is the ultimate arbiter of geometry. For example, we can define the differential ds of arc length (differential separation of two neighboring points x^μ and $x^\mu + dx^\mu$) by defining a velocity vector $\boldsymbol{v}$ for a particle such that it moves from x^μ to $x^\mu + dx^\mu$ in infinitesimal time $d\lambda$. Then

$$ds^2 = |\boldsymbol{v}|^2\, d\lambda^2 = g_{\alpha\beta} v^\alpha v^\beta\, d\lambda^2 = g_{\alpha\beta} \frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\, d\lambda^2 = g_{\alpha\beta} dx^\alpha dx^\beta. \tag{8}$$

This relation – essentially a generalization of the Pythagorean theorem for infinitesimal triangles and with legs aligned with the coordinate directions rather than orthogonal to each other – contains the same information as $g_{\alpha\beta}$. The differential of arc length, or *line element*, is often used to describe a dot product.

C. Example I: Polar coordinates

The very first, and simplest, example we will consider is the Cartesian to polar coordinate transformation in $\mathbb{R}^2$. As you are well aware, one can define Cartesian coordinates (x^1, x^2) but also polar coordinates $(r = x^{1'}, \theta = x^{2'})$ via

$$x^1 = r\cos\theta, \quad x^2 = r\sin\theta \qquad \leftrightarrow \qquad r = \sqrt{(x^1)^2 + (x^2)^2}, \quad \theta = \arctan\frac{x^2}{x^1}. \tag{9}$$

The Jacobian matrix associated with this transformation are:

$$L^\beta{}_{\alpha'} = \begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix}. \tag{10}$$

One can see that $\det \mathbf{L} = r$, so the coordinate transformation is nonsingular except at $r = 0$.

The polar coordinate system admits a pair of basis vectors $\boldsymbol{e}_r$ and $\boldsymbol{e}_\theta$. Using the rule of basis transformations $\mathbf{e}_{i'} = L^j{}_{i'}\mathbf{e}_j$, we see that

$$\mathbf{e}_r = \cos\theta\, \mathbf{e}_1 + \sin\theta\, \mathbf{e}_2 \quad \text{and} \quad \mathbf{e}_\theta = -r\sin\theta\, \mathbf{e}_1 + r\cos\theta\, \mathbf{e}_2. \tag{11}$$

Observe that $\mathbf{e}_r$ and $\mathbf{e}_\theta$ are **not orthonormal** and **not constant** (the notion of "constant" will have limited applicability in curved spacetime). This is a general property of most coordinate systems. It also means that there is a complicated relation between the components of a vector in the polar coordinate basis and e.g. its length. The way one finds this relation is to use the metric tensor, which was $g_{ij} = \delta_{ij}$ in the Cartesian basis. In the general basis it is

$$g_{i'j'} = L^i{}_{i'} L^j{}_{j'} g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}. \tag{12}$$

Therefore the square-norm of a vector is

$$|\boldsymbol{v}|^2 = \boldsymbol{v}\cdot\boldsymbol{v} = g_{i'j'} v^{i'} v^{j'} = (v^r)^2 + r^2 (v^\theta)^2. \tag{13}$$

The line element written in polar coordinates is

$$ds^2 = dr^2 + r^2 d\theta^2, \tag{14}$$

and is what one would usually write down if asked for the "metric" in this coordinate system.

