# Lecture VIII: Algebra and calculus with curved coordinate systems

(Dated: September 13, 2019)

## I. OVERVIEW

This lecture covers the key material from Chapter 5. This will deal with curved coordinate systems, which may be either on curved spaces or on flat spaces (e.g., polar coordinates). We first review how algebra changes when going from one system to another, and then consider what changes when we try to do calculus.

## II. PREFACE: POLAR COORDINATES AND VECTORS

Our "most basic" example of a curved coordinate system is the polar coordinate system, $(r, \phi)$, in Euclidean $\mathbb{R}^2$. This relates to the usual Cartesian coordinates by

$$x = r\cos\phi \quad \text{and} \quad y = r\sin\phi. \tag{1}$$

When we talk about vectors or tensors in polar coordinates, a key change relative to Cartesian coordinates is that I have to tell you where a vector is located: the "$r$ component" of a vector corresponds to a different direction depending on where in the plane we are. In polar coordinates – which are of course a curved coordinate system on a flat space – the vector has a direction whose meaning is independent of where it is located, and it is only the components that depend on where I choose to place the vector. In a curved space, by contrast, a vector will simply live in a different vector space depending on where it is. Thus we take our first step toward curved spacetime: to say that a "vector" $\boldsymbol{v}$ is located at a point $\mathcal{P}$.

Let us first think about how we describe the components of a vector. In flat space, we talked about a "vector" as describing the derivative of a particle's position with respect to some parameter $\sigma$: $v^\mu = \mathrm{d}x^\mu/\mathrm{d}\sigma$. In an alternative coordinate system, the vector components transformed according to the Jacobian:

$$v^{\bar\mu} = J^{\bar\mu}{}_\nu v^\nu, \quad J^{\bar\mu}{}_\nu = \frac{\partial x^{\bar\mu}}{\partial x^\nu}. \tag{2}$$

This concept carries over perfectly well to curved coordinate systems, now the only difference is that the Jacobian can depend on position and can be an arbitrary matrix (not necessarily a rotation matrix or Lorentz transformation). So if the barred system is the Cartesian coordinate ($x^{\bar 1} = x$, $x^{\bar 2} = y$) and the unbarred system is polar ($x^1 = r$, $x^2 = \phi$) then

$$\mathbf{J} = \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix}, \tag{3}$$

and we have

$$v^{\bar 1} = \cos\phi\, v^1 - r\sin\phi\, v^2 \quad \text{and} \quad v^{\bar 2} = \sin\phi\, v^1 + r\cos\phi\, v^2. \tag{4}$$

(Usually we write $v^1 = v^r$ and $v^2 = v^\phi$, as it is easier to keep track this way.) The transformation rule of Eq. (2) could also be used in reverse, or could allow us to do any transformation of coordinates (as long as it is differentiable).

There is a set of *coordinate basis vectors*, $\{\boldsymbol{e}_r, \boldsymbol{e}_\phi\}$, which are the velocities of particles with $\dot r = 1$, $\dot\phi = 0$ and $\dot r = 0$, $\dot\phi = 1$ respectively. These don't necessarily form an orthonormal basis.

We can also write 1-forms, such as the gradient: if $f$ is a scalar, then the object $\boldsymbol{k} = \tilde{\boldsymbol{d}}f$ with components $k_\mu = f_{,\mu} = \partial f/\partial\mu$ transforms as

$$k_{\bar\mu} = \frac{\partial f}{\partial x^{\bar\mu}} = \frac{\partial x^\nu}{\partial x^{\bar\mu}}\frac{\partial f}{\partial x^\nu} = [\mathbf{J}^{-1}]^\nu{}_{\bar\mu} k_\nu. \tag{5}$$

So in this sense, the rules for transformations of tensors carry over to curved coordinates in a rather trivial way.

