Vectors and tensors in curved space time

Asaf Pe'er[1]

April 23, 2020

This part of the course is based on Refs. [1], [2] and [3].

1. Introduction

Using the equivalence principle, we have studied the trajectories of free test particles in curved space time. We argued that from the point of view of the test particle, it is in free motion, and remains in free motion even in the presence of a gravitational field (think of astronauts orbiting earth... they are weightless !). The inclusion of gravity changes the curvature of space time. In the presence of gravity, space-time can no longer be considered "flat", but "curved". This curvature is what gives rise to what we call gravitational "acceleration", which is mathematically described by the affine connection.

Our ultimate goal is to understand **how** does the presence of gravitational field curves space-time. This will eventually lead us to **Einstein's field equation**. However, before we can get there, we still need to bridge a mathematical gap. We first need to understand how to describe mathematical quantities such as vectors and tensors, from which physical equations are derived, in curved space time.

2. The principle of general covariance

We want to understand how the laws of physics, beyond those governing freely-falling particles described by the geodesic equation, adapt to the curvature of space-time. The procedure essentially follows the paradigm established in arguing that free particles move along geodesics.

- First, we consider an equation that describes a law of physics in flat space-time, traditionally written in terms of partial derivatives and the flat metric.
- Second, according to the equivalence principle this equation will hold in the presence of gravity, provided that the equation is **generally covariant**, namely, it preserves its form under general coordinate transformation, $x \to x'$.

[1]Department of Physics, Bar-Ilan University

This is known as **the principle of general covariance**.

Let us first show that the principle of general covariance follows from the equivalence principle. Assume that we are in an arbitrary gravitational field, and consider any equation that satisfies both conditions. From the second condition, it follows that if the equation is true in one coordinate system, it is true in any other system, as it preserves its form. Next, the equivalence principle tells us that at any point we can construct a locally inertial system, in which the effects of gravity are absent. Finally, from the first criterion we know that the equation holds in this system, hence we conclude that it must hold in all other coordinate systems.

Another way of thinking at the principle of general covariance is that it is a consequence of the equivalence principle, plus the requirement that the laws of physics be independent of coordinates. (The requirement that laws of physics be independent of coordinates is essentially impossible to even imagine being untrue. Given some experiment, if one person uses one coordinate system to predict a result and another one uses a different coordinate system, they had better agree.)

The principle of general covariance manifests the importance of vectors and tensors introduced earlier: as these have simple transformation laws, equations composed of vectors and tensors can (relatively) easy be made invariant under general coordinate transformations.

The purpose of the rest of this chapter is therefore to generalize the notion of vectors and tensors introduced in flat space-time, so that we could use them in arbitrary curved space-time.

3. Generalization of the definition of vectors and tensors to curved space-time.

3.1. Vectors as directional derivatives

In special relativity, we emphasized the fact that vectors belong to the *tangent space*, composed of the set of all vectors at a single point in space-time. The crucial point was to emphasis the fact that vectors are objects **associated with a single point**. By doing so, we had to pay a price: we lost the sense of *direction*. We could not use a statement like “the vector points in the x direction” - this doesn’t make sense if the tangent space is merely an abstract vector space associated with each point !. Now it is time to take care of this problem.

3.1.1. Manifolds

Before we continue, since I will often use the term "manifold" to describe curved space time, I should briefly introduce it. For completeness, I provide a full mathematical definition in the appendix. I simply think that it is better at this stage to focus on physical/intuitive understanding of the concept, rather than give the precise mathematical definition, which can be somewhat confusing at first.

Very crudely speaking, without getting into the math, a **manifold** is an n-dimensional space that near each of its points resembles an n-dimensions Euclidean space. Thus, while locally it looks Euclidean, globally it is not - exactly what happens to space-time in the presence of gravity. Simple examples are n-dimensional sphere, torus ("bagel"), and Riemann surface of genus g which is a two-torus with g holes (see Figure 1).

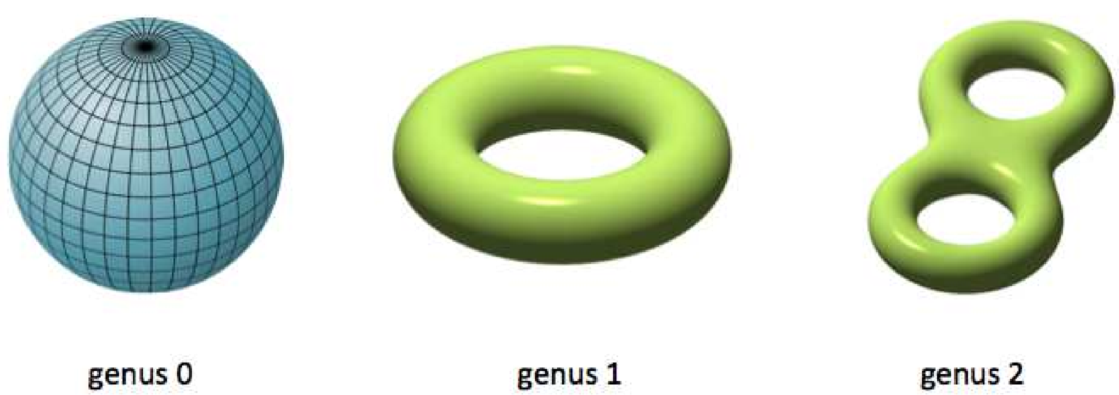

Fig. 1.— Riemann surfaces. A Riemann surface of genus 0 is a 2-dimensions sphere (also known as S^2), and Riemann surface of genus g is two-dimensional torus with g holes.

3.1.2. Vectors in curved spaces

We next want to put structures on manifolds. The first will be vectors and tangent spaces. How do we do that ?

We want to construct the tangent space at a point p in a curved space time (or in a manifold M), **using only things that are intrinsic to** M (no embedding in higher-dimensional spaces etc.). There is a little bit of mathematical subtlety here, so lets go over it carefully.

We know that the tangent space consists of all objects that are "tangent vectors to curves". Thus, we can draw a set of curves on the manifold that all pass through the point p. The temptation is to define the tangent space as simply the space of all tangent vectors to these curves at the point p.

The problem with this direct approach is that we still don't have an exact definition of what is exactly a "tangent vector to a curve" (at point p), apart from it being an element of the tangent space T_p - which we are currently trying to define... (see Figure 2, left).

Suppose we have a coordinate system x^μ which is defined on M in the vicinity of the point p (we can always do that - see appendix). In this coordinate system any curve through p defines an element of $\mathbf{R}^n$ specified by the n real numbers $dx^\mu/d\lambda$ (at the point p, where λ is the parameter along the curve). However, this map is clearly coordinate-dependent, which is not what we want (Figure 2, middle).

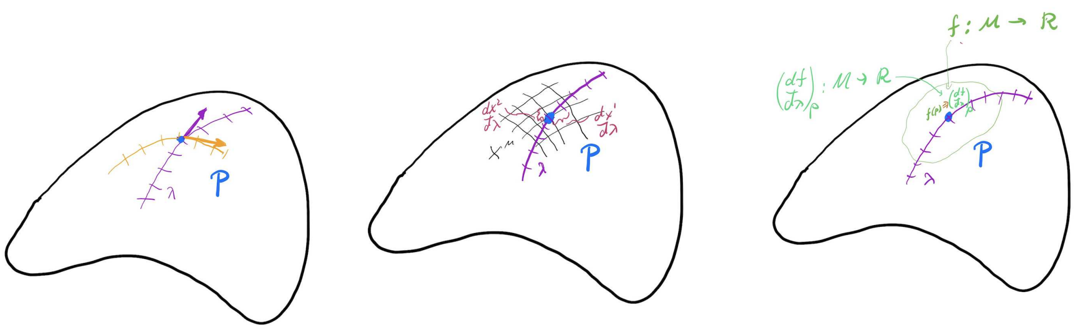

Fig. 2.— Constructing a tangent space at p. Left: first attemps - using tangent vectors to curves. But these are not defined yet. Middle: second attempt - in a given coordinate system, a curve defines a map from M to $\mathbf{R}^n$, by specifying $dx^\mu/d\lambda$ (at the point p); but this is coordinate dependent. Right: Coordinate independent way. Use the curve to define an operator which transforms any given function $f : M \to \mathbf{R}$ to $df/d\lambda$ at point p. This way specifies the tangent vector to the curve without referring to any given pre-set of coordinates.