D. Local orthonormal frames

In Euclidean geometry or in special relativity it was possible to have a global orthonormal coordinate system. That is no longer the case. However, we can still build *local orthonormal frames* by choosing basis vectors that are not necessarily coordinate vectors. These will be a set of vectors $\{\boldsymbol{e}_{\hat\alpha}\}$ with

$$\boldsymbol{e}_{\hat i} \cdot \boldsymbol{e}_{\hat j} = \delta_{\hat i \hat j} \text{ (Euclidean signature)} \quad \text{or} \quad \boldsymbol{e}_{\hat\alpha} \cdot \boldsymbol{e}_{\hat\beta} = \eta_{\hat\alpha\hat\beta} \text{ (Minkowski signature)}. \tag{15}$$

The metric is simple in such a frame, and indeed it is convenient to use such a thing to describe measurements. (What is the momentum of a photon seen by an observer? One writes the 4-momentum $\boldsymbol{p}$ in the observer's local orthonormal frame and finds its components.) But in using such a frame, one still makes reference to a coordinate system $\{x^\mu\}$. It is then necessary to describe how the basis vectors of the orthonormal frame relate to those of the coordinate frame, i.e. one must describe the coefficients of the $n \times n$ transformation matrix $\mathbf{L}$.

Example: In polar coordinates, one can write the vectors

$$\boldsymbol{e}_{\hat{r}} = \boldsymbol{e}_r = \frac{\partial}{\partial r} \quad \text{and} \quad \boldsymbol{e}_{\hat{\theta}} = \frac{1}{r}\boldsymbol{e}_\theta = \frac{1}{r}\frac{\partial}{\partial\theta}, \tag{16}$$

i.e. related to the coordinate basis via the transformation matrix

$$\mathbf{L} = \begin{pmatrix} 1 & 0 \\ 0 & r^{-1} \end{pmatrix}. \tag{17}$$

The basis vectors of orthonormal frames in general relativity are often called *vierbeins* or *tetrads* (both indicating the number "4," but by convention often used in $n \neq 4$ dimensions). Because the metric components are trivial in the orthonormal frame, specifying the orthonormal frame, or the transformation matrix $\mathbf{L}$, completely specifies the metric. However, you will note that $\mathbf{L}$ has n^2 components whereas the metric had $\frac{1}{2}n(n+1)$ components. This is because $\mathbf{L}$ contains $\frac{1}{2}n(n-1)$ extra pieces of information: it specifies not just the geometry, but also chooses a particular reference frame at each location.

E. Example II: Unit sphere

Let us now return to the unit sphere, S^2, which is a manifold of dimension $n = 2$, defined by embedding in $\mathbb{R}^3$. We will consider a patch of the sphere that can be covered by the usual coordinate system (θ, ϕ). [At least 3 charts of this type are required to cover all of S^2 but we need consider only one.] The coordinate basis vectors are then $\boldsymbol{e}_\theta$ and $\boldsymbol{e}_\phi$.

To determine the metric tensor in this coordinate system, we need to compute the dot products of all of the basis vectors. Recall that $\boldsymbol{e}_\theta$ is the velocity of a particle with $d\theta/d\lambda = 1$ and $d\phi/d\lambda = 0$; it is easily seen that the magnitude of this velocity is unity, so

$$g_{\theta\theta} = \boldsymbol{e}_\theta \cdot \boldsymbol{e}_\theta = 1. \tag{18}$$

On the other hand, $\boldsymbol{e}_\phi$ is the velocity of a particle with $d\phi/d\lambda = 1$ and $d\theta/d\lambda = 0$. Using standard geometry one sees that the magnitude of the velocity is $\sin\theta$, so

$$g_{\phi\phi} = \boldsymbol{e}_\phi \cdot \boldsymbol{e}_\phi = \sin^2\theta. \tag{19}$$

Finally, these two velocity vectors are orthogonal, so $g_{\theta\phi} = 0$. Thus see that the metric tensor components are

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}, \tag{20}$$

and the line element is

$$ds^2 = d\theta^2 + \sin^2\theta\, d\phi^2. \tag{21}$$

The latter is of course the most common form as it is simple and compact to write.