What does not carry over so trivially is the inner product. Recall that the inner product of two vectors is

$$\boldsymbol{u}\cdot\boldsymbol{v} = g_{\alpha\beta}u^\alpha v^\beta, \tag{6}$$

where $g_{\alpha\beta}$ is the Kronecker delta (for Euclidean space with Cartesian coordinates) or $\eta_{\alpha\beta}$ (in special relativity). We call $\mathbf{g}$ the *metric tensor*. If we are considering curved coordinates in flat space, then one should obtain $g_{\alpha\beta}$ by transforming from Cartesian coordinates:

$$g_{\alpha\beta} = J^{\bar{\mu}}{}_{\alpha} J^{\bar{\nu}}{}_{\beta} g_{\bar{\mu}\bar{\nu}}. \tag{7}$$

In this case, in matrix form:

$$\mathbf{g} = \mathbf{J}^{\mathrm{T}}\bar{\mathbf{g}}\mathbf{J} = \mathbf{J}^{\mathrm{T}}\mathbb{I}\mathbf{J} = \mathbf{J}^{\mathrm{T}}\mathbf{J} = \begin{pmatrix} \cos\phi & \sin\phi \\ -r\sin\phi & r\cos\phi \end{pmatrix} \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}. \tag{8}$$

In components: $g_{rr} = 1$, $g_{\phi\phi} = r^2$, and $g_{r\phi} = g_{\phi r} = 0$.

Note that the coordinate basis vectors satsify the usual rule

$$\boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\beta = g_{\alpha\beta}; \tag{9}$$

thus in the case above, $\boldsymbol{e}_\phi$ is not a unit vector (it has square norm $r^2$).

If we want the lowering of indices to be consistent, in the sense that lowering an index and then changing coordinate systems is the same as changing coordinate systems and then lowering an index, we must lower indices with $g_{\alpha\beta}$. Similarly, raising of indices must take place with $g^{\mu\nu}$, which is the matrix inverse of $g_{\alpha\beta}$: $g^{\mu\nu}g_{\nu\gamma} = \delta^{\mu}{}_{\gamma}$. We will always require $g_{\alpha\beta}$ to be symmetric and non-singular. It should be symmetric since dot products should be symmetric ($\boldsymbol{u}\cdot\boldsymbol{v} = \boldsymbol{v}\cdot\boldsymbol{u}$). If it were singular with an eigenvector corresponding to eigenvalue 0, i.e., $g_{\alpha\beta}v^\beta = 0$ for some non-zero $\boldsymbol{v}$, then we would have $\boldsymbol{u}\cdot\boldsymbol{v} = 0\ \forall\boldsymbol{u}$. This would ruin the ability to raise and lower indices, and also ruins a whole bunch of mathematical theorems based on dot products.

For the polar coordinate case, the inverse metric is

$$g^{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & 1/r^2 \end{pmatrix}. \tag{10}$$

Thus, in the case of polar coordinates, the raising and lowering follows the rules:

$$v_r = v^r, \quad v_\phi = r^2 v^\phi, \quad v^r = v_r, \quad \text{and} \quad v^\phi = \frac{1}{r^2}v_\phi. \tag{11}$$

As you can see, it becomes very important to be careful of what index is up and what is down.

The metric is often written instead using the *line element* $\mathrm{d}s^2 = g_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta$. In the case of polar coordinates, we write

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\,\mathrm{d}\phi^2. \tag{12}$$

Since $g_{\alpha\beta}$ is symmetric, the line element contains exactly the same information as the metric tensor.

The Levi-Civita tensor is still fully antisymmetric. However, it now obeys the transformation rule

$$\varepsilon_{\bar{1}\bar{2}\bar{3}\ldots\bar{D}} = (\det\mathbf{J})\varepsilon_{123\ldots D}, \tag{13}$$

and we have allowed transformations with $\det\mathbf{J} \neq 1$. We note that $\det\bar{\mathbf{g}} = (\det\mathbf{J})^2\det\mathbf{g}$. Thus it is consistent to choose the normalization

$$\varepsilon_{123\ldots D} = \sqrt{|\det\mathbf{g}|} \tag{14}$$

for a right-handed coordinate system. This has the advantage of giving the usual result in the original Cartesian frame (in the flat case).

