Another chart (U_2, ϕ_2) can be obtained by projecting from the south pole to the plane $x^3 = +1$. The resulting coordinates cover the sphere minus the south pole, and are given by:

$$\phi_2(x^1, x^2, x^3) \equiv (z^1, z^2) = \left(\frac{2\,x^1}{1+x^3}, \frac{2\,x^2}{1+x^3} \right) .$$

Together, these two charts cover the entire manifold, and they overlap in the region $-1 < x^3 < 1$. One can check that the composition $\phi_2 \circ \phi_1^{-1}$ is given by

$$z^i = \frac{4\,y^i}{\sqrt{(y^1)^2 + (y^2)^2}} , \quad i = 1, 2 ,$$

which is C^∞ in the overlap region.

Consider the maps $f : \mathbb{R}^m \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^l$, and the composition map $(g \circ f) : \mathbb{R}^m \to \mathbb{R}^l$. We can label points on each space in terms of the usual Cartesian coordinates: x^a on $\mathbb{R}^m$, y^b on $\mathbb{R}^n$ and z^c on $\mathbb{R}^l$, where the indices range over the appropriate values. The chain rule relates the partial derivatives of the composition to the partial derivatives of the individual maps:

$$\frac{\partial}{\partial x^a}(g \circ f)^c = \sum_b \frac{\partial f^b}{\partial x^a} \frac{\partial g^c}{\partial y^b} .$$

This is usually abbreviated to

$$\frac{\partial}{\partial x^a} = \sum_b \frac{\partial y^b}{\partial x^a} \frac{\partial}{\partial y^b} . \tag{4.2}$$

When $m = n$, the determinant of the matrix $\frac{\partial y^b}{\partial x^a}$ is called Jacobian of the map, and the map is invertible whenever the Jacobian is non-zero.

4.2 Vectors

To construct the tangent space T_p using objects that are intrinsic to M we proceed as follows. Let $\mathcal{F}$ be the space of all smooth functions on M, i.e., C^∞ maps $f : M \to \mathbb{R}$. Each curve through p defines an operator on this space, namely the directional derivative, which maps $f \to \frac{df}{d\lambda}$ at p. Then the tangent space T_p can be identified with the space of directional derivative operators along curves through p. To see that this is indeed the case, note that two operators $\frac{d}{d\lambda}$ and $\frac{d}{d\eta}$ representing derivatives along two curves $x^a(\lambda)$ and $x^a(\eta)$ through p can be added and scaled by real numbers to give another operator $a\frac{d}{d\lambda} + b\frac{d}{d\eta}$. This new operator clearly acts linearly on functions and it can also be shown to satisfy the Leibniz rule. Therefore, the set of directional derivatives forms a vector space.

To identify the vector space of directional derivatives with the tangent space T_p we need to show the directional derivatives form a suitable basis for this space. To construct such a basis, consider a coordinate chart with coordinates x^a. An obvious set of n-directional derivatives at p are the partial derivatives ∂_a at this point. Note that this is the definition of partial derivative with respect to x^a: the directional derivative along a curve defined by $x^b =$ constant for all $b \neq a$, parametrised by x^a itself. We are now going to show that the partial derivative operators $\{\partial_a\}$ at p form a basis for the tangent space T_p. To do this, we are going to show that any directional derivative can be decomposed