One may define a local orthonormal frame via

$$\boldsymbol{e}_{\hat{\theta}} = \boldsymbol{e}_\theta \quad \text{and} \quad \boldsymbol{e}_{\hat{\phi}} = \csc\theta\, \boldsymbol{e}_\phi. \tag{22}$$

Generally, defining orthonormal basis vectors in the coordinate directions (but with different normalization than the coordinate basis vectors) is possible if $g_{\alpha\beta}$ is diagonal. This is true for the standard forms of many metrics (Friedmann-Robertson-Walker, Schwarzschild, Tolman-Oppenheimer-Volkoff) but **not** in general. In the general case the local orthonormal frame must contain linear combinations of the coordinate basis vectors.

IV. VECTOR CALCULUS

The above machinery suffices to do vector algebra in curved spacetime. It's remarkably simple. But what about vector calculus? This requires us to take gradients. We will also want to do tensor calculus ... but just as we had to define vectors first, and then 1-forms and tensors, so we will do vector calculus first, and then form and tensor calculus.

A. Gradient of a scalar

It's easy to take the gradient of a scalar f. Since the 1-form $\boldsymbol{d}f$ was defined without reference to a metric, one may simply ignore the fact that space is "curved." One simply write down everything we said in flat spacetime: if we follow along any trajectory with velocity $\boldsymbol{v} = d\mathcal{P}/d\lambda$, then

$$\frac{df}{d\lambda} = \langle \boldsymbol{d}f, \boldsymbol{v} \rangle = \nabla_{\boldsymbol{v}} f, \tag{23}$$

or equivalently that

$$(\boldsymbol{d}f)_\alpha = \frac{\partial f}{\partial x^\alpha} = f_{,\alpha}. \tag{24}$$

B. Gradient of a vector

The gradient of a vector field $\boldsymbol{w}(\mathcal{P})$, however, presents a difficulty. In flat spacetime, we could write

$$\frac{d\boldsymbol{w}}{d\lambda} = \nabla_{\boldsymbol{v}} \boldsymbol{w}, \tag{25}$$

or in component language

$$\frac{dw^\beta}{d\lambda} = v^\alpha w^\beta{}_{,\alpha}. \tag{26}$$

The new problem in curved spacetime is that we can't subtract vectors that "live" in different places. So if we write

$$\frac{d\boldsymbol{w}}{d\lambda} \stackrel{?}{=} \lim_{\epsilon\to 0} \frac{\boldsymbol{w}[\mathcal{P}(\lambda+\epsilon)] - \boldsymbol{w}[\mathcal{P}(\lambda)]}{\epsilon}, \tag{27}$$

the right-hand side is now meaningless! One could of course subtract the components of $\boldsymbol{w}$, but since the basis varies from place to place (see the polar coordinate case where $\boldsymbol{e}_r$ and $\boldsymbol{e}_\theta$ were inconstant) this is not a coordinate-independent exercise.

There are several equivalent ways to solve this problem and define a *covariant derivative* $\nabla_{\boldsymbol{v}}\boldsymbol{w}$. We will follow the line of argumentation in MTW §8.5, but with more (albeit still incomplete) mathematical rigor.

C. The local Lorentz (or Cartesian) coordinate system

It is not generally possible to define Lorentz (or Cartesian, depending on the metric signature) coordinates on curved manifolds. However, it is always possible to define a coordinate system that is locally close to Lorentz. Consider a map of the Pacific Ocean; it is highly distorted by the coordinate system used by the cartographer. Now consider a map of Pasadena; the streets can now be represented nearly undistorted on flat paper. That small regions of spacetime look flat is the entire basis for the utility of special relativity in a universe whose global configuration is unknown.

[Note: MTW states the existence of such a frame but the text does not give an elementary proof.]