## III. LOCAL ORTHONORMAL FRAMES

You have seen that in the case above, it was possible at any point $\mathcal{P}$ to choose an orthonormal set of basis vectors. If such a set of basis vectors exists, then we can write

$$\boldsymbol{e}_{\hat{\alpha}} \cdot \boldsymbol{e}_{\hat{\beta}} = \eta_{\hat{\alpha}\hat{\beta}} \quad (\text{or } \delta_{\hat{\alpha}\hat{\beta}} \text{ for Euclidean metrics}). \tag{15}$$

Clearly, when describing a flat space, we can use the original Cartesian basis vectors as the orthonormal set (in our example: $\boldsymbol{e}_{\hat{x}}$ and $\boldsymbol{e}_{\hat{y}}$). However, there are infinitely many possibilities, each related by a rotation (Euclidean) or Lorentz transformation (special relativity). In the case of polar coordinates, we may choose for example the orthonormal basis

$$\boldsymbol{e}_{\hat{r}} = \boldsymbol{e}_r \to (1,0) \quad \text{and} \quad \boldsymbol{e}_{\hat{\phi}} = \frac{1}{r}\boldsymbol{e}_\phi \to \left(0, \frac{1}{r}\right). \tag{16}$$

One can check by explicit computation that, e.g., $g_{\alpha\beta}(\boldsymbol{e}_{\hat{\phi}})^\alpha(\boldsymbol{e}_{\hat{\phi}})^\beta = r^2(1/r)(1/r) = 1$. Generally, if the metric tensor is diagonal in a given coordinate system, that means that the coordinate basis vectors are already orthogonal (any two distinct coordinate basis vectors have dot product zero); then we can normalize them by taking

$$\boldsymbol{e}_{\hat{\alpha}} = \frac{1}{\sqrt{g_{\alpha\alpha}}}\boldsymbol{e}_\alpha \quad (\text{no E}\Sigma\text{C}). \tag{17}$$

Many highly symmetrical spacetimes can be written with coordinates where $\mathbf{g}$ is diagonal. If a spacetime is symmetrical enough and/or in low enough dimension, then there may be many ways to do this. But most general spacetimes in 4 dimensions cannot.

Note that if $g_{\alpha\beta}$ is symmetric and non-singular, then we can find an orthonormal basis at any point $\mathcal{P}$ by diagonalizing $g_{\alpha\beta}(\mathcal{P})$ and then normalizing each eigenvector. The pattern of positive and negative eigenvalues, called the *signature* of the metric, cannot be changed in this process. This pattern is $++$ in locally Euclidean 2D manifolds, $+++$ in locally Euclidean 3D manifolds, and $-+++$ in relativity.

The local orthonormal frame is a very important concept in geometry. If each $g_{\alpha\beta}$ is a differentiable function of the coordinates, then in the case of the orthonormal frame described by $\{\boldsymbol{e}_{\hat{\alpha}}\}$, we may take any point $\mathcal{P}$ and define a hatted frame:

$$x^\mu = x^\mu(\mathcal{P}) + (\boldsymbol{e}_{\hat{\alpha}})^\mu x^{\bar{\alpha}}. \tag{18}$$

In this frame, $\mathcal{P}$ has coordinates $x^{\bar{\alpha}}(\mathcal{P}) = 0$, i.e., it is at the origin. Moreover, in the hatted frame, we have

$$g_{\bar{\alpha}\bar{\beta}}(0) = \frac{\partial x^\mu}{\partial x^{\bar{\alpha}}}\frac{\partial x^\nu}{\partial x^{\bar{\beta}}} g_{\mu\nu} = (\boldsymbol{e}_{\hat{\alpha}})^\mu(\boldsymbol{e}_{\hat{\beta}})^\nu g_{\mu\nu} = \boldsymbol{e}_{\hat{\alpha}} \cdot \boldsymbol{e}_{\hat{\beta}} = \eta_{\hat{\alpha}\hat{\beta}}. \tag{19}$$

Thus one can construct a local coordinate system around point $\mathcal{P}$ that looks like Minkowski geometry or Euclidean geometry. This is what we do when we make a map of a city on Earth. It is also what we do when we consider a local observer in a globally curved Universe (e.g., a student in the lab with a clock and rulers to define $txyz$, but who lives in the gravitational field of the Earth). Note that $g_{\bar{\alpha}\bar{\beta}}$ is equal to $\mathbb{I}$ or $\boldsymbol{\eta}$ at the origin ($\mathcal{P}$); as one moves away from the origin, this result may no longer hold. This means that the coordinate system is **only locally** Euclidean (or Minkowskian).