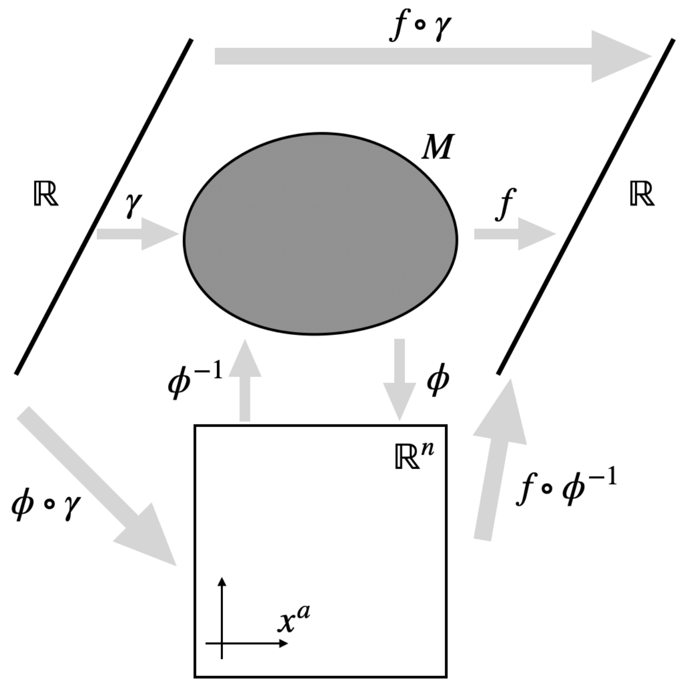

Figure 4.3: Decomposing the tangent vector to a curve $\gamma : \mathbb{R} \to M$ in terms of partial derivatives with respect to local coordinates on M.

into a linear combination of partial derivatives. Consider an n-dimensional manifold M, a coordinate chart $\phi : M \to \mathbb{R}^n$, a curve $\gamma : \mathbb{R} \to M$, and a function $f : M \to \mathbb{R}$. This leads to the tangle of maps shown in Fig. 4.3. If λ is the parameter along γ, we want to express the vector/operator $\frac{d}{d\lambda}$ in terms of partial derivatives ∂_a. Using the chain rule, we have

$$\begin{aligned} \frac{d}{d\lambda} f &= \frac{d}{d\lambda}(f \circ \gamma) \\ &= \frac{d}{d\lambda}[(f \circ \phi^{-1}) \circ (\phi \circ \gamma)] \\ &= \frac{d(\phi \circ \gamma)^a}{d\lambda} \frac{\partial (f \circ \phi^{-1})}{\partial x^a} \\ &= \frac{dx^a}{d\lambda} \partial_a f \,. \end{aligned} \tag{4.3}$$

The first line simply takes the informal expression on the left hand side and writes it as an honest derivative of the function $(f \circ \gamma) : \mathbb{R} \to \mathbb{R}$. The second line just comes from the definition of inverse map ϕ^{-1}. The third line is just the formal chain rule and the last line is a return to the informal notation of the start. Since the function f is arbitrary, we have

$$\frac{d}{d\lambda} = \frac{dx^a}{d\lambda} \partial_a \,.$$

Thus, the partial derivatives $\{\partial_a\}$ do indeed represent a good basis for the vector space of directional derivatives, and hence we can identify the latter with the tangent space.

This particular basis, i.e., $\hat{e}_{(a)} = \partial_a$, is known as a coordinate basis for T_p; it is the formalisation of the notion of setting up basis vectors to point along the coordinate axes. In general, we do not have to limit ourselves to coordinate bases when we consider tangent vectors. For example, the coordinate basis vectors are typically not normalised to unity nor orthogonal to each other. In fact, on a curved manifold, the coordinate basis will never be orthogonal throughout the neighbourhood of a point where the curvature does not vanish. One can define non-coordinate orthonormal bases by giving their components in a coordinate basis but we will not use this in these lectures.

One of the advantages of the abstract point of view that we have taken regarding vectors is that now the transformation law under changes of coordinates is immediate. Since the basis vectors are $\hat{e}_{(a)} = \partial_a$, the basis vectors in some new coordinate system $x^{a'}$ are simply given by the chain rule (4.2):

$$\partial_{a'} = \frac{\partial x^a}{\partial x^{a'}} \partial_a \,.$$

Just as in flat space, we can get the transformation law for the components of a vector V by demanding that $V = V^a \partial_a$ is unchanged by a change of basis:

$$\begin{aligned} V^a \partial_a &= V^{a'} \partial_{a'} \\ &= V^{a'} \frac{\partial x^a}{\partial x^{a'}} \partial_a \,, \end{aligned} \tag{4.4}$$

and hence, since the matrix $\frac{\partial x^{a'}}{\partial x^a}$ is the inverse of the matrix $\frac{\partial x^a}{\partial x^{a'}}$, we have

$$V^{a'} = \frac{\partial x^{a'}}{\partial x^a} V^a \,. \tag{4.5}$$

Since the basis vectors are usually not written explicitly, the rule (4.5) for transforming components of vectors is what we call "vector transformation law". Note that this law is compatible with the transformation of vector components in Special Relativity under Lorentz transformations, $V^{a'} = L^{a'}{}_a V^a$ since Lorentz transformations are just a very special kind of coordinate transformations, namely $x^{a'} = L^{a'}{}_a x^a$. However, (4.5) is completely general and it determines the transformation of vectors under arbitrary changes of coordinates.