So let us try to formalize this. Given any point $\mathcal{P}$ with original (unprimed) coordinates $x^\mu(\mathcal{P})$, is it possible to build a new (primed) coordinate system centered at $\mathcal{P}$ that looks flat near $\mathcal{P}$? Clearly we can take the origin of said coordinate system to be at $\mathcal{P}$: $x^{\alpha'}(\mathcal{P}) = 0$. Let's then construct the old coordinates in terms of the new using a Taylor expansion:

$$x^\mu(x^{\alpha'}) = x^\mu(\mathcal{P}) + A^\mu{}_{\alpha'} x^{\alpha'} + \frac{1}{2} B^\mu{}_{\alpha'\beta'} x^{\alpha'} x^{\beta'} + \dots\,, \tag{28}$$

where $A^\mu{}_{\alpha'}$ and $B^\mu{}_{\alpha'\beta'}$ denote coefficients (and B is symmetric, $B^\mu{}_{\alpha'\beta'} = B^\mu{}_{\beta'\alpha'}$). The matrix relating the old and new coordinate basis vectors is

$$L^\mu{}_{\alpha'} = \frac{\partial x^\mu}{\partial x^{\alpha'}} = A^\mu{}_{\alpha'} + B^\mu{}_{\alpha'\beta'} x^{\beta'} + \dots\,. \tag{29}$$

Warning: **A** and **B** are coefficients in a Taylor expansion. They are **not** tensors because they only transform as such under global linear coordinate transformations (hence the index locations) – but under general coordinate transformations they have no special behavior.

Now let's consider the metric tensor in the new coordinate system. In the old system we can Taylor expand it as

$$g_{\mu\nu} = g_{\mu\nu}(\mathcal{P}) + g_{\mu\nu,\sigma}(\mathcal{P})[x^\sigma - x^\sigma(\mathcal{P})] + \dots \,. \tag{30}$$

In the new coordinate system, it is (sorry for the messy algebra):

$$\begin{aligned}
g_{\alpha'\beta'} &= L^\mu{}_{\alpha'} L^\nu{}_{\beta'} g_{\mu\nu} \\
&= [A^\mu{}_{\alpha'} + B^\mu{}_{\alpha'\gamma'} x^{\gamma'}][A^\nu{}_{\beta'} + B^\nu{}_{\beta'\delta'} x^{\delta'}]\{g_{\mu\nu}(\mathcal{P}) + g_{\mu\nu,\sigma}(\mathcal{P})[x^\sigma - x^\sigma(\mathcal{P})]\} + \mathcal{O}(x'^2) \\
&= [A^\mu{}_{\alpha'} + B^\mu{}_{\alpha'\gamma'} x^{\gamma'}][A^\nu{}_{\beta'} + B^\nu{}_{\beta'\delta'} x^{\delta'}][g_{\mu\nu}(\mathcal{P}) + g_{\mu\nu,\sigma}(\mathcal{P}) A^\sigma{}_{\epsilon'} x^{\epsilon'}] + \mathcal{O}(x'^2) \\
&= A^\mu{}_{\alpha'} A^\nu{}_{\beta'} g_{\mu\nu}(\mathcal{P}) + [B^\mu{}_{\alpha'\epsilon'} A^\nu{}_{\beta'} g_{\mu\nu}(\mathcal{P}) + A^\mu{}_{\alpha'} B^\nu{}_{\beta'\epsilon'} g_{\mu\nu}(\mathcal{P}) + A^\mu{}_{\alpha'} A^\nu{}_{\beta'} A^\sigma{}_{\epsilon'} g_{\mu\nu,\sigma}(\mathcal{P})] x^{\epsilon'} + \mathcal{O}(x'^2).
\end{aligned} \tag{31}$$

Since the metric has Lorentz signature, it is always possible to choose a nonsingular $A^\mu{}_{\alpha'}$ to set

$$g_{\alpha'\beta'}(0) = A^\mu{}_{\alpha'} A^\nu{}_{\beta'} g_{\mu\nu}(\mathcal{P}) = \eta_{\alpha'\beta'}. \tag{32}$$

(This is a change of basis, as proved on HW#1.) In fact, there are many choices of $A^\mu{}_{\alpha'}$, with 6 free parameters characterizing the Lorentz transformation. In Euclidean signature spaces, one may set $g_{\alpha'\beta'}(0) = \delta_{\alpha'\beta'}$, with $\frac{1}{2}n(n-1)$ free parameters since one can rotate the basis set while preserving orthonormality of the coordinate basis.