We want to be more general that this: what we want is to use definitions which are **independent on the coordinates**. Here is how we do it (Figure 2, right). We define $\mathcal{F}$ to be the space of all smooth functions on M (in mathematical lingo, we say C^∞ maps $f : M \to \mathbf{R}$). Each curve through p defines *an operator on this space*: this is the directional derivative, which maps $f \to df/d\lambda$ (at p).[2] The curve, thus, defines an operation at the point p, in a coordinate-independent way, as we wanted.

We can now make the following claim: **the tangent space T_p can be identified with the space of directional derivative operators along curves through** p. To establish this idea we must demonstrate two things: (I) furst, we need to show that the space of directional derivatives is indeed a vector space; and (II) that it is the vector space we want (it has the same dimensionality as M, yields a natural idea of a vector pointing along a certain direction, and so on).

The first claim, that directional derivatives form a vector space, seems straightforward enough. Imagine two operators $\frac{d}{d\lambda}$ and $\frac{d}{d\eta}$ representing derivatives along two curves through p - $x^\mu(\lambda)$ and $x^\mu(\eta)$ (to distinguish between the curves I specify each point by its coordinates). These can be added and multiplied by real numbers, to obtain a new operator $a\frac{d}{d\lambda} + b\frac{d}{d\eta}$.

It is left to check that the space closes; *i.e.*, that the resulting operator is itself a derivative operator. A good derivative operator is one that acts linearly on functions, and obeys the conventional Leibniz (product) rule on products of functions [3]. Our new operator is manifestly linear, so we need to verify that it obeys the Leibniz rule. We have

$$\begin{aligned} \left(a\frac{d}{d\lambda} + b\frac{d}{d\eta}\right)(fg) &= af\frac{dg}{d\lambda} + ag\frac{df}{d\lambda} + bf\frac{dg}{d\eta} + bg\frac{df}{d\eta} \\ &= \left(a\frac{df}{d\lambda} + b\frac{df}{d\eta}\right)g + \left(a\frac{dg}{d\lambda} + b\frac{dg}{d\eta}\right)f \ . \end{aligned} \tag{1}$$

Here, $f, g \in \mathcal{F}$. The product rule is thus satisfied, which proves that the new operator is indeed a derivative operator. We can therefore conclude that the set of directional derivatives is indeed a vector space.

Now for the second point: is it the vector space that we can identify with the tangent space? Obviously, the answer is yes. The easiest way to become convinced is to find a basis for the space. Consider again a coordinate chart with coordinates x^μ.[4] Then there is an obvious set of n directional derivatives at p, namely the partial derivatives ∂_μ at p, which define directional derivatives along curves that keep all of the other coordinates constant

---

[2]First, note that I use λ to refer to the specific curve., while it was also used to parametrize the curve Second, note that the function f is arbitray; at every point on M it produces a number, specifically at the point p. So is the new function $df/d\lambda$, which is defined at p.

[3]Leibniz rule states that for two functions f and g, $(f \cdot g)' = f' \cdot g + f \cdot g'$.

[4]A coordinate chart, or coordinate system is a way of expressing the points of a small neighborhood of a point p on a manifold M, as coordinates in Euclidean space. Technically, this is a one to one mapping $\phi : U \to \mathbf{R}^n$ from an open set U in M to an open set in $\mathbf{R}^n$. See appendix A for more details.

(see Figure 3). [5]

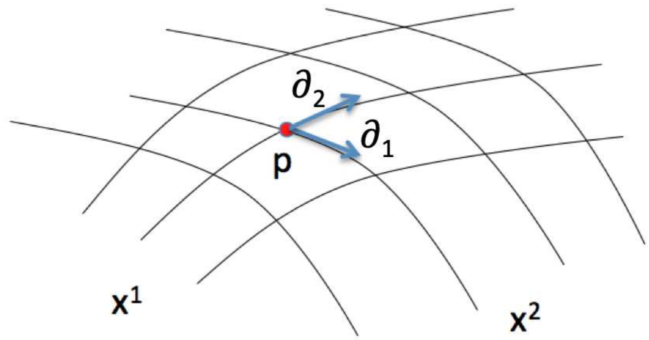

Fig. 3.— Coordinate chart on a (curved) manifold M provides a natural way to form the basis of the tangent space, by using the partial derivatives $\{\partial_\mu\}$ at p as a basis. These partial derivatives define directional derivatives along curves that keep all of the other coordinates constant.

We are now going to claim that the set of partial derivative operators $\{\partial_\mu\}$ at p form a basis for the tangent space T_p. (It follows immediately that T_p is n-dimensional, since that is the number of basis vectors.) To see this we need to show that any directional derivative can be decomposed into a sum of real numbers times partial derivatives. But this is in fact just the familiar expression for the components of a tangent vector. Consider a general n-dimensional manifold M, a curve $\gamma : \mathbf{R} \to M$, and a function $f : M \to \mathbf{R}$. If λ is the parameter along the curve γ, we can expand the vector/operator $\frac{d}{d\lambda}$ in terms of the partials ∂_μ

$$\frac{d}{d\lambda} f = \lim_{\epsilon \to 0} \frac{f(x^\mu(\lambda + \epsilon)) - f(x^\mu(\lambda))}{\epsilon} = \frac{dx^\mu}{d\lambda} \partial_\mu f \ . \tag{2}$$

Since the function f is arbitrary, we can write

$$\frac{d}{d\lambda} = \frac{dx^\mu}{d\lambda} \partial_\mu \ . \tag{3}$$

Thus, the partials $\{\partial_\mu\}$ do indeed represent a good basis for the vector space of directional derivatives, which we can therefore safely identify with the tangent space.

Of course, the vector represented by $\frac{d}{d\lambda}$ is one we already know; it is the tangent vector to the curve with parameter λ. Thus Equation 3 can be thought of as a restatement of

---

[5] Note that this is really the **definition** of the partial derivative with respect to x^μ: the directional derivative along a curve defined by $x^\nu = Const$, for all $\nu \neq \mu$. For this cure, we set as parameter $\lambda = x^\mu$.

Equation 28 in the SR chapter, where we claimed that the components of the tangent vector were simply $dx^\mu/d\lambda$. The only difference is that now we are working on an arbitrary manifold, and we have specified our basis vectors to be $\hat{e}_{(\mu)} = \partial_\mu$.

This particular basis ($\hat{e}_{(\mu)} = \partial_\mu$) is known as a **coordinate basis** for T_p; it is the formalization of the notion of setting up the basis vectors to point along the coordinate axes. There is no reason why we are limited to coordinate bases when we consider tangent vectors; it is sometimes more convenient, for example, to use orthonormal bases of some sort. However, the coordinate basis is very simple and natural, and we will use it almost exclusively throughout the course.