As a general note, in vector calculus you probably learned about the "$r$" and "$\phi$" components of a vector in polar coordinates. In most undergraduate texts, these components were the orthonormal frame components, $v^{\hat{r}}$ and $v^{\hat{\phi}}$. Note that in the orthonormal frame, indices are raised and lowered according to the flat space rules (i.e., spatial indices can be up or down with no consequence, and if present time indices can move up or down with a $-$ sign). For polar coordinates,

$$v_{\hat{r}} = v^{\hat{r}} = v^r \quad \text{and} \quad v_{\hat{\phi}} = v^{\hat{\phi}} = r v^\phi. \tag{20}$$

In what follows, we require the existence of a metric and/or orthonormal basis at each point (one implies the other). We will assume the existence of as many derivatives of $g_{\alpha\beta}$ as we need.

## IV. COVARIANT DERIVATIVES AND CHRISTOFFEL SYMBOLS

Now let's consider a path $x^\mu(\sigma)$ through a space or spacetime, with "velocity" vector $v^\mu = \mathrm{d}x^\mu/\mathrm{d}\sigma$. We know that we can take the directional derivative of a scalar $f$:

$$\boldsymbol{\nabla}_{\boldsymbol{v}} f = \lim_{\epsilon\to 0} \frac{f(x^\mu(\sigma+\epsilon)) - f(x^\mu(\sigma))}{\epsilon} = \frac{\mathrm{d}f(x^\mu(\sigma))}{\mathrm{d}\sigma} = \frac{\mathrm{d}x^\mu(\sigma)}{\mathrm{d}\sigma}\frac{\partial f}{\partial x^\mu} = v^\mu f_{,\mu}. \tag{21}$$

We would now like to be able to take the directional derivative of a vector field $\boldsymbol{u}$ instead of a scalar $f$. Unfortunately, we have a problem. In the scalar case, $f(x^\mu(\sigma))$ and $f(x^\mu(\sigma+\epsilon))$ live in the same space, $\mathbb{R}$ (or $\mathbb{C}$ if $f$ is complex).

But in the vector case, $\boldsymbol{u}(x^\mu(\sigma))$ is a vector at the point with coordinates $x^\mu(\sigma)$ and $\boldsymbol{u}(x^\mu(\sigma+\epsilon))$ is a vector at the point with coordinates $x^\mu(\sigma+\epsilon)$. So we need a way to compare vectors at neighboring points.

If we are just working with curved coordinates in flat spacetime, then of course one can compute the difference

$$\boldsymbol{u}(x^\mu(\sigma+\epsilon)) - \boldsymbol{u}(x^\mu(\sigma)) \tag{22}$$

by going to a global rectilinear frame; but we won't be able to do this in curved spacetime, so we must take a different approach. I will take an axiomatic approach here, by invoking postulates about the derivative of a vector until there is exactly one way to do it. As long as these postulates are independent of the choice of coordinate system, the resulting operation will be independent of the choice of coordinates, too.

We begin by postulating that the directional derivative $\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{u}$ at the point $\mathcal{P}$ is a function only of $\boldsymbol{v}$ at $\mathcal{P}$; it is linear in the components of $\boldsymbol{v}$ and contains no derivatives of $\boldsymbol{v}$. Furthermore, we require it to obey the usual vector calculus rules:

- Linearity in $\boldsymbol{u}$: $\boldsymbol{\nabla}_{\boldsymbol{v}}(\boldsymbol{u}+\boldsymbol{p}) = \boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{u} + \boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{p}$ and $\boldsymbol{\nabla}_{\boldsymbol{v}}(\lambda\boldsymbol{u}) = \lambda\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{u}$ for general vector fields $\boldsymbol{u}$ and $\boldsymbol{p}$, and for constant $\lambda$.
- The product rule for scalar multiplication: if $f$ is a scalar field and $\boldsymbol{u}$ is a vector field, then

$$\boldsymbol{\nabla}_{\boldsymbol{v}}(f\boldsymbol{u}) = f\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{u} + (\boldsymbol{\nabla}_{\boldsymbol{v}}f)\boldsymbol{u}. \tag{23}$$

- The product rule for dot products:

$$\boldsymbol{\nabla}_{\boldsymbol{v}}(\boldsymbol{u}\cdot\boldsymbol{p}) = (\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{u})\cdot\boldsymbol{p} + \boldsymbol{u}\cdot(\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{p}). \tag{24}$$

(It turns out this isn't sufficient to determine everything; we will therefore add one more condition later.)