Since a vector at a point can be thought of as directional derivative along a path through that point, a vector field defines a map from smooth functions to smooth functions all over the manifold by taking a derivative at each point. Given two vector fields X and Y, we can define the commutator $[X, Y]$ by its action on an arbitrary function $f(x^a)$:

$$[X, Y](f) \equiv X(Y(f)) - Y(X(f)) \,. \tag{4.6}$$

Clearly this operator is independent of the coordinates. Moreover, the commutator of two vector fields is itself a vector field: if f and g are functions and a and b are real numbers, the commutator is linear:

$$[X, Y](af + bg) = a[X, Y](f) + b[X, Y](g) \,,$$

and it obeys the Leibniz rule,

$$[X, Y](fg) = f[X, Y](g) + g[X, Y](f) \,.$$

Exercise: show that the components of the vector field $[X, Y]^a$ are given by

$$[X, Y]^a = X^b \partial_b Y^a - Y^c \partial_c X^a \,.$$

Exercise: using the result above show that the components of $[X, Y]^a$ transforms as a vector under general coordinate transformations.
Remark. The commutator is a special case of the Lie derivative.

4.3 Tensors

Having defined vectors on general manifolds, we can now consider dual vectors (one-forms). Once again, the co-tangent space T_p^* can be thought of as the set of linear maps $\omega : T_p \to \mathbb{R}$. The canonical example of a one-form is the gradient of a function f, denoted by df. Its action on a vector $\frac{d}{d\lambda}$ is the directional derivative of the function:

$$df\left(\frac{d}{d\lambda}\right) = \frac{df}{d\lambda}\,. \tag{4.7}$$

Like in the case of a vector, a one-form exists only at the point where it is defined and it does not depend on information at other points in M.

Just as the partial derivatives along the coordinate axes provide a natural basis for the tangent space, the gradients of the coordinate functions x^a provide a natural basis for the co-tangent space. In flat space we constructed a basis for T_p^* by demanding that $\hat{\theta}^{(a)}(\hat{e}_{(b)} = \delta^a_b$. On an arbitrary manifold M we can do the same and (4.7) leads to

$$dx^a(\partial_b) = \frac{\partial x^a}{\partial x^b} = \delta^a_b\,. \tag{4.8}$$

Therefore the gradients $\{dx^a\}$ are an appropriate basis of one-forms, and hence an arbitrary one-form can be expanded into components as $\omega = \omega_a\, dx^a$.

The transformation rules of basis dual vectors and components follow from the usual procedure. For the basis one-forms, we get

$$dx^{a'} = \frac{\partial x^{a'}}{\partial x^a}\, dx^a\,, \tag{4.9}$$

and for the components,

$$\omega_{a'} = \frac{\partial x^a}{\partial x^{a'}}\, \omega_a\,. \tag{4.10}$$

We will usually write the components ω_a we we refer to a one-form ω.

Just as in flat space, a $(k.l)$ tensor is a multilinear map from k dual vectors and l vectors to $\mathbb{R}$. Its components in a coordinate basis can be obtained by acting the tensor on the basis of one-forms and vectors,

$$T^{a_1\ldots a_k}{}_{b_1\ldots b_l} = T(dx^{a_1}, \ldots, dx^{a_k}, \partial_{b_1}, \ldots, \partial_{b_l})\,. \tag{4.11}$$

This is equivalent to the expansion

$$T = T^{a_1\ldots a_k}{}_{b_1\ldots b_l}\, \partial_{a_1} \otimes \cdots \otimes \partial_{a_k} \otimes dx^{b_1} \otimes \cdots \otimes dx^{b_l}\,.$$

The transformation law for general tensors under general changes of coordinates follows exactly the same pattern as in flat space, now replacing the Lorentz transformation matrix used in flat space with the Jacobian of the general coordinate transformation:

$$T^{a'_1\ldots a'_k}{}_{b'_1\ldots b'_l} = \frac{\partial x^{a'_1}}{\partial x^{a_1}} \cdots \frac{\partial x^{a'_k}}{\partial x^{a_k}} \frac{\partial x^{b_1}}{\partial x^{b'_1}} \cdots \frac{\partial x^{b_l}}{\partial x^{b'_l}}\, T^{a_1\ldots a_k}{}_{b_1\ldots b_l}\,. \tag{4.12}$$

It is often easier (but entirely equivalent) to transform a tensor under coordinate transformations by taking the identity of basis vectors and one-forms as partial derivatives and gradients at face value, and simply substituting in the coordinate transformation.