One of the advantages of the rather abstract point of view we have taken toward vectors is that the transformation law is immediate. Since the basis vectors are $\hat{e}_{(\mu)} = \partial_\mu$, the basis vectors in some new coordinate system $x^{\mu'}$ are given by the chain rule as

$$\partial_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \partial_\mu \ . \tag{4}$$

We can get the transformation law for vector components by the same technique used in flat space, demanding the vector $V = V^\mu \partial_\mu$ be unchanged by a change of basis. We have

$$\begin{aligned} V^\mu \partial_\mu &= V^{\mu'} \partial_{\mu'} \\ &= V^{\mu'} \tfrac{\partial x^\mu}{\partial x^{\mu'}} \partial_\mu \ , \end{aligned} \tag{5}$$

and hence (since the matrix $\partial x^{\mu'}/\partial x^\mu$ is the inverse of the matrix $\partial x^\mu/\partial x^{\mu'}$),

$$V^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} V^\mu \ . \tag{6}$$

Since the basis vectors are usually not written explicitly, the rule in Equation 6 for transforming components is what we call the **"transformation law of (contravariant) vector."** An object that transforms according to Equation 6 when the coordinates are changed from $x^\mu \to x^{\mu'}$ is a **contravariant vector**. Thus, **we identified vectors with directional derivatives.**

Of course, the transformation law in Equation 6 is compatible with the transformation of contravariant vector components in special relativity. Under Lorentz transformations, we had $V^{\mu'} = \Lambda^{\mu'}{}_\mu V^\mu$, but a Lorentz transformation is a special kind of coordinate transformation, with $x^{\mu'} = \Lambda^{\mu'}{}_\mu x^\mu$. Equation 6, though is much more general, as it encompasses the behavior of vectors under arbitrary changes of coordinates (and therefore bases), not just linear transformations. As such, it can be used in curved space time, not only a flat one.

To conclude, we are trying to emphasize a somewhat subtle ontological distinction — tensor components do not change when we change coordinates, but they change when we change the basis in the tangent space. Since we have decided to use the coordinates to define our basis, a change of coordinates induces a change of basis (see Figure 4), which, in turn, induces a change in the tensor components.

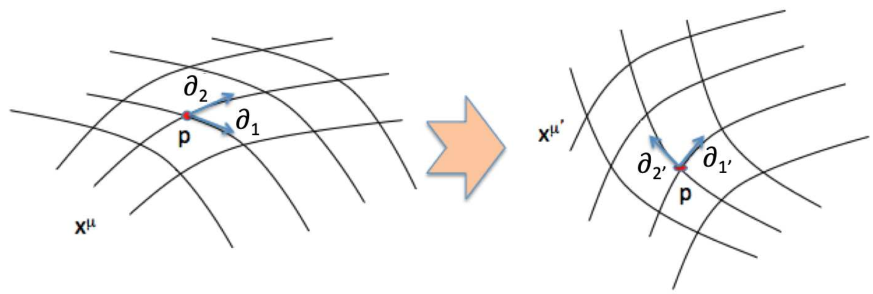

Fig. 4.— When we change the coordinate from x^μ to $x^{\mu'}$, we induce a change in the basis. This basis change leads to a change in the components of a tensor.

3.2. Covariant vectors

Equation 6 thus provides a general definition of a contravariant vector. We can now continue to follow the steps we took in flat space (SR), and consider the dual vectors (one forms). Once again the cotangent space T_p^* is the set of linear maps $\omega : T_p \to \mathbf{R}$. The canonical example of a one-form is the gradient of a function f, denoted $\mathrm{d}f$. Its action on a vector $\frac{d}{d\lambda}$ is exactly the directional derivative of the function:

$$\mathrm{d}f\left(\frac{d}{d\lambda}\right) = \frac{df}{d\lambda} \ . \tag{7}$$

Note the following: it's tempting to think, "why shouldn't the function f itself be considered the one-form, and $df/d\lambda$ its action?" The point is that a one-form, like a vector, exists only at the point it is defined, and does not depend on information at other points on M. If you know a function in some neighborhood of a point you can take its derivative, but not just from knowing its value at the point; the gradient, on the other hand, encodes precisely the information necessary to take the directional derivative along any curve through p, fulfilling its role as a dual vector.

Just as the partial derivatives along coordinate axes provide a natural basis for the tangent space, **the gradients of the coordinate functions x^μ provide a natural basis for the cotangent space**. Recall that in flat space we constructed a basis for T_p^* by demanding that $\hat{\theta}^{(\mu)}(\hat{e}_{(\nu)}) = \delta^\mu_\nu$. Continuing the same philosophy on an arbitrary manifold, we find that Equation 7 leads to

$$\mathrm{d}x^\mu(\partial_\nu) = \frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu \ . \tag{8}$$

Therefore the gradients $\{\mathrm{d}x^\mu\}$ are an appropriate set of basis one-forms; an arbitrary one-form is expanded into components as $\omega = \omega_\mu \,\mathrm{d}x^\mu$.

The transformation properties of basis dual vectors and components follow from what is by now the usual procedure. We obtain, for basis one-forms,

$$\mathrm{d}x^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu}\,\mathrm{d}x^\mu \ , \tag{9}$$

and for components,

$$\omega_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}}\omega_\mu \ . \tag{10}$$

We will usually write the components ω_μ when we speak about a one-form ω. Thus, equation 10 can be viewed as defining **the transformation law of the covariant vector (or one-form)** ω.

3.3. Tensors

The transformation law for general tensors follows this same pattern of replacing the Lorentz transformation matrix used in flat space with a matrix representing more general coordinate transformations. A (k, l) tensor T can be expanded

$$T = T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l}\partial_{\mu_1}\otimes\cdots\otimes\partial_{\mu_k}\otimes\mathrm{d}x^{\nu_1}\otimes\cdots\otimes\mathrm{d}x^{\nu_l} \ , \tag{11}$$

where I have used the symbol $\otimes$ to describe a **tensor product** (also known as **outer product**).[6]

Under a coordinate transformation the components change like the product of contravariant vectors and covariant vectors,

$$T^{\mu'_1\cdots\mu'_k}{}_{\nu'_1\cdots\nu'_l} = \frac{\partial x^{\mu'_1}}{\partial x^{\mu_1}}\cdots\frac{\partial x^{\mu'_k}}{\partial x^{\mu_k}}\frac{\partial x^{\nu_1}}{\partial x^{\nu'_1}}\cdots\frac{\partial x^{\nu_l}}{\partial x^{\nu'_l}}T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} \ . \tag{12}$$

---

[6]Without getting into precise mathematical definition, if T and S are tensors in the sense that each acts on a set of dual vectors and vectors, then $T\otimes S$ can be thought of as first act T on the appropriate set of dual vectors and vectors, then act S on the remainder, and finally multiply the answers. Note that, in general, $T\otimes S \neq S\otimes T$.

This tensor transformation law is straightforward to remember, since there really isn't anything else it could be, given the placement of indices. Equation 12 thus defines **the transformation law of tensors**.

3.3.1. Example.

Let us consider a symmetric $(0, 2)$ tensor S on a 2-dimensional curved space (manifold). Let us take as coordinate system on the manifold $(x^1 = x, x^2 = y)$. Let us assume that the components of the tensor are given by

$$S_{\mu\nu} = \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} . \tag{13}$$

This can be written equivalently as

$$S = S_{\mu\nu}(\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu) = x(\mathrm{d}x)^2 + (\mathrm{d}y)^2 , \tag{14}$$

Let us now change the coordinate system: consider new coordinates, say

$$\begin{aligned} x' &= x^{1/3} \\ y' &= e^{x+y} . \end{aligned} \tag{15}$$

This leads us to express the old basis $\mathrm{d}x^\mu$ in terms of the new basis, $\mathrm{d}x^{\mu'}$:

$$\begin{aligned} x &= (x')^3 \\ y &= \ln(y') - (x')^3 \\ \mathrm{d}x &= 3(x')^2\,\mathrm{d}x' \\ \mathrm{d}y &= \tfrac{1}{y'}\,\mathrm{d}y' - 3(x')^2\,\mathrm{d}x' . \end{aligned} \tag{16}$$

We need only plug these expressions directly into Equation 14 to write the components of S in terms of the new coordinates x', y'. (Remembering that tensor products don't commute, so $\mathrm{d}x'\,\mathrm{d}y' \neq \mathrm{d}y'\,\mathrm{d}x'$):

$$S = 9(x')^4[1 + (x')^3](\mathrm{d}x')^2 - 3\frac{(x')^2}{y'}(\mathrm{d}x'\,\mathrm{d}y' + \mathrm{d}y'\,\mathrm{d}x') + \frac{1}{(y')^2}(\mathrm{d}y')^2 , \tag{17}$$

or

$$S_{\mu'\nu'} = \begin{pmatrix} 9(x')^4[1 + (x')^3] & -3\frac{(x')^2}{y'} \\ -3\frac{(x')^2}{y'} & \frac{1}{(y')^2} \end{pmatrix} . \tag{18}$$

Notice that the tensor S is still symmetric. We did not use the transformation law (Equation 12) directly, but doing so would have yielded the same result, as you can check.