The linearity and scalar multiplication rules allow us to write all directional derivatives in terms of those of the coordinate basis vectors. We see that

$$\boldsymbol{\nabla}_{\boldsymbol{v}}\boldsymbol{u} = v^\beta\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}(u^\alpha\boldsymbol{e}_\alpha) = v^\beta\left[(\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}u^\alpha)\boldsymbol{e}_\alpha + u^\alpha\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}\boldsymbol{e}_\alpha\right] = v^\beta\left(u^\mu{}_{,\beta} + \Gamma^\mu{}_{\alpha\beta}u^\alpha\right)\boldsymbol{e}_\mu, \tag{25}$$

where we have defined the *Christoffel symbol* $\Gamma^\mu{}_{\alpha\beta}$ by

$$\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}\boldsymbol{e}_\alpha = \Gamma^\mu{}_{\alpha\beta}\boldsymbol{e}_\mu. \tag{26}$$

These $D^3$ coefficients (in $D$ dimensions) contain the information necessary to take derivatives of vectors.

**Warning**: The Christoffel symbols are not true tensors, because if we change coordinate systems their transformation law depends not only on the Jacobian $\mathbf{J}$ but on higher derivatives. Nevertheless, we write them with up and down indices that are appropriate for EΣC, and we can raise and lower the indices using the metric to define Christoffel symbols with other indices in the up or down position.

The linearity and scalar multiplication rules, however, tell us nothing about what the Christoffel symbols actually equal. To determine their values, we look at the special case of Eq. (24) with all unit vectors (this contains all the information we need):

$$\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}(\boldsymbol{e}_\gamma\cdot\boldsymbol{e}_\delta) = (\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}\boldsymbol{e}_\gamma)\cdot\boldsymbol{e}_\delta + \boldsymbol{e}_\gamma\cdot(\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}\boldsymbol{e}_\delta) \tag{27}$$

or

$$g_{\gamma\delta,\beta} = \Gamma^\mu{}_{\gamma\beta}\boldsymbol{e}_\mu\cdot\boldsymbol{e}_\delta + \boldsymbol{e}_\gamma\cdot\Gamma^\mu{}_{\delta\beta}\boldsymbol{e}_\mu = \Gamma^\mu{}_{\gamma\beta}g_{\mu\delta} + \Gamma^\mu{}_{\delta\beta}g_{\mu\gamma} = \Gamma_{\delta\gamma\beta} + \Gamma_{\gamma\delta\beta}. \tag{28}$$

Equation (28) is $D^3$ equations for $D^3$ unknowns, so you might think it enables us to determine $\Gamma_{\delta\gamma\beta}$. Unfortunately, it does not: the right-hand side is automatically symmetric in $\gamma$ and $\delta$, so in fact there are only $D^2(D+1)/2$ independent equations. We will come back to this later, where we introduce another condition to make the solution unique.

The vector directional derivative is important enough that its components get a special name. We define the *covariant derivative* of a vector to be

$$\nabla_\beta u^\mu \equiv u^\mu{}_{;\beta} = u^\mu{}_{,\beta} + \Gamma^\mu{}_{\alpha\beta}u^\alpha. \tag{29}$$

This will turn out to be a second rank tensor. The semicolon is introduced to distinguish it from a partial derivative (with a comma). We will use either the $\nabla$ or semicolon notation as appropriate. Note that $\nabla_\beta$ denotes a covariant derivative, whereas $\partial_\beta$ denotes a partial derivative.