Example: Consider a symmetric $(0,2)$ tensor S on a two-dimensional manifold whose components in a coordinate system $(x^1 = x, x^2 = y)$ are given by:

$$S_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & x^2 \end{pmatrix} .$$

This can be equivalently written as

$$S = S_{ab}\, dx^a \otimes dx^b = (dx)^2 + x^2\,(dy)^2 \,, \tag{4.13}$$

where in the last equality we suppress the tensor product symbols for brevity (as this is common practice!). Consider the coordinate transformation:

$$x' = \frac{2\,x}{y} \,, \quad y' = \frac{y}{2} \,,$$

(valid for example when $x > 0$ and $y > 0$). This coordinate transformation can be straightforwardly inverted:

$$x = x'\,y' \,, \quad y = 2\,y' \,. \tag{4.14}$$

To calculate the components of S in the new coordinate system, $S_{a'b'}$, we could easily compute the matrix $\frac{\partial x^a}{\partial x^{a'}}$ (and its inverse $\frac{\partial x^{a'}}{\partial x^a}$ if needed) and apply the general formula for the tensor transformation law (4.12). Instead, we will use the fact that we compute the derivatives of the old coordinates in terms of the new ones using (4.14) to express dx^a in terms of $dx^{a'}$:

$$\begin{aligned} dx &= y'\,dx' + x'\,dy' \\ dy &= 2\,dy' \,. \end{aligned}$$

Plugging these expressions into (4.13) and remembering that the tensor products do not commute ($dx'\,dy' \neq dy'\,dx'$), we obtain,

$$\begin{aligned} S &= (dx)^2 + x^2\,(dy)^2 \\ &= (y'\,dx' + x'\,dy')(y'\,dx' + x'\,dy') + 2\,(2\,dy')(2\,dy') \\ &= (y')^2(dx')^2 + x'\,y'(dx'\,dy' + dy'\,dx') + [(x')^2 + 4\,(x'\,y')^2](dy')^2 \,, \end{aligned}$$

which is equivalent to

$$S_{a'b'} = \begin{pmatrix} (y')^2 & x'\,y' \\ x'\,y' & (x')^2 + 4\,(x'\,y')^2 \end{pmatrix} .$$

Most tensor operations defined in flat space, e.g., contraction, symmetrisation, etc., are unaltered on a general manifold. However, the partial derivative of a general tensor is not, in general, a new tensor. The gradient, which is the partial derivative of a scalar is an honest $(0,1)$ tensor, but the partial derivative of a higher rank tensor is not a tensor. To see this, for example, let's consider the transformation of $\partial_a W_b$, where W_a is a $(0,1)$ tensor, under a general coordinate transformation:

$$\begin{aligned} \frac{\partial}{\partial x^{a'}} W_{b'} &= \frac{\partial x^a}{\partial x^{a'}} \frac{\partial}{\partial x^a} \left(\frac{\partial x^b}{\partial x^{b'}}\, W_b \right) \\ &= \frac{\partial x^a}{\partial x^{a'}} \frac{\partial x^b}{\partial x^{b'}} \left(\frac{\partial}{\partial x^a}\, W_b \right) + W_b \frac{\partial x^a}{\partial x^{a'}} \frac{\partial}{\partial x^a} \left(\frac{\partial x^b}{\partial x^{b'}} \right) . \end{aligned} \tag{4.15}$$

4.4 The metric

The metric tensor in a general curved space is denoted by the symbol g_{ab} (while η_{ab} is specifically reserved for the Minkowski metric). The metric tensor g_{ab} is a symmetric $(0,2)$ tensor and, at least in these lectures, we will take it to be non-degenerate, that is, its determinant $g \equiv |g_{ab}|$ does not vanish. In this case, we can define the inverse metric g^{ab} via

$$g^{ab}\, g_{bc} = g_{cd}\, g^{da} = \delta^a_c\,.$$

Since g_{ab} is symmetric, the inverse metric g^{ab} is also symmetric. Just as in special relativity, the metric and its inverse can be used to raise and lower indices of tensors.