For the most part the various tensor operations we defined in flat space are unaltered in a more general setting: contraction, symmetrization, etc. There are three important exceptions: partial derivatives, the metric, and the Levi-Civita tensor. Let's look at the metric first.

4. Volume elements in curved space time and tensor densities

Clearly, the metric tensor,

$$g_{\mu\nu} \equiv \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\nu} \tag{19}$$

transforms as

$$g_{\mu'\nu'} = \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^{\mu'}} \frac{\partial \xi^\beta}{\partial x^{\nu'}} = \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial \xi^\beta}{\partial x^\nu} \frac{\partial x^\nu}{\partial x^{\nu'}} \tag{20}$$

or

$$g_{\mu'\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} g_{\mu\nu} \ . \tag{21}$$

Thus, we see that $g_{\mu\nu}$ is indeed a covariant tensor. Its inverse, $g^{\mu\nu}$ is a contravariant tensor. The Kronecker symbol, δ^μ_ν is a mixed tensor. However, not everything is tensor!. For example, the affine connection, $\Gamma^\lambda_{\mu\nu}$ is **not** a tensor (as we will see below).

One important example of a non-tensor quantity is the determinant of the metric tensor:

$$g \equiv -\mathrm{Det}\ (g_{\mu\nu}) = |g_{\mu\nu}| \tag{22}$$

Using the transformation rule of the metric tensor (Equation 21) and taking its determinant,

$$g(x^{\mu'}) = \left| \frac{\partial x^{\mu'}}{\partial x^\mu} \right|^{-2} g(x^\mu) \ . \tag{23}$$

which can be written as

$$g' = \left| \frac{\partial x'}{\partial x} \right|^{-2} g \tag{24}$$

Since $|\partial x'/\partial x|$ is the **Jacobian** of the transformation $x \to x'$. Thus, g transforms like a scalar, except for an extra factor of the Jacobian. It is thus called **scalar density** (which is a special case of **tensor density**). The number of factors of $|\partial x'/\partial x|$ is called the **weight** of the density; thus g is scalar density of weight -2.

The importance of the tensor density arise from the fact that under a general coordinate transformation $x \to x'$, the volume element $d^n x$ picks up a factor of the Jacobian,

$$d^n x' = \left| \frac{\partial x^{\mu'}}{\partial x^\mu} \right| d^n x \ . \tag{25}$$

Thus, $\sqrt{g}d^nx$ **is an invariant volume element.**

Example.

Consider the flat 3-d space written in spherical coordinates: $ds^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)$. We can thus write the metric as

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix} . \tag{26}$$

Thus, $g = r^4\sin^2\theta$, and a volume element is $\sqrt{g}drd\theta d\phi = r^2\sin\theta drd\theta d\phi$.

Notation.

We say that the metric is in **canonical form** if it is written in the form

$$g_{\mu\nu} = \mathrm{diag}(-1, -1, ..., -1, +1, +1, ..., +1, 0, 0, ..., 0).$$

If any of the eigenvalues is zero, the metric is "degenerate", and its inverse does not exist. However, we will always deal in this course with continuous, non-degenerate metrics. If all the signs are positive, the metric is called **Euclidean**, or **Riemannian**, while if there is a single minus, it is called **pseudo-Riemannian**, or **Lorentzian**. In GR, we deal with pseudo-Riemannian metric.

4.1. The Levi-Civita tensor density

The final change we have to make to our tensor knowledge now that we have dropped the assumption of flat space has to do with the Levi-Civita tensor, $\epsilon_{\mu_1\mu_2\cdots\mu_n}$. Remember that the flat-space version of this object, which we will now denote by $\tilde{\epsilon}_{\mu_1\mu_2\cdots\mu_n}$, was defined as

$$\tilde{\epsilon}_{\mu_1\mu_2\cdots\mu_n} = \begin{cases} +1 \text{ if } \mu_1\mu_2\cdots\mu_n \text{ is an even permutation of } 01\cdots(n-1) \ , \\ -1 \text{ if } \mu_1\mu_2\cdots\mu_n \text{ is an odd permutation of } 01\cdots(n-1) \ , \\ 0 \text{ otherwise} \ . \end{cases} \tag{27}$$

We will now define the **Levi-Civita symbol** to be exactly this $\tilde{\epsilon}_{\mu_1\mu_2\cdots\mu_n}$ — that is, an object with n indices which has the components specified above *in any coordinate system.* This is called a "symbol," of course, because it is **not** a tensor; it is defined not to change under coordinate transformations.

We can relate its behavior to that of an ordinary tensor by looking at the determinant of the matrix $\partial x^\mu/\partial x^{\mu'}$, which obeys

$$\tilde{\epsilon}_{\mu_1'\mu_2'\cdots\mu_n'} = \left|\frac{\partial x^{\mu'}}{\partial x^\mu}\right| \tilde{\epsilon}_{\mu_1\mu_2\cdots\mu_n} \frac{\partial x^{\mu_1}}{\partial x^{\mu_1'}}\frac{\partial x^{\mu_2}}{\partial x^{\mu_2'}}\cdots\frac{\partial x^{\mu_n}}{\partial x^{\mu_n'}} \ . \tag{28}$$

(This can be found in any linear algebra book..). Thus, the Levi-Civita symbol is a tensor density of weight 1.

However, we prefer to work with tensors, rather than tensor densities. There is a simple way to convert a density into an honest tensor — multiply by $|g|^{w/2}$, where w is the weight of the density (the absolute value signs are there because $g < 0$ for Lorentzian metrics). The result will transform according to the tensor transformation law. Therefore, for example, we can define the Levi-Civita tensor as

$$\epsilon_{\mu_1\mu_2\cdots\mu_n} = \sqrt{|g|}\,\tilde{\epsilon}_{\mu_1\mu_2\cdots\mu_n} \ . \tag{29}$$

Since this is a real tensor, we can raise indices, etc. Sometimes people define a version of the Levi-Civita symbol with upper indices, $\tilde{\epsilon}^{\mu_1\mu_2\cdots\mu_n}$, whose components are numerically equal to the symbol with lower indices. This turns out to be a density of weight -1, and is related to the tensor with upper indices by

$$\epsilon^{\mu_1\mu_2\cdots\mu_n} = \ \mathrm{sgn}(g)\frac{1}{\sqrt{|g|}}\,\tilde{\epsilon}^{\mu_1\mu_2\cdots\mu_n} \ . \tag{30}$$

As an aside, (for those of you who like math) we can point out that, even with the factor of $\sqrt{|g|}$, the Levi-Civita tensor is in some sense not a true tensor, because on some manifolds it cannot be globally defined. Those manifolds on which it can be defined are called **orientable**, and we will deal exclusively with orientable manifolds in this course. An example of a non-orientable manifold is the Möbius strip; see Schutz's *Geometrical Methods in Mathematical Physics* (or a similar text) for a discussion.