So it is possible to choose a coordinate system that looks Lorentz-like at a given point. But we can do even better, and choose a new coordinate system that additionally has $g_{\alpha'\beta',\epsilon'} = 0$ at this single point. To see that this is indeed the case, define the symbol

$$H_{\alpha'\beta'\epsilon'} = A^\mu{}_{\alpha'} B^\nu{}_{\beta'\epsilon'} g_{\mu\nu}(\mathcal{P}) \quad \leftrightarrow \quad B^\nu{}_{\beta'\epsilon'} = [\mathbf{A}^{-1}]^{\alpha'}{}_\mu g^{\mu\nu}(\mathcal{P}) H_{\alpha'\beta'\epsilon'}. \tag{33}$$

Here $H_{\alpha'\beta'\epsilon'}$ is symmetric on the last two indices, and subject to this condition, any $H_{\alpha'\beta'\epsilon'}$ can be generated by some $B^\nu{}_{\beta'\epsilon'}$. Then using the last line of Eq. (31), we have

$$g_{\alpha'\beta',\epsilon'}(0) = H_{\beta'\alpha'\epsilon'} + H_{\alpha'\beta'\epsilon'} + I_{\alpha'\beta'\epsilon'}, \tag{34}$$

where

$$I_{\alpha'\beta'\epsilon'} = A^\mu{}_{\alpha'} A^\nu{}_{\beta'} A^\sigma{}_{\epsilon'} g_{\mu\nu,\sigma}(\mathcal{P}) \tag{35}$$

is symmetric on the *first* two indices. It turns out to be possible to set $g_{\alpha'\beta',\epsilon'}(0) = 0$ by choosing

$$H_{\alpha'\beta'\epsilon'} = \frac{1}{2}(I_{\epsilon'\beta'\alpha'} - I_{\alpha'\beta'\epsilon'} - I_{\epsilon'\alpha'\beta'}). \tag{36}$$

It also turns out (homework!) that once we have chosen $A^\mu{}_{\alpha'}$, then $H_{\alpha'\beta'\epsilon}$ is unique.

Putting the above equations together, we can derive a formula for $H_{\alpha'\beta'\epsilon'}$:

$$H_{\alpha'\beta'\epsilon'} = \frac{1}{2} A^\rho{}_{\alpha'} A^\sigma{}_{\beta'} A^\tau{}_{\epsilon'} (g_{\tau\sigma,\rho} - g_{\rho\sigma,\tau} - g_{\tau\rho,\sigma}). \tag{37}$$

It then follows that

$$B^\nu{}_{\beta'\epsilon'} = \frac{1}{2} g^{\rho\nu} A^\sigma{}_{\beta'} A^\tau{}_{\epsilon'} (g_{\tau\sigma,\rho} - g_{\rho\sigma,\tau} - g_{\tau\rho,\sigma}), \tag{38}$$

where everything on the right-hand side is evaluated at $\mathcal{P}$.

We will find that it is not in general possible to set the second derivatives of the metric to zero: indeed, we will find that the metric is the GR analogue of the gravitational potential, the first derivative to the gravitational field, and the second derivatives correspond to the tidal field. A freely falling observer can pretend that he/she is at zero potential, and not in the presence of a gravitational field, but the tidal field cannot be eliminated.

Therefore, we conclude that **given any point $\mathcal{P} \in \mathcal{M}$, there exists a coordinate system with $x^{\alpha'}(\mathcal{P}) = 0$, with metric $g_{\alpha'\beta'} = \eta_{\alpha'\beta'}$ at $\mathcal{P}$, and $g_{\alpha'\beta',\epsilon'} = 0$ at $\mathcal{P}$.** The coordinate system that does this is unique up to second order in x', aside from the 6 Lorentz transformation degrees of freedom.