## V. COVARIANT DERIVATIVES OF 1-FORMS AND TENSORS

Just as we had directional and covariant derivatives of vectors, we should be able to do the same for 1-forms. Since there is a vector-to-1-form mapping, we have nothing new. We may write

$$\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}(\tilde{\boldsymbol{\omega}}^\delta \cdot \boldsymbol{e}_\alpha) = 0, \tag{30}$$

so – by applying the product rule to the left-hand side and rearranging –

$$(\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}\tilde{\boldsymbol{\omega}}^\delta) \cdot \boldsymbol{e}_\alpha = -\tilde{\boldsymbol{\omega}}^\delta \cdot (\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}\boldsymbol{e}_\alpha) = -\tilde{\boldsymbol{\omega}}^\delta \cdot \Gamma^\mu{}_{\alpha\beta}\boldsymbol{e}_\mu = -\Gamma^\delta{}_{\alpha\beta}. \tag{31}$$

Thus

$$\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}\tilde{\boldsymbol{\omega}}^\delta = -\Gamma^\delta{}_{\alpha\beta}\tilde{\boldsymbol{\omega}}_\alpha. \tag{32}$$

Writing a general 1-form $\boldsymbol{p}$ as $p_\delta\tilde{\boldsymbol{\omega}}^\delta$, we have

$$\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}\boldsymbol{p} = p_{\delta,\beta}\tilde{\boldsymbol{\omega}}^\delta + p_\delta\boldsymbol{\nabla}_{\boldsymbol{e}_\beta}\tilde{\boldsymbol{\omega}}^\delta = (p_{\alpha,\beta} - \Gamma^\delta{}_{\alpha\beta}p_\delta)\tilde{\boldsymbol{\omega}}_\alpha. \tag{33}$$

In general, we write the components as

$$p_{\alpha;\beta} = p_{\alpha,\beta} - \Gamma^\delta{}_{\alpha\beta}p_\delta. \tag{34}$$

A rank $\binom{m}{n}$ tensor can be written with basis vectors and 1-forms multiplied together: $\boldsymbol{e}_{\alpha_1} \otimes ...\boldsymbol{e}_{\alpha_m} \otimes \tilde{\boldsymbol{\omega}}^{\mu_1} \otimes ... \otimes \tilde{\boldsymbol{\omega}}^{\mu_n}$. The derivative of this, if we define it with the product rule, gets us $m+n$ terms, each with a Christoffel symbol. The result for a rank $\binom{2}{2}$ tensor is:

$$T^{\alpha_1\alpha_2}{}_{\mu_1\mu_2;\beta} = T^{\alpha_1\alpha_2}{}_{\mu_1\mu_2,\beta} + \Gamma^{\alpha_1}{}_{\delta\beta}T^{\delta\alpha_2}{}_{\mu_1\mu_2} + \Gamma^{\alpha_2}{}_{\delta\beta}T^{\alpha_1\delta}{}_{\mu_1\mu_2} - \Gamma^\delta{}_{\mu_1\beta}T^{\alpha_1\alpha_2}{}_{\delta\mu_2} - \Gamma^\delta{}_{\mu_2\beta}T^{\alpha_1\alpha_2}{}_{\mu_1\delta}. \tag{35}$$

That is, the covariant derivative of a tensor has a partial derivative, and then a sequence of correction terms for the fact that the basis vector or 1-form corresponding to each index is changing as one moves in direction $\beta$. For each correction term, the index being corrected moves onto the Christoffel symbol; the other indices are spectators. The Christoffel symbols have an index at the end indicating the direction (here $\beta$) and a dummy index (here $\delta$) that shows how the basis vector being corrected rotates into $\boldsymbol{e}_\delta$ or $\tilde{\boldsymbol{\omega}}_\delta$. Note that "up" and "down" indices are corrected with opposite signs.

A scalar has no indices, and its covariant derivative has no correction terms: $f_{;\beta} = f_{,\beta}$.

You will see that the up and down index pattern on the first two indices of the Christoffel symbol looks a lot like the rules for transformations of up- and down-indices when we do a coordinate transformation. This is not an accident: the Christoffel symbol describes the infinitesimal transformation of the basis vectors when we map from the space of vectors at point $\mathcal{P}$ to a neighboring point where the $x^\beta$ coordinate is slightly larger.