5. Covariant derivatives

The unfortunate fact is that the partial derivative of a tensor is not, in general, a new tensor. For example, if we take the contravariant vector V^μ, whose transformation law is given by Equation 6,

$$V^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} V^\mu$$

and we differentiate with respect to $x^{\lambda'}$, we get

$$\frac{\partial V^{\mu'}}{\partial x^{\lambda'}} = \frac{\partial}{\partial x^{\lambda'}}\left(\frac{\partial x^{\mu'}}{\partial x^\mu}V^\mu\right) = \left(\frac{\partial x^{\mu'}}{\partial x^\mu}\right)\left(\frac{\partial x^\rho}{\partial x^{\lambda'}}\frac{\partial V^\mu}{\partial x^\rho}\right) + \left(\frac{\partial^2 x^{\mu'}}{\partial x^\mu \partial x^\rho}\frac{\partial x^\rho}{\partial x^{\lambda'}}\right)V^\mu. \tag{31}$$

The first term on the right hand side is what we would expect if $\partial V^\mu / \partial x^\lambda$ was a tensor. It is this second term that destroys the tensor behavior.

This is a very problematic result, as derivatives of tensors are obvious ingredients in physical equations. We somehow need to find a way to generalize equations such as $\partial_\mu T^{\mu\nu} = 0$ to curved space time. Thus, what we really look for, is an operator which reduces to the partial derivative in flat space with Cartesian coordinates, but transforms as a tensor on an arbitrary manifold. Such an operator is called **covariant derivative**.

In order to construct a covariant derivative, lets have a look first at the transformation law of the affine connection.

5.1. Transformation law of the affine connection

Recall the definition of the affine connection,

$$\Gamma^\lambda_{\mu\nu} \equiv \frac{\partial x^\lambda}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\nu}, \tag{32}$$

where $\xi^\alpha(x)$ is the locally inertial coordinate system. When changing coordinates from x^μ to $x^{\mu'}$, the affine connection transforms as

$$\begin{aligned} \Gamma^{\lambda'}_{\mu'\nu'} &\equiv \frac{\partial x^{\lambda'}}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x^{\mu'} \partial x^{\nu'}}, \\ &= \frac{\partial x^{\lambda'}}{\partial x^\rho} \frac{\partial x^\rho}{\partial \xi^\alpha} \frac{\partial}{\partial x^{\mu'}} \left(\frac{\partial x^\sigma}{\partial x^{\nu'}} \frac{\partial \xi^\alpha}{\partial x^\sigma} \right) \\ &= \frac{\partial x^{\lambda'}}{\partial x^\rho} \frac{\partial x^\rho}{\partial \xi^\alpha} \left(\frac{\partial x^\sigma}{\partial x^{\nu'}} \frac{\partial x^\tau}{\partial x^{\mu'}} \frac{\partial^2 \xi^\alpha}{\partial x^\tau \partial x^\sigma} + \frac{\partial^2 x^\sigma}{\partial x^{\nu'} \partial x^{\mu'}} \frac{\partial \xi^\alpha}{\partial x^\sigma} \right). \end{aligned} \tag{33}$$

Undoubtedly, lovely. Using again Equation 32, we can write this as

$$\Gamma^{\lambda'}_{\mu'\nu'} = \frac{\partial x^{\lambda'}}{\partial x^\rho} \frac{\partial x^\sigma}{\partial x^{\nu'}} \frac{\partial x^\tau}{\partial x^{\mu'}} \Gamma^\rho_{\tau\sigma} + \frac{\partial x^{\lambda'}}{\partial x^\rho} \frac{\partial^2 x^\rho}{\partial x^{\nu'} \partial x^{\mu'}} \tag{34}$$

Clearly, the first term is what we would get if the affine connection was a tensor. The second term is inhomogeneous, and makes it a non-tensor.

We will write this in a slightly different form, using the identity

$$\frac{\partial x^{\lambda'}}{\partial x^\rho} \frac{\partial x^\rho}{\partial x^{\nu'}} = \delta^{\lambda'}_{\nu'}, \tag{35}$$

and differentiating with respect to $x^{\mu'}$ to write

$$\frac{\partial x^{\lambda'}}{\partial x^\rho} \frac{\partial^2 x^\rho}{\partial x^{\mu'} \partial x^{\nu'}} = -\frac{\partial^2 x^{\lambda'}}{\partial x^\rho \partial x^\sigma} \frac{\partial x^\sigma}{\partial x^{\mu'}} \frac{\partial x^\rho}{\partial x^{\nu'}}. \tag{36}$$

The transformation of the affine connection, equation 34 thus becomes

$$\Gamma^{\lambda'}_{\mu'\nu'} = \frac{\partial x^{\lambda'}}{\partial x^\rho} \frac{\partial x^\sigma}{\partial x^{\nu'}} \frac{\partial x^\tau}{\partial x^{\mu'}} \Gamma^\rho_{\tau\sigma} - \frac{\partial x^\sigma}{\partial x^{\mu'}} \frac{\partial x^\rho}{\partial x^{\nu'}} \frac{\partial^2 x^{\lambda'}}{\partial x^\rho \partial x^\sigma} \tag{37}$$

5.2. Covariant derivatives of vectors and tensors

Although $\partial V^\mu/\partial x^\lambda$ is not a tensor, the results of equation 37 can be used to construct a tensor. This is done by looking at the transformation law of $\Gamma^{\lambda'}_{\mu'\nu'}V^{\nu'}$,

$$\begin{aligned}\Gamma^{\lambda'}_{\mu'\nu'}V^{\nu'} &= \left[\frac{\partial x^{\lambda'}}{\partial x^\rho}\frac{\partial x^\sigma}{\partial x^{\nu'}}\frac{\partial x^\tau}{\partial x^{\mu'}}\Gamma^\rho_{\tau\sigma} - \frac{\partial x^\sigma}{\partial x^{\mu'}}\frac{\partial x^\rho}{\partial x^{\nu'}}\frac{\partial^2 x^{\lambda'}}{\partial x^\rho \partial x^\sigma}\right]\frac{\partial x^{\nu'}}{\partial x^\kappa}V^\kappa \\ &= \frac{\partial x^{\lambda'}}{\partial x^\rho}\frac{\partial x^\tau}{\partial x^{\mu'}}\Gamma^\rho_{\tau\sigma}V^\sigma - \frac{\partial x^\sigma}{\partial x^{\mu'}}\frac{\partial^2 x^{\lambda'}}{\partial x^\rho \partial x^\sigma}V^\rho\end{aligned} \tag{38}$$

Adding this to Equation 31 (replacing the indices $\nu' \leftrightarrow \lambda'$), one gets

$$\frac{\partial V^{\mu'}}{\partial x^{\lambda'}} + \Gamma^{\mu'}_{\lambda'\nu'}V^{\nu'} = \frac{\partial x^{\mu'}}{\partial x^\mu}\frac{\partial x^\rho}{\partial x^{\lambda'}}\left(\frac{\partial V^\mu}{\partial x^\rho} + \Gamma^\mu_{\rho\sigma}V^\sigma\right) \tag{39}$$

which is basically what we wanted !. We are thus led to define a **covariant derivative**

$$\nabla_\lambda V^\mu \equiv V^\mu{}_{;\lambda} \equiv \frac{\partial V^\mu}{\partial x^\lambda} + \Gamma^\mu_{\lambda\sigma}V^\sigma. \tag{40}$$

Equation 39 tells us that $V^\mu{}_{;\lambda}$ is a tensor, since

$$V^{\mu'}{}_{;\lambda'} = \frac{\partial x^{\mu'}}{\partial x^\mu}\frac{\partial x^\rho}{\partial x^{\lambda'}}V^\mu{}_{;\rho} \tag{41}$$

In an identical way, we can define the covariant derivate of a covariant vector ω_μ. Recall the transformation law (Equation 10),

$$\omega_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}}\omega_\mu,$$

and differentiate with respect to $x^{\nu'}$, to get

$$\frac{\partial \omega_{\mu'}}{\partial x^{\nu'}} = \frac{\partial x^\rho}{\partial x^{\mu'}}\frac{\partial x^\sigma}{\partial x^{\nu'}}\frac{\partial \omega_\rho}{\partial x^\sigma} + \frac{\partial^2 x^\rho}{\partial x^{\mu'}\partial x^{\nu'}}\omega_\rho. \tag{42}$$