Here we outline some of the reasons why the metric is such an important object:

1. The metric provides a notion of "past" and "future".
2. The metric allows to compute the length of paths and proper time.
3. The metric determines the "shortest" distance between two points (and therefore the motion of test particles).
4. In general relativity, the metric replaces the Newtonian gravitational field ϕ.
5. The metric provides a notion of locally inertial frames and therefore a sense of "no rotation".
6. The metric determines causality, by defining the speed of light faster than which no signal can travel.
7. The metric replaces the usual Euclidean three-dimensional dot product.

In our discussion of proper time and special relativity we introduced the line element $ds^2 = \eta_{ab}dx^a dx^b$, which we used to calculate the length of the path. Now we now that dx^a is really a basis of dual vectors. In a general curved manifold, the line element is given by

$$ds^2 = g_{ab}(x)\, dx^a\, dx^b\,. \tag{4.16}$$

For example, the line element of three-dimensional Euclidean space in Cartesian coordinates is

$$ds^2 = (dx)^2 + (dy)^2 + (dz)^2\,.$$

In spherical coordinates,

$$\begin{aligned} x &= r\, \sin\theta\, \cos\theta\,,\\ y &= r\, \sin\theta\, \sin\theta\,,\\ z &= r\, \cos\theta\,, \end{aligned}$$

the line element becomes (**exercise**),

$$ds^2 = dr^2 + r^2\, d\theta^2 + \sin^2\theta\, d\phi^2\,.$$

It can be shown that, at some given point p on a manifold M, the metric g_{ab} can always be put into its canonical form, where its components are

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

Furthermore, $\partial_c g_{ab}|_p = 0$ but $\partial_c \partial_d g_{ab}|_p \neq 0$. The signature of the metric is the number of both positive and negative eigenvalues. If all signs are positive, the metric is called Euclidean or Riemannian (or just positive definite), while if there is a single minus it is called Lorentzian or pseudo-Riemannian, and any metric with some $+1$'s and some -1's is called indefinite.

With the definition (4.16) of a general metric g_{ab} on a curved space, we can straightforwardly generalise many of the notions we had in Minkowski space:

Norm of a vector: Given a vector V^a, the norm is defined via

$$|V|^2 \equiv g_{ab} V^a V^b.$$

If $|V|^2 > 0$ (or $|V|^2 < 0$) for all vectors V^a, the metric is called positive definite (or negative definite) —this is the Riemannian case. Otherwise it is called indefinite —this includes the Lorentzian case.

Inner (scalar) product between two vectors: Given two arbitrary vectors A^a and B^a, their inner (scalar) product is defined as

$$A \cdot B \equiv g_{ab} A^a B^b.$$

If $g_{ab} A^a B^b = 0$, then A^a and B^b are said to be orthogonal.

Null vectors: For indefinite metrics, null vectors are those that have zero norm, i.e., they are orthogonal to themselves: For indefinite metrics there are vectors that are orthogonal to themselves. That is,

$$g_{ab} A^a A^b = 0.$$

Note for indefinite metrics, that this does *not* imply that A^a is the zero vector ($A^a = 0$).

4.5 Covariant derivatives

In flat space in inertial coordinates, the partial derivative operator ∂_a is a map from (k, l) tensor fields to $(k, l+1)$ tensor fields which acts linearly on its arguments and obeys the Leibniz rule on tensor products. However, we have seen in (4.15) that on a general curved manifold, it is no longer true that the partial derivative operator ∂_a operator acting on a tensor produces another tensor. Therefore, we need to define a new derivative operator, called covariant derivative and denoted by the symbol ∇, which is independent of the coordinates and that maps (k, l) tensor fields to $(k, l+1)$ tensor fields. Given two arbitrary tensor fields T and S, then we demand that ∇ obeys:

1. Linearity: $\nabla(T + S) = \nabla T + \nabla S$.
2. Leibniz rule: $\nabla(T \otimes S) = (\nabla T) \otimes S + T \otimes (\nabla S)$.