## VI. COMPUTING THE CHRISTOFFEL SYMBOLS

We now return to the problem of computing the Christoffel symbols. Recall we said we had not imposed enough postulates yet to determine them. We now want to invoke a new postulate to make the Christoffel symbols unique. We want to generalize the concept of an exterior derivative so as to generalize Stokes's theorem to curved spacetime. The obvious way to do this is to replace $\partial$ with $\nabla$ in the definition: for a $k$-form $\boldsymbol{F}$,

$$[\tilde{\boldsymbol{d}}\boldsymbol{F}]_{\alpha_1...\alpha_{k+1}} = (k+1)\nabla_{[\alpha_1}F_{\alpha_2...\alpha_{k+1}]}. \tag{36}$$

That looks easy enough, but some weird things can happen. If we take a 0-form $f$, then we find

$$[\tilde{\boldsymbol{d}}\tilde{\boldsymbol{d}}f]_{\alpha\beta} = \nabla_\alpha[\tilde{\boldsymbol{d}}f]_\beta - \nabla_\alpha[\tilde{\boldsymbol{d}}f]_\alpha = \nabla_\alpha\nabla_\beta f - \nabla_\beta\nabla_\alpha f = \nabla_\alpha\partial_\beta f - \nabla_\beta\partial_\alpha f = \partial_\alpha\partial_\beta f - \Gamma^\delta{}_{\beta\alpha}\partial_\delta f - \partial_\beta\partial_\alpha f + \Gamma^\delta{}_{\alpha\beta}\partial_\delta f. \tag{37}$$

This is only zero (as it needs to be for exterior derivative calculus to work) if $\Gamma^\delta{}_{\beta\alpha} = \Gamma^\delta{}_{\alpha\beta}$. We will postulate that $\Gamma$ be symmetric in the last two indices. This is called the *torsion-free postulate.* It has $D^2(D-1)/2$ independent equations (all cases where $\alpha \neq \beta$), and so with Eq. (28) this gives the correct number of equations for the $D^3$ components of $\Gamma$.

*Comment* — The symmetry of $\Gamma$, and the full antisymmetry of the exterior derivative, imply that all the correction terms cancel:

$$[\tilde{\boldsymbol{d}}\boldsymbol{F}]_{\alpha_1...\alpha_{k+1}} = (k+1)\nabla_{[\alpha_1}F_{\alpha_2...\alpha_{k+1}]} = (k+1)\partial_{[\alpha_1}F_{\alpha_2...\alpha_{k+1}]}. \tag{38}$$

The exterior derivative, with all indices down, works the same way in curved as in flat space. The generalized Stokes's theorem, etc., also work the same way if the area and volume indices are up.

At this point, we are finally ready to compute the Christoffel symbols. Recall Eq. (28):

$$g_{\gamma\delta,\beta} = \Gamma_{\delta\gamma\beta} + \Gamma_{\gamma\delta\beta}. \tag{39}$$

By permuting the indices three times, and subtracting the last permutation:

$$g_{\gamma\delta,\beta} + g_{\delta\beta,\gamma} - g_{\beta\gamma,\delta} = \Gamma_{\delta\gamma\beta} + \Gamma_{\gamma\delta\beta}\Gamma_{\beta\delta\gamma} + \Gamma_{\delta\beta\gamma} - \Gamma_{\gamma\beta\delta} - \Gamma_{\beta\gamma\delta} = 2\Gamma_{\delta\gamma\beta}, \tag{40}$$

where we have used symmetry on the last two indices. This implies

$$\Gamma_{\delta\gamma\beta} = \frac{1}{2}(g_{\gamma\delta,\beta} + g_{\delta\beta,\gamma} - g_{\beta\gamma,\delta}). \tag{41}$$

The "up" Christoffel symbols, which are what you usually need to use in practice, are obtained by raising an index. Equation (41) is one of the most important results in doing computations in GR.

The most trivial case is a rectilinear coordinate system in flat space, where $g_{\alpha\beta}$ is constant and the Christoffel symbols are all zero. We consider a non-trivial but tractable case next.