Using Equation 34, we have

$$\begin{aligned}\Gamma^{\lambda'}_{\mu'\nu'}\omega_{\lambda'} &= \left[\frac{\partial x^{\lambda'}}{\partial x^\rho}\frac{\partial x^\sigma}{\partial x^{\nu'}}\frac{\partial x^\tau}{\partial x^{\mu'}}\Gamma^\rho_{\tau\sigma} + \frac{\partial x^{\lambda'}}{\partial x^\rho}\frac{\partial^2 x^\rho}{\partial x^{\nu'}\partial x^{\mu'}}\right]\frac{\partial x^\kappa}{\partial x^{\lambda'}}\omega_\kappa \\ &= \frac{\partial x^\sigma}{\partial x^{\nu'}}\frac{\partial x^\tau}{\partial x^{\mu'}}\Gamma^\kappa_{\tau\sigma}\omega_\kappa + \frac{\partial^2 x^\kappa}{\partial x^{\nu'}\partial x^{\mu'}}\omega_\kappa\end{aligned} \tag{43}$$

Subtracting Equation 43 from 42, we cancel the inhomogeneous terms and obtain

$$\frac{\partial \omega_{\mu'}}{\partial x^{\nu'}} - \Gamma^{\lambda'}_{\mu'\nu'}\omega_{\lambda'} = \frac{\partial x^\rho}{\partial x^{\mu'}}\frac{\partial x^\sigma}{\partial x^{\nu'}}\left(\frac{\partial \omega_\rho}{\partial x^\sigma} - \Gamma^\kappa_{\mu\nu}\omega_\kappa\right) \tag{44}$$

Thus, we define the **covariant derivative of a covariant vector** by

$$\nabla_\lambda \omega_\mu \equiv \omega_{\mu;\lambda} \equiv \frac{\partial \omega_\mu}{\partial x^\lambda} - \Gamma^\sigma_{\lambda\mu}\omega_\sigma. \tag{45}$$

Clearly, from Equation 44, $\omega_{\mu;\lambda}$ is a tensor, since

$$\omega_{\mu';\lambda'} = \frac{\partial x^\rho}{\partial x^{\mu'}}\frac{\partial x^\sigma}{\partial x^{\lambda'}}\omega_{\rho;\sigma} \tag{46}$$

Generalizing these definitions to arbitrary tensors is now straightforward. For each upper index we introduce a term with a $+\Gamma$, and for each lower index a term with a $-\Gamma$. For example,

$$T^{\mu_1\mu_2}{}_{\nu_1;\sigma} = \frac{\partial T^{\mu_1\mu_2}{}_{\nu_1}}{\partial x^\sigma} + \Gamma^{\mu_1}_{\sigma\lambda}T^{\lambda\mu_2}{}_{\nu_1} + \Gamma^{\mu_2}_{\sigma\lambda}T^{\mu_1\lambda}{}_{\nu_1} - \Gamma^{\lambda}_{\sigma\nu_1}T^{\mu_1\mu_2}{}_{\lambda}. \tag{47}$$

Clearly, this is a tensor.

The idea of covariant derivative can be extended to tensor densities, but I will check whether it is absolutely needed before continuing in this direction.

6. Importance of covariant derivatives, and the derivative of the metric

Let us stop for a moment and see what we got. By introducing the concept of covariant derivatives, combined with the algebraic properties of tensors [linearity, external (direct) product and contraction ($T^{\mu\nu} \equiv T^\mu{}_\rho{}^{\nu\rho}$)], we were able to extend the concept of partial derivatives from flat space time to a curved one. Moreover, we did it without being dependent on the particular coordinate system used. In particular, we found that the covariant derivatives are:

1. Linear:
   $(\alpha T + \beta S)_{;\lambda} = \alpha T_{;\lambda} + \beta S_{;\lambda}$, where α and β are numbers, T and S are tensors.

2. Obey the Leibniz (product) rule:
   $\nabla(T \otimes S) = (\nabla T) \otimes S + T \otimes (\nabla S)$, or $(TS)_{;\lambda} = T_{;\lambda}S + TS_{;\lambda}$.

The covariant derivative of the metric tensor is 0. This can be understood "intuitively", as in the local inertial frame it vanishes, and being a tensor, if it is 0 in one coordinate system it is 0 in any coordinate system. This can also be seen directly,

$$g_{\mu\nu;\lambda} = \frac{\partial g_{\mu\nu}}{\partial x^\lambda} - \Gamma^\rho_{\lambda\mu}g_{\rho\nu} - \Gamma^\rho_{\lambda\nu}g_{\rho\mu}, \tag{48}$$

and using the definition of the affine connection. In a very similar way, $g^{\mu\nu}{}_{;\lambda} = 0$, and $\delta^{\mu}_{\nu;\lambda} = 0$. Equation 48 is known as the **metric compatibility**.

The importance of covariant derivative arise from two of its properties:

1. It converts tensors to other tensors.
2. It reduces to ordinary differentiation in the absence of gravitation, $\Gamma^{\lambda}_{\mu\nu} = 0$, namely in flat space-time and Cartesian coordinates.

These properties thus suggest an easy algorithm to asses the effects of gravitation on physical systems: **(1) Write the appropriate SR equation that hold in the absence of gravity; then (2) replace $\eta_{\mu\nu}$ with $g_{\mu\nu}$ and all derivatives with covariant derivatives**. The resulting equations will be generally covariant, and true in the absence of gravitation. According to the principle of general covariance, they will be true in the presence of gravitational field (provided that we work in sufficiently small region of space).

7. Geometric interpretation of covariant derivatives

Let us have another look at covariant derivatives, as these are crucial when working in curved space-time. Let us look first at flat space-time. When we want to take a derivative of a vector, we consider two vectors $V(x^{\alpha})$ and $V(x^{\alpha} + dx^{\alpha})$ separated by an infinitesimal displacement dx^{α} along the direction of the derivative. Thus, to construct the derivative, we *first* transport the vector $V(x^{\alpha} + dx^{\alpha})$ parallel to itself back to the point x^{α}, to give the vector $V_{\|}(x^{\alpha})$. Only then it is in the tangent space of x^{α}, and then at a *second* step, we subtract the vector $V(x^{\alpha})$ from it, using the parallelogram rule (see Figure 5). The key thing that we do is **parallel transport**.

We now turn our attention to curved space time. We can perform parallel transport in curved space time, because locally we have a local inertial frame which is equivalent to flat space time. However, when we do that, **the coordinates of a vector change**. This results from the *change in the angle the vector make with the basis vectors*. This is demonstrated in Figure 6. This change is linear in the vector components. We therefore expect a term of the form $\nabla_{\beta} V^{\alpha} = \partial V^{\alpha}/\partial x^{\beta} + \overline{\Gamma}^{\alpha}_{\beta\gamma} V^{\gamma}$. Thus, while the first term $\partial V^{\alpha}/\partial x^{\beta}$ arise from a change in the vector field V between x^{α} and $x^{\alpha} + dx^{\alpha}$, the second term arise from a change in the basis vectors between the two points.