It can be shown that if ∇ has to satisfy the Leibniz rule, then it can always be written as the partial derivative plus some linear transformation. For the case of a vector field V^a, this implies

$$\nabla_a V^b = \partial_a V^b + \Gamma^b{}_{ac} V^c\,, \tag{4.17}$$

where the ${\Gamma^b}_{ac}$'s are called connection coefficients. We can determine the transformation rule for the connection coefficients by demanding that the covariant derivative of a vector (4.17) transforms as a $(1,1)$ tensor under coordinate transformations, namely

$$\nabla_{a'}V^{b'} = \frac{\partial x^a}{\partial x^{a'}}\frac{\partial x^{b'}}{\partial x^b}\nabla_a V^b \,.$$

Expanding the left hand side,

$$\begin{aligned}\nabla_{a'}V^{b'} &= \frac{\partial x^a}{\partial x^{a'}}\frac{\partial x^{b'}}{\partial x^b}\nabla_a V^b \\ &= \frac{\partial x^a}{\partial x^{a'}}\frac{\partial x^{b'}}{\partial x^b}\partial_a V^b + \frac{\partial x^a}{\partial x^{a'}}V^b\frac{\partial}{\partial x^a}\left(\frac{\partial x^{b'}}{\partial x^b}\right) + {\Gamma^{b'}}_{a'c'}\frac{\partial x^{c'}}{\partial x^c}V^c \,.\end{aligned}$$

Expanding now the right hand side:

$$\frac{\partial x^a}{\partial x^{a'}}\frac{\partial x^{b'}}{\partial x^b}\nabla_a V^b = \frac{\partial x^a}{\partial x^{a'}}\frac{\partial x^{b'}}{\partial x^b}\partial_a V^b + \frac{\partial x^a}{\partial x^{a'}}\frac{\partial x^{b'}}{\partial x^b}{\Gamma^b}_{ac}V^c \,.$$

The last two expressions have to be equated; the first terms in each side are the same and hence they cancel, so we are left with

$${\Gamma^{b'}}_{a'c'}\frac{\partial x^{c'}}{\partial x^c}V^c + \frac{\partial x^a}{\partial x^{a'}}V^c\frac{\partial}{\partial x^a}\left(\frac{\partial x^{b'}}{\partial x^c}\right) = \frac{\partial x^a}{\partial x^{a'}}\frac{\partial x^{b'}}{\partial x^b}{\Gamma^b}_{ac}V^c \,.$$

Notice that in the second term on the left hand side of this equation we have changed the dummy index $b \to c$. Multiplying this equation by $\frac{\partial x^c}{\partial x^{d'}}$ and relabelling $d' \to c'$ we find

$${\Gamma^{b'}}_{a'c'} = \frac{\partial x^a}{\partial x^{a'}}\frac{\partial x^c}{\partial x^{c'}}\frac{\partial x^{b'}}{\partial x^b}{\Gamma^b}_{ac} - \frac{\partial x^a}{\partial x^{a'}}\frac{\partial x^c}{\partial x^{c'}}\frac{\partial^2 x^{b'}}{\partial x^a \partial x^c} \,. \tag{4.18}$$

This is not a tensor transformation law since the second term on the right hand side spoils it; therefore, it is clear that the connection coefficients are not the components of a tensor. They are precisely constructed to be non-tensorial but in way such that the combination (4.17) transforms as a tensor.

Equation (4.18) does not determine the connection coefficients uniquely. To further constraint the connection, we impose that the covariant derivative satisfies the following two properties, in addition to the previous ones:

3. Commutes with contractions: $\nabla_a({T^c}_{cb}) = {(\nabla T)_a}^c{}_{cb}$.

4. Reduces to the partial derivative when acting on scalars: $\nabla_a\phi = \partial\phi$.

The first property is equivalent to demand that the Kronecker delta (the identity map) is covariantly constant: $\nabla_a\delta^b_c = 0$, which resonable to impose since the components of δ^a_b are constants (zeros and ones).

Let's see what these new properties imply. Consider an arbitrary one-form ω_a and an arbitrary vector field V^a; we can compute the covariant derivative of the scalar defined by $\omega_a V^a$ to obtain:

$$\begin{aligned}\nabla_a(\omega_b V^b) &= (\nabla_a\omega_b)\,V^b + \omega_b(\nabla_a V^b) \\ &= (\nabla_a\omega_b)\,V^b + \omega_b(\partial_a V^b + {\Gamma^b}_{ac}V^c) \,.\end{aligned}$$

But

$$\nabla_a(\omega_b V^b) = \partial_a(\omega_b V^b)$$
$$= V^b \partial_a \omega_b + \omega_b \partial_a V^b$$

Equating these two expressions, re-labelling the dummy indices $b \to c$ and $c \to b$ in the term with the connection in the first expression and recalling that V^a is arbitrary, we obtain the formula for the covariant derivative of a one-form

$$\nabla_a \omega_b = \partial_a \omega_b - \Gamma^c{}_{ab}\, \omega_c\,. \tag{4.19}$$