## VII. EXAMPLE: POLAR COORDINATES

Let's see how this all works in the case of polar coordinates. We have metric coefficients

$$g_{rr} = 1, \quad g_{r\phi} = g_{\phi r} = 0, \quad \text{and} \quad g_{\phi\phi} = r^2. \tag{42}$$

Of the $D^3 = 8$ possible derivatives of the metric coefficients with respect to the coordinates, we have only one that is not zero:

$$g_{\phi\phi,r} = 2r. \tag{43}$$

This leads, via Eq. (41), to

$$\Gamma_{r\phi\phi} = -r, \quad \Gamma_{\phi r\phi} = \Gamma_{\phi\phi r} = r, \quad \text{others zero.} \tag{44}$$

We can raise the indices with the inverse metric:

$$g^{rr} = 1, \quad g^{r\phi} = g^{\phi r} = 0, \quad \text{and} \quad g^{\phi\phi} = r^2. \tag{45}$$

This leads to the "up" Christoffel symbols:

$$\Gamma^r{}_{\phi\phi} = -r, \quad \Gamma^\phi{}_{r\phi} = \Gamma^\phi{}_{\phi r} = \frac{1}{r}, \quad \text{others zero.} \tag{46}$$

(In most cases, this will look more complicated.)

We can now work some examples in polar coordinates. For example, we can compute the direction derivatives of $\boldsymbol{e}_r$:

$$\nabla_{\boldsymbol{e}_r}\boldsymbol{e}_r = \Gamma^\mu{}_{rr}\boldsymbol{e}_\mu = 0 \quad \text{and} \quad \nabla_{\boldsymbol{e}_\phi}\boldsymbol{e}_r = \Gamma^\mu{}_{r\phi}\boldsymbol{e}_\mu = \frac{1}{r}\boldsymbol{e}_\phi. \tag{47}$$

You are familiar with the latter statement: if you are at $\phi = 0$ and move in the $\phi$-direction, the $\boldsymbol{e}_r$ vector rotates counterclockwise, i.e., up – that's in the $\phi$-direction. Similarly

$$\nabla_{\boldsymbol{e}_r}\boldsymbol{e}_\phi = \Gamma^\mu{}_{\phi r}\boldsymbol{e}_\mu = \frac{1}{r}\boldsymbol{e}_\phi \quad \text{and} \quad \nabla_{\boldsymbol{e}_\phi}\boldsymbol{e}_\phi = \Gamma^\mu{}_{\phi\phi}\boldsymbol{e}_\mu = -r\boldsymbol{e}_r. \tag{48}$$

We may take the Laplacian of a scalar:

$$\begin{aligned}
\nabla^2 f &= g^{\alpha\beta}\nabla_\alpha\nabla_\beta f \\
&= g^{\alpha\beta}\nabla_\alpha\partial_\beta f \\
&= g^{\alpha\beta}(\partial_\alpha\partial_\beta f - \Gamma^\mu{}_{\beta\alpha}\partial_\mu f) \\
&= g^{\alpha\beta}\partial_\alpha\partial_\beta f - g^{\alpha\beta}\Gamma^\mu{}_{\beta\alpha}\partial_\mu f \\
&= \partial_r^2 f + \frac{1}{r^2}\partial_\phi^2 f - \frac{1}{r^2}(-r)\partial_r f \\
&= \partial_r^2 f + \frac{1}{r^2}\partial_\phi^2 f + \frac{1}{r}\partial_r f.
\end{aligned} \tag{49}$$

(The only contribution from the Christoffel symbol term is $\mu = r$, $\alpha = \beta = \phi$.) You are probably familiar with this result from vector calculus, but it is nice to see.

Note that the relations may look different in the coordinate basis than in the orthonormal basis: recall $v^r = v^{\hat{r}}$ but $v^\phi = v^{\hat{\phi}}/r$. Thus for the divergence of a vector,

$$\nabla_\alpha v^\alpha = \partial_\alpha v^\alpha + \Gamma^\alpha{}_{\beta\alpha} v^\beta = \partial_r v^r + \partial_\phi v^\phi + \frac{1}{r} v^r = \partial_r v^{\hat{r}} + \frac{1}{r}\partial_\phi v^{\hat{\phi}} + \frac{1}{r} v^{\hat{r}}. \tag{50}$$

My usual recommendation is to do calculus in the coordinate basis and convert to an orthonormal basis at the end if you think that is a better way to present the results (or more intuitive in that particular case, or more closely related to what you measure in an experiment, or what the problem asks for). There is a way to do calculus with orthonormal basis vectors or with other vectors satisfying particular dot product rules. That formalism is useful – especially for dealing with spinors and for perturbation theory around black holes – but I don't have time to cover it in this class.