IV. VECTOR AND TENSOR CALCULUS ON MANIFOLDS

The laws of physics are differential equations involving (mostly) tensor-valued objects.

We therefore need to understand how to take derivatives of vectors and tensors on a general manifold, i.e., to develop vector and tensor calculus.

The issue we face is that, on a general manifold, tensors at different points inhabit separate (tangent) vector spaces and there is no unique way to compare tensors at different points.

In this topic we shall see how to construct tensor-valued *covariant derivatives* of tensors, and in so doing connect together tangent spaces at different points.

We shall also look at *geodesic curves* as an important application.

1 Covariant derivatives

1.1 Derivatives of scalar fields

Consider a scalar field $\phi(x)$ which is differentiable function of the coordinates x^a.

We saw in the last handout that the partial derivatives $\partial\phi/\partial x^a$ form the components of a dual vector, which we call the *gradient* of ϕ, since, under a change of coordinates, $\phi'(x') = \phi(x)$ and

$$\frac{\partial\phi'}{\partial x'^a} = \frac{\partial x^b}{\partial x'^a}\frac{\partial\phi}{\partial x^b}\,. \tag{1}$$

We are used to thinking of the gradient as a vector, and we can always associate a vector by forming $g^{ab}\partial\phi/\partial x^b$.

However, the gradient is more naturally thought of as a dual vector, i.e., a linear map from vectors to real numbers.

This is because the gradient maps an infinitesimal displacement – a vector with components δx^a – into the change in the function between points with coordinate separation δx^a as

$$\delta\phi = \frac{\partial\phi}{\partial x^a}\delta x^a \,. \tag{2}$$

1.2 Covariant derivatives of tensor fields

We want to work with derivatives that preserve the tensorial nature of the object being differentiated.

In Euclidean space, this is straightforward: we work in global Cartesian coordinates and take the partial derivatives of the Cartesian components of tensors.

The resulting object transforms as a Cartesian tensor under orthogonal coordinate transformations.

However, on a general manifold we cannot do this as there are no global Cartesian coordinates.

Even in Euclidean space, if we want to work in a general coordinate system the partial derivatives of the components of a tensor do not transform as a tensor.

To see the problem, consider a vector field $v^a(x)$ and construct the derivative $\partial v^b/\partial x^a$.

Now transform to some other coordinates, x'^a, in which case the vector field has components $v'^a(x')$, and take the derivative with respect to the new coordinates; we have

$$\begin{aligned}\frac{\partial v'^b}{\partial x'^a} &= \frac{\partial}{\partial x'^a}\left(\frac{\partial x'^b}{\partial x^c}v^c\right) \\ &= \frac{\partial x^d}{\partial x'^a}\frac{\partial}{\partial x^d}\left(\frac{\partial x'^b}{\partial x^c}v^c\right) \\ &= \frac{\partial x^d}{\partial x'^a}\frac{\partial x'^b}{\partial x^c}\frac{\partial v^c}{\partial x^d} + \frac{\partial x^d}{\partial x'^a}\frac{\partial^2 x'^b}{\partial x^d \partial x^c}v^c\,. \end{aligned} \tag{3}$$

The first term on the right is the usual transformation law for a type-$(1,1)$ tensor, but the second term means that $\partial v^b/\partial x^a$ do *not* form the components of a tensor.

To fix this problem requires the introduction of a more complicated derivative construction, called the *covariant derivative.*

The covariant derivative of a type-(k,l) tensor $T^{a_1\ldots a_k}{}_{b_1\ldots b_l}$ is a type-$(k,l+1)$ tensor, denoted by $\nabla_c T^{a_1\ldots a_k}{}_{b_1\ldots b_l}$, which satisfies the following usual properties of a derivative.