The question now, is why do $\overline{\Gamma}^{\alpha}_{\beta\gamma}$ turn out to be identical to the affine connection, $\Gamma^{\alpha}_{\beta\gamma}$ - surely this is no coincidence (?). Of course it isn't. We saw that the geodesic, in locally

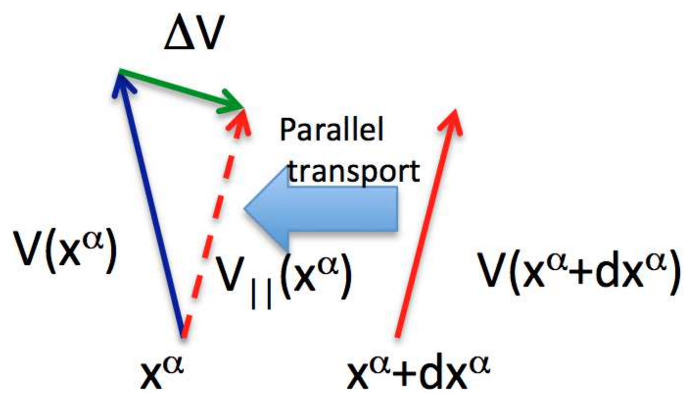

Fig. 5.— The derivative of a vector in flat space time includes to stages: First, the vector $V(x^\alpha + dx^\alpha)$ is *being transported parallel to itself* from $x^\alpha + dx^\alpha$ to x^α. The transported vector is subtracted from $V(x^\alpha)$ to obtain the difference $\Delta V(x^\alpha)$. The derivative is the difference $\Delta V(x^\alpha)/dx^\alpha$ in the limit $dx^\alpha \to 0$.

flat space time is a straight line. Now, we defined a "straight line" as the curve of extremal (minimal) distance between points. However, an alternative definition, is a curve whose unit tangent vector is parallel to itself (see Figure 7). Let us call this tangent vector u: then its covariant derivative in its own direction must vanish,

$$\nabla_u u^\alpha = u^\beta \left(\frac{\partial u^\alpha}{\partial x^\beta} + \overline{\Gamma}^\alpha_{\beta\gamma} u^\gamma \right) = 0, \tag{49}$$

where $u^\alpha = dx^\alpha/d\tau$. However, we already know that u fulfills the geodesic equation, which we can write as

$$u^\beta \left(\frac{\partial u^\alpha}{\partial x^\beta} + \Gamma^\alpha_{\beta\gamma} u^\gamma \right) = 0. \tag{50}$$

Comparing Equations 49 and 50 retrieves that indeed, $\overline{\Gamma}^\alpha_{\beta\gamma} = \Gamma^\alpha_{\beta\gamma}$, as we expected.

This argument can in fact be turned around, to give an elegant version of the geodesic equation in terms of the covariant derivative. A geodesic is a curve whose tangent vector u obeys

$$\nabla_u u = 0. \tag{51}$$

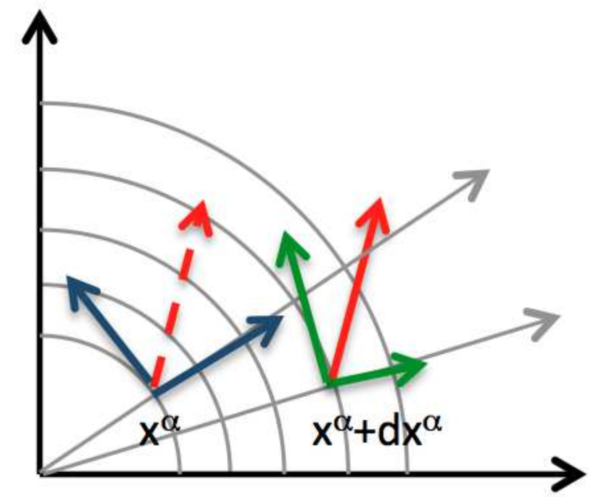

Fig. 6.— When parallel transporting a vector in non-Cartesian coordinates, the components of the vector change, due to change in the basis vectors: in this example, we use polar coordinates, and while the vector itself does not change when parallel-transported, its components do.

8. Gradient, divergence and curl

The equations of electromagnetism, fluid mechanics and many other areas of classical physics make use of the three-dimensional vector calculus employing functions such as gradient, divergence, curl and Laplacian. You have seen explicit forms of these functions in non-Cartesian coordinate systems, such as cylindrical or spherical. The concept of covariant derivative provides a unified picture of all these derivatives and a direct route to the explicit forms in given coordinate systems.

We have already seen that the covariant derivative of a scalar is just the ordinary gradient:

$$S_{;\mu} = \frac{\partial S}{\partial x^{\mu}} \tag{52}$$

Another special case is the covariant of the curl. Using Equation 45, $\omega_{\mu;\nu} = \partial\omega_{\mu}/\partial x^{\nu} - \Gamma^{\lambda}_{\mu\nu}\omega_{\lambda}$, and the fact that $\Gamma^{\lambda}_{\mu\nu}$ is symmetric in μ and ν, the covariant curl is just the ordinary curl,

$$\omega_{\mu;\nu} - \omega_{\nu;\mu} = \frac{\partial\omega_{\mu}}{\partial x^{\nu}} - \frac{\partial\omega_{\nu}}{\partial x^{\mu}}. \tag{53}$$

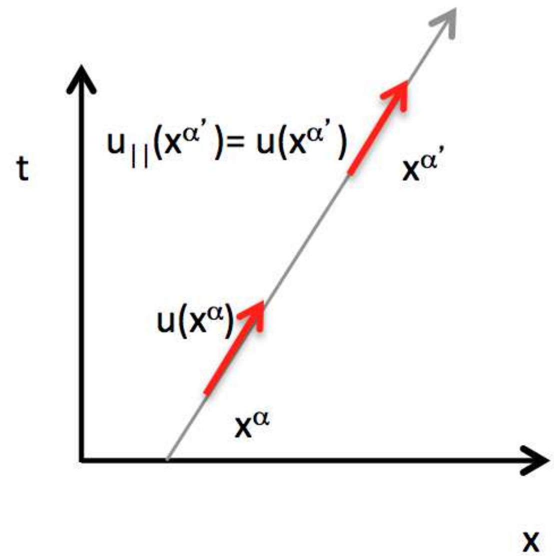

Fig. 7.— A geodesic can be thought of as a line for which a tangent vector V at x^α is parallel-transported to $x^{\alpha'}$, the obtained vector $V_\|(x^{\alpha'})$ coincides with the tangent vector $V(x^{\alpha'})$.

The covariant divergence of a contravariant vector is

$$V^\mu{}_{;\mu} = \frac{\partial V^\mu}{\partial x^\mu} + \Gamma^\mu_{\mu\lambda} V^\lambda. \tag{54}$$

We can use the symmetry properties of $\Gamma^\mu_{\mu\lambda}$ to write it as

$$\Gamma^\mu_{\mu\lambda} = \frac{1}{2} g^{\mu\rho} \left(\frac{\partial g_{\rho\mu}}{\partial x^\lambda} + \frac{\partial g_{\rho\lambda}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\rho} \right) = \frac{1}{2} g^{\mu\rho} \frac{\partial g_{\rho\mu}}{\partial x^\lambda} \tag{55}$$

Using the algebraic identity

$$Tr \left(M^{-1}(x) \frac{\partial}{\partial x^\lambda} M(x) \right) = \frac{\partial}{\partial x^\lambda} \ln \mathrm{Det}\,[M(x)] \tag{56}$$

and applying it to the matrix $M = g_{\rho\mu}$, leads to

$$\Gamma^\mu_{\mu\lambda} = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^\mu} \sqrt{|g|}, \tag{57}$$

and the covariant divergence is

$$V^\mu{}_{;\mu} = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^\mu} (\sqrt{|g|} V^\mu) \tag{58}$$

This immediately leads to the covariant form of Gauss's theorem: if V^μ vanishes at infinity, then

$$\int d^4x\sqrt{|g|}V^\mu{}_{;\mu} = 0. \tag{59}$$

(Note the appearance of $\sqrt{|g|}$ that makes $d^4x\sqrt{|g|}$ invariant).

In 3-dimensions, the Laplacian of a scalar S is just the divergence of its gradient, namely

$$\nabla^2 S = (g^{ij}S_{;i})_{;j} \tag{60}$$

which, using Equations 52 and 58 is

$$\nabla^2 S = \frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^j}\left(\sqrt{|g|}g^{ij}\frac{\partial S}{\partial x^i}\right). \tag{61}$$

Appendix A: Manifolds

Let us look at the exact (mathematical) definition of a manifold.

I stated that, very loosely speaking, a **manifold** is an n-dimensional space, that near each of its points resembles an n-dimensional Euclidean space; i.e., it is **locally** Euclidean, while **globally** it is not.