D. Covariant derivative

The existence of a coordinate system in which spacetime looks flat to lowest order suggests that we could define the gradient of a vector at a point $\mathcal{P}$ by the rule:

$$(\nabla_{\boldsymbol{v}}\boldsymbol{w})^{\alpha'} = v^{\beta'}w^{\alpha'}{}_{,\beta'} \quad \text{or} \quad \nabla_{\beta'}w^{\alpha'} = w^{\alpha'}{}_{,\beta'}, \tag{39}$$

where the partial derivative is to be evaluated in a local Lorentz coordinate system. If one wants the components of $\boldsymbol{\nabla w}$ in the original (unprimed) coordinate basis, or in any other basis, one does a basis transformation.

To use this definition, one must do two things: (i) find an actual computational method to find the covariant derivative; and (ii) show that it is not dependent on the choice of local Lorentz coordinate system (i.e. on which of the many legal $A^\rho{}_{\alpha'}$ we choose). Our preference is to do (i) first, since then (ii) will be trivial.

To compute the covariant derivative of a vector field $\boldsymbol{w}$, let's imagine that we are given the original components w^μ everywhere. Then

$$\begin{aligned}
\nabla_{\beta'}w^{\alpha'}(\mathcal{P}) &= \frac{\partial w^{\alpha'}}{\partial x^{\beta'}}(\mathcal{P}) \\
&= \frac{\partial}{\partial x^{\beta'}}([\mathbf{L}^{-1}]^{\alpha'}{}_\mu w^\mu) \\
&= w^\mu \frac{\partial[\mathbf{L}^{-1}]^{\alpha'}{}_\mu}{\partial x^{\beta'}} + [\mathbf{L}^{-1}]^{\alpha'}{}_\mu \frac{\partial w^\mu}{\partial x^{\beta'}}.
\end{aligned} \tag{40}$$

The last step used the product rule. We now need the partial derivative of a matrix inverse. This we can obtain by some simple manipulations, again involving the product rule:

$$\begin{aligned}
\frac{\partial[\mathbf{L}^{-1}]^{\alpha'}{}_\mu}{\partial x^{\beta'}} &= [\mathbf{L}^{-1}]^{\alpha'}{}_\sigma L^\sigma{}_{\gamma'} \frac{\partial[\mathbf{L}^{-1}]^{\gamma'}{}_\mu}{\partial x^{\beta'}} \\
&= [\mathbf{L}^{-1}]^{\alpha'}{}_\sigma \left\{ \frac{\partial}{\partial x^{\beta'}}(L^\sigma{}_{\gamma'}[\mathbf{L}^{-1}]^{\gamma'}{}_\mu) - [\mathbf{L}^{-1}]^{\beta'}{}_\mu \frac{\partial L^\sigma{}_{\beta'}}{\partial x^{\beta'}} \right\} \\
&= -[\mathbf{L}^{-1}]^{\alpha'}{}_\sigma [\mathbf{L}^{-1}]^{\gamma'}{}_\mu \frac{\partial L^\sigma{}_{\gamma'}}{\partial x^{\beta'}}.
\end{aligned} \tag{41}$$

(You might consider this a matrix quotient rule.) In the last line, we have recalled that $L^\sigma{}_{\gamma'}[\mathbf{L}^{-1}]^{\gamma'}{}_\mu = \delta^\sigma_\mu$, so its partial derivative is zero.