It is now straightforward to determine the formula for the covariant derivative acting on an arbitrary (k, l) tensor. We find,

$$\begin{aligned}\nabla_c T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l} &= \partial_c T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l} \\ &+ \Gamma^{a_1}{}_{cd}\, T^{d a_2 \ldots a_k}{}_{b_1 \ldots b_l} + \cdots + \Gamma^{a_k}{}_{cd}\, T^{a_1 \ldots a_{k-1} d}{}_{b_1 \ldots b_l} \\ &- \Gamma^d{}_{a b_1}\, T^{a_1 \ldots a_k}{}_{d b_2 \ldots b_l} - \cdots - \Gamma^d{}_{a b_l}\, T^{a_1 \ldots a_k}{}_{b_1 \ldots b_{l-1} d}\,. \end{aligned} \tag{4.20}$$

Sometimes an alternative notation is used; just as commas are used to denote partial derivatives, semi-colons are used for the covariant derivatives:

$$\nabla_c T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l} = T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l ; c}\,.$$

In these lectures we will mostly use ∇_a for the covariant derivative.

Still, we have not fully specified the connection and in fact, one can define many different connections on a manifold satisfying the previous four requirements. It turns out though that every metric defines a unique connection, which is the one that is used in general relativity.

To see this, the first thing to notice is that the difference between two connections, say Γ and $\tilde{\Gamma}$, is a tensor. Indeed,

$$\nabla_a V^b - \tilde{\nabla}_a V^b = \partial_a V^b + \Gamma^b{}_{ac}\, V^c - \left(\partial_a V^b + \tilde{\Gamma}^b{}_{ac}\, V^c\right)$$
$$= (\Gamma^b{}_{ac} - \tilde{\Gamma}^b{}_{ac}) V^c\,.$$

Since the left hand side is a tensor by definition of covariant derivative, the right hand side must also be a tensor. Hence,

$$\Gamma^b{}_{ac} - \tilde{\Gamma}^b{}_{ac} = S^b{}_{ac}$$

where $S^b{}_{ac}$ is a tensor. Next, notice that given a connection $\Gamma^c{}_{ab}$, one can form another connection by simply permuting the lower indices; namely the coefficients $\Gamma^c{}_{ba}$ also transform as (4.18) and hence they determine a distinct connection. Therefore, to every connection we can associate a tensor, known as the torsion tensor, given by

$$T^c{}_{ab} = \Gamma^c{}_{ab} - \Gamma^c{}_{ba} = 2\,\Gamma^c{}_{[ab]}\,. \tag{4.21}$$

It is clear that the torsion tensor in anti-symmetric in its lower indices, and a connection that is symmetric in its lower indices (and hence the torsion tensor vanishes) is called "torsion-free".

We can now determine the unique connection on a manifold with a metric g_{ab} by introducing two additional properties:

5. Torsion-free: ${\Gamma^c}_{ab} = {\Gamma^c}_{ba}$.

6. Metric compatibility: $\nabla_c g_{ab} = 0$.

The torsion free condition implies that covariant derivatives acting on a scalar field commute:

$$\nabla_a \nabla_b \phi = \nabla_b \nabla_a \phi \,.$$

A connection is metric compatible if the covariant derivative of the metric with respect to that connection vanishes everywhere, and such a connection is known as the Levi-Civita connection.

The metric-compatibility condition implies the following for the covariant derivative of the inverse metric:

$$\begin{aligned} 0 &= \nabla_a \delta^b_c \\ &= \nabla_a (g^{bd}\, g_{dc}) \\ &= g_{dc} \nabla_a g^{bd} + g^{bd} \nabla_a g_{dc} \\ &= \nabla_a (g^{bd})\, g_{dc} \\ \Rightarrow \nabla_a g^{bd} &= 0 \,. \end{aligned}$$