Since manifolds - or more precisely **differentiable manifolds** are fundamental concepts, I wish to bring here the exact, mathematical definition. For that we need a few other definitions. Informally, manifold is a space that consists of patches, that locally look like $\mathbf{R}^n$, and are smoothly sewn together. Let us define these two notions.

We begin with some formal definitions:

1. A **map** between two sets: given two sets, M and N, a map $\phi : M \rightarrow N$ is a relationship that assigns, to each element of M exactly one element of N. Basically, a map is a generalization of a function.

2. **Composition of maps**: Given two maps, $\phi : A \rightarrow B$ and $\psi : B \rightarrow C$, the composition $\psi \circ \phi : A \rightarrow C$ is defined by the operation $(\psi \circ \phi)(a) = \psi(\phi(a))$, where $a \in A$. Clearly, $(\psi \circ \phi)(a) \in C$.

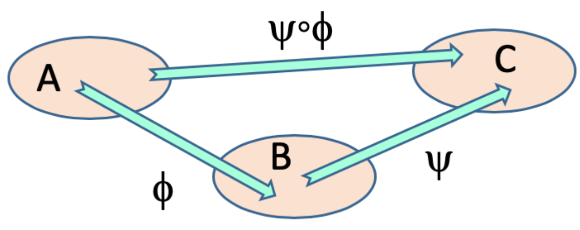

Fig. 8.— Composition of maps.

3. A map ϕ between two sets, M and N is called **one-to-one** if each element of N has at most one element of M mapped into it. It is **onto** if each element of N has at least one element of M mapped into it. For example, the map $\phi : \mathbf{R} \rightarrow \mathbf{R}$ defined such that $\phi(x) = e^x$ is one-to-one, but not onto. A map that is both one-to-one and onto is known as **invertible**.

4. The set M is known as the **domain** (Hebrew: Thum) of the map ϕ. The set of points in N that M gets mapped into is called the **image** of ϕ.

5. The **smoothness** of a function measures the number of derivatives it has that are continuous. Consider the map $\phi : \mathbf{R}^m \to \mathbf{R}^n$: that is, this map taken m-tuple $(x^1, ..., x^m)$ and maps it into n-tuple $(y^1, ..., y^n)$. We can think of it as a collection of n functions, ϕ^i, such that $y^i = \phi^i(x^1, ..., x^m)$, $(i = 1..n)$. If the pth-derivative of each of this functions exists and is continuous, we refer to that function as C^p. If this is the case for all functions, we refer to the map $\phi : \mathbf{R}^m \to \mathbf{R}^n$ as C^p. C^∞ maps are called **smooth**.

6. Two sets, M and N are called **diffeomorphic** if there exists a C^∞ map $\phi : M \to N$, with a C^∞ inverse $\phi^{-1} : N \to M$. The map ϕ is then called a **diffeomorphism**: basically it means that it is an invertible function that maps one differentiable manifold to another such that both the function and its inverse are smooth.

Let us now apply these general definitions to manifolds.

1. An **open ball** is the set of all points x in $\mathbf{R^n}$ such that $|x - y| < r$ for some fixed $y \in \mathbf{R}^n$ and $r \in \mathbf{R}$. Here, $|x - y| = \left[\sum_i (x^i - y^i)^2\right]^{1/2}$ (Note the inequality).

2. An **open set** in $\mathbf{R}^n$ is a set constructed from an arbitrary (possibly, infinite) union of open balls. Alternatively: $V \subset \mathbf{R}^n$ is open if, for any $y \in V$ there is an open ball centered at y, that is completely inside V.

3. A **chart** or **coordinate system** consists of a subset U of a set M, along with a one-to-one map $\phi : U \to \mathbf{R}^n$, such that the image $\phi(U)$ is open in $\mathbf{R}^n$. Since any map is onto its image, the map ϕ is invertible if it is one-to-one. Thus, U is an open set in M.

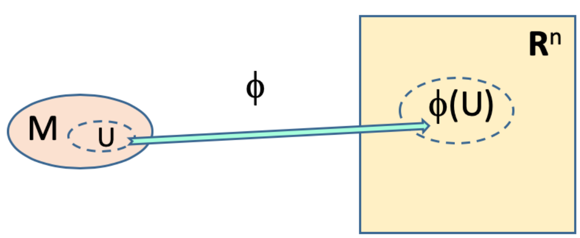

Fig. 9.— A coordinate chart maps an open subset U of M into an open subset of $\mathbf{R}^n$.

Think of it for a second, and convince yourself that a chart is what we normally think of as a coordinate system on some open set.

4. A C^∞ **atlas** is an indexed collection of charts $\{(U_\alpha, \phi_\alpha)\}$ that satisfies two conditions:

   (a) The union of U_α is equal to M. In other words, the U_α cover M.

   (b) The charts are smoothly sewn together in the following way: if two charts overlap, namely $U_\alpha \cap U_\beta \neq \oslash$, then the map $(\phi_\alpha \circ \phi_\beta^{-1})$ takes a point in $\mathbf{R}^n$: $\phi_\beta(U_\alpha \cap U_\beta) \in \mathbf{R}^n$ *onto* an open set $\phi_\alpha(U_\alpha \cap U_\beta) \in \mathbf{R}^n$. All the maps are C^∞ where they are defined.

   Thus, an atlas is a system of charts that are smoothly related on their overlaps.

5. Finally, a C^∞, n-dimensional **manifold** is a set M along with a maximal atlas - one that contains every possible compatible charts.

A very important point is that defined this way, a manifold does not need to be embedded in any higher-dimensional space; it is perfectly defined without it.

As an example, consider the 1-dimensional circle, S^1, and the coordinate system $\theta : S^1 \to \mathbf{R}$. However, since the image $\theta(S^1)$ must be open in $\mathbf{R}$, we must exclude the points $\theta = 0, 2\pi$. We thus need at least two charts to cover S^1.

Appendix B: Proof of Equation 56

Consider an arbitrary matrix M, whose components depend on x, thus $M = M(x)$. Let us look at the variation of ln(Det(M) that occurs when x^λ is changed by δx^λ.

$$\begin{aligned} \delta \ln \mathrm{Det} M &\equiv \ln \mathrm{Det}(M + \delta M) - \ln \mathrm{Det} M \\ &= \ln \left(\frac{\mathrm{Det}(M+\delta M)}{\mathrm{Det} M} \right) \\ &= \ln \mathrm{Det} \left[M^{-1}(M + \delta M) \right] \\ &= \ln \mathrm{Det} \left[1 + M^{-1} \delta M \right] \end{aligned} \quad (62)$$

where in the 3rd line we used the fact that for any two matrices M and N, Det MDet $N =$ Det (MN).

In order to proceed, we recall that if $\lambda_1 ... \lambda_n$ are the eigenvalues of the matrix M, then

$$\mathrm{Det}\ M = \Pi_i \lambda_i \quad (63)$$

Furthermore, the trace of the matrix is

$$\mathrm{Tr}\ M = \sum_i \lambda_i \quad (64)$$

Thus, for small change $\delta M \ll 1$, one gets $\mathrm{Det}(1+\delta M) = \Pi_i(1+\lambda_i) \approx 1 + \sum_i \lambda_i = 1 + \mathrm{Tr}\ \delta M$.

Overall, we get

$$\begin{aligned} \delta \ln \mathrm{Det} M &\approx \ln(1 + \mathrm{Tr}\ M^{-1} \delta M) \\ &\approx (\mathrm{Tr}\ M^{-1} \delta M) \end{aligned} \quad (65)$$

where in the last line we Taylor expanded $\ln(1 + x) \approx x$, assuming $x \ll 1$. Equation 56 immediately follows.

REFERENCES

[1] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (John Wiley & Sons), chapter 4.

[2] S. Carroll, *Lecture Notes on General Relativity*, part 2. Manifolds (http://preposterousuniverse.com/grnotes/).

[3] J. Hartle, *Gravity: An Introduction to Einstein's General Relativity* (Addison-Wesley), chapter 7 and chapter 20.