Returning to Eq. (40), we have

$$\nabla_{\beta'}w^{\alpha'}(\mathcal{P}) = -w^\mu[\mathbf{L}^{-1}]^{\alpha'}{}_\sigma[\mathbf{L}^{-1}]^{\gamma'}{}_\mu \frac{\partial L^\sigma{}_{\gamma'}}{\partial x^{\beta'}} + [\mathbf{L}^{-1}]^{\alpha'}{}_\mu \frac{\partial w^\mu}{\partial x^{\beta'}}. \tag{42}$$

Using that $L^\mu{}_{\alpha'} = A^\mu{}_{\alpha'}$ and $L^\mu{}_{\alpha',\beta'} = B^\mu{}_{\alpha'\beta'}$ at $\mathcal{P}$, we can re-write this as

$$\nabla_{\beta'}w^{\alpha'}(\mathcal{P}) = -w^\mu[\mathbf{A}^{-1}]^{\alpha'}{}_\sigma[\mathbf{A}^{-1}]^{\gamma'}{}_\mu B^\sigma{}_{\gamma'\beta'} + [\mathbf{A}^{-1}]^{\alpha'}{}_\mu \frac{\partial w^\mu}{\partial x^{\beta'}}. \tag{43}$$

We may re-write the last term using the chain rule,

$$\nabla_{\beta'}w^{\alpha'}(\mathcal{P}) = -w^\mu[\mathbf{A}^{-1}]^{\alpha'}{}_\sigma[\mathbf{A}^{-1}]^{\gamma'}{}_\mu B^\sigma{}_{\gamma'\beta'} + [\mathbf{A}^{-1}]^{\alpha'}{}_\mu A^\kappa{}_{\beta'} w^\mu{}_{,\kappa}. \tag{44}$$

Then we substitute in our previous result for B:

$$\nabla_{\beta'}w^{\alpha'}(\mathcal{P}) = -\frac{1}{2}w^\mu[\mathbf{A}^{-1}]^{\alpha'}{}_\sigma[\mathbf{A}^{-1}]^{\gamma'}{}_\mu g^{\sigma\tau}A^\eta{}_{\gamma'}A^\pi{}_{\beta'}(g_{\pi\eta,\tau} - g_{\eta\tau,\pi} - g_{\pi\tau,\eta}) + [\mathbf{A}^{-1}]^{\alpha'}{}_\mu A^\kappa{}_{\beta'} w^\mu{}_{,\kappa} \tag{45}$$

and simplify the first term:

$$\nabla_{\beta'}w^{\alpha'}(\mathcal{P}) = -\frac{1}{2}w^\eta[\mathbf{A}^{-1}]^{\alpha'}{}_\sigma g^{\sigma\tau}A^\pi{}_{\beta'}(g_{\pi\eta,\tau} - g_{\eta\tau,\pi} - g_{\pi\tau,\eta}) + [\mathbf{A}^{-1}]^{\alpha'}{}_\mu A^\kappa{}_{\beta'} w^\mu{}_{,\kappa}. \tag{46}$$

Finally, we use the transformation matrix to take us back to the unprimed coordinates:

$$\begin{aligned}
\nabla_\delta w^\epsilon(\mathcal{P}) &= [\mathbf{A}^{-1}]^{\beta'}{}_\delta A^\epsilon{}_{\alpha'}\nabla_{\beta'}w^{\alpha'}(\mathcal{P}) \\
&= -\frac{1}{2}w^\eta[\mathbf{A}^{-1}]^{\beta'}{}_\delta A^\epsilon{}_{\alpha'}[\mathbf{A}^{-1}]^{\alpha'}{}_\sigma g^{\sigma\tau}A^\pi{}_{\beta'}(g_{\pi\eta,\tau} - g_{\eta\tau,\pi} - g_{\pi\tau,\eta}) + [\mathbf{A}^{-1}]^{\beta'}{}_\delta A^\epsilon{}_{\alpha'}[\mathbf{A}^{-1}]^{\alpha'}{}_\mu A^\kappa{}_{\beta'} w^\mu{}_{,\kappa} \\
&= -\frac{1}{2}w^\eta g^{\epsilon\tau}(g_{\delta\eta,\tau} - g_{\eta\tau,\delta} - g_{\delta\tau,\eta}) + w^\epsilon{}_{,\delta}.
\end{aligned} \tag{47}$$

Now **A** has dropped out, so this is independent of what local Lorentz coordinate system we chose. (As an aside, we never even used the signature ... so this equation is equally valid for Euclidean signature spaces with no modification.) Furthermore we have a computational tool. Mission accomplished!