In addition, a metric-compatible covariant derivative commutes with raising and lowering of indices. For example, for an arbitrary vector field V^a,

$$g_{ac} \nabla_b V^c = \nabla_b (g_{ac} V^c) = \nabla_b V_a \,.$$

Now we will show both the existence and uniqueness of the metric-compatible connection by explicitly determining the connection coefficients in terms of the metric. To do so, consider the following three expressions for the expanded metric compatibility condition obtained by permuting the indices:

$$\begin{aligned} \nabla_c g_{ab} &= \partial_c g_{ab} - {\Gamma^d}_{ca}\, g_{db} - {\Gamma^d}_{cb}\, g_{ad} = 0 \\ \nabla_a g_{bc} &= \partial_a g_{bc} - {\Gamma^d}_{ab}\, g_{dc} - {\Gamma^d}_{ac}\, g_{bd} = 0 \\ \nabla_b g_{ca} &= \partial_b g_{ab} - {\Gamma^d}_{bc}\, g_{da} - {\Gamma^d}_{ba}\, g_{cd} = 0 \end{aligned}$$

Subtracting the second and third equations from the first and using the symmetry properties of the connection, we obtain

$$\partial_c g_{ab} - \partial_a g_{bc} - \partial_b g_{ca} + 2\, {\Gamma^d}_{ab}\, g_{dc} = 0 \,.$$

Multiplying this expression by g^{ec} and re-labelling the indices, we find the final expression for the Levi-Civita connection:

$${\Gamma^c}_{ab} = \frac{1}{2}\, g^{cd} \left(\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab} \right) \,. \tag{4.22}$$

This connection is also known as the Christoffel connection and the symbols in (4.22) are referred to as the Christoffel symbols.

Example: Christoffel symbols of the flat Euclidean metric in 2 dimensions in polar coordinates:

$$ds^2 = dr^2 + r^2\, d\theta^2\,.$$

The non-zero components of the metric are $g_{rr} = 1$, $g_{\theta\theta} = r^2$, while the non-zero components of the inverse metric are $g^{rr} = 1$ and $g^{\theta\theta} = \frac{1}{r^2}$. Notice that we use r and θ as indices in a notation that should be obvious; we will continue to do so in the rest of these lectures. Using (4.22), we compute

$$\begin{aligned}
\Gamma^r{}_{rr} &= \frac{1}{2}\, g^{ra}(\partial_r g_{ra} + \partial_r g_{ra} - \partial_a g_{rr}) \\
&= \frac{1}{2}\, g^{rr}(\partial_r g_{rr} + \partial_r g_{rr} - \partial_r g_{rr}) + \frac{1}{2}\, g^{r\theta}(\partial_r g_{r\theta} + \partial_r g_{r\theta} - \partial_\theta g_{rr}) \\
&= \frac{1}{2}(1)(0+0-0) + \frac{1}{2}(0)(0+0-0) \\
&= 0\,.
\end{aligned}$$

Similarly, we compute

$$\begin{aligned}
\Gamma^r{}_{\theta\theta} &= \frac{1}{2}\, g^{ra}(\partial_\theta g_{\theta a} + \partial_\theta g_{\theta a} - \partial_a g_\theta) \\
&= \frac{1}{2}\, g^{rr}(\partial_\theta g_{\theta r} + \partial_\theta g_{\theta r} - \partial_r g_{\theta\theta}) \\
&= -r\,.
\end{aligned}$$

Proceeding in exactly the same manner as above, we find that the remaining components of the Christoffel symbols are given by

$$\begin{aligned}
&\Gamma^r{}_{\theta r} = \Gamma^r{}_{r\theta} = 0 \\
&\Gamma^\theta{}_{rr} = 0 \\
&\Gamma^\theta{}_{r\theta} = \Gamma^\theta{}_{\theta r} = \frac{1}{r} \\
&\Gamma^\theta{}_{\theta\theta} = 0\,.
\end{aligned}$$

Remark. The Christoffel symbols of the flat metric in Cartesian coordinates vanish identically (**exercise**).

4.6 Parallel transport and geodesics

In flat space, parallel transport of a vector along a curve intuitively means "keeping the vector constant" as we move it along the curve. More precisely, given a curve $x^b(\lambda)$, imposing that an arbitrary vector V^a is constant along this curve in flat space corresponds to:

$$\frac{d}{d\lambda}V^a = \frac{dx^b}{d\lambda}\frac{\partial}{\partial x^b}V^a = 0\,.$$

The crucial difference between flat and curved spaces is that, in a curved space, the result of parallel transporting a vector from one point to another will depend on the path taken between the points, see Fig. 4.4.

The generalisation of this concept to curved manifolds amounts to replace the partial derivative by a covariant derivative. Therefore, a vector V^a is said to be parallely transported along W^b if

$$W^b \nabla_b V^a = 0.$$