E. Notation

Vector calculus using Eq. (47) directly is possible, but is hampered by carrying around the funny combinations of partial derivatives of the metric tensor. There are several pieces of notation that simplify this. First, we define the *connection coefficients* or *Christoffel symbols*

$$\Gamma^{\epsilon}{}_{\delta\eta} \equiv \frac{1}{2} g^{\epsilon\tau}(-g_{\delta\eta,\tau} + g_{\eta\tau,\delta} + g_{\delta\tau,\eta}). \tag{48}$$

These coefficients are **not** tensors, since they don't transform with factors of $\mathbf{L}$ and $\mathbf{L}^{-1}$ under general coordinate transformations (example: in Cartesian coordinates, these vanish in $\mathbb{R}^2$, but in polar coordinates $\Gamma^{r}{}_{\theta\theta} \neq 0$). However, it is convenient to put the indices in the given locations, both for notational convenience and since under global linear changes of the coordinates they do transform in the expected way. In contrast, $\nabla_\delta w^\epsilon$ **is** a tensor, since we constructed it to be the same in all coordinate systems. If Alice has one system and Bob has another, then to construct $\nabla_\delta w^\epsilon(\mathcal{P})$ they both build locally flat coordinates at $\mathcal{P}$, which differ only by an irrelevant Lorentz transformation and higher-order terms, compute equivalent answers, and convert back to their chosen coordinate systems. Therefore their results differ only by a coordinate change.

Note that the Christoffel symbols are symmetric on the last two indices. Therefore, the number of Christoffel symbols required is $\frac{1}{2}n^2(n+1)$. This is 1 in 1D; 6 in 2D; 18 in 3D; and 40 in 4D. Fortunately, in symmetrical spacetimes many of these are zero, but their computation is still a formidable if straightforward task.

It is also common to write the covariant derivative of a vector with a semicolon:

$$w^{\epsilon}{}_{;\delta} \equiv \nabla_\delta w^\epsilon. \tag{49}$$

The semicolon distinguishes the covariant derivative from the ordinary partial derivative (denoted by a comma).

We may write the relation for the covariant derivative of a vector in this notation:

$$w^{\epsilon}{}_{;\delta} = w^{\epsilon}{}_{,\delta} + \Gamma^{\epsilon}{}_{\delta\eta} w^\eta. \tag{50}$$

Finally, we may relate the Christoffel symbols to the covariant derivatives of the basis vectors. The coordinate basis vector $\boldsymbol{e}_\alpha$ has constant components (0,...,0,1,0,...,0). If we take its directional derivative with respect to $\boldsymbol{e}_\beta$, we get a vector with components

$$(\nabla_{\boldsymbol{e}_\beta} \boldsymbol{e}_\alpha)^\epsilon = (\boldsymbol{e}_\beta)^\delta (\boldsymbol{e}_\alpha)^{\epsilon}{}_{;\delta} = \Gamma^{\epsilon}{}_{\beta\alpha}. \tag{51}$$

Since a basis 1-form extracts a component, it follows that

$$\langle \boldsymbol{\omega}^\epsilon, \nabla_{\boldsymbol{e}_\beta} \boldsymbol{e}_\alpha \rangle = \Gamma^{\epsilon}{}_{\beta\alpha}. \tag{52}$$

Finally, note that the covariant derivative immediately allows us to compute the divergence of a vector field:

$$\boldsymbol{\nabla} \cdot \boldsymbol{v} = v^{\alpha}{}_{;\alpha} = v^{\alpha}{}_{,\alpha} + \Gamma^{\alpha}{}_{\beta\alpha} v^\beta. \tag{53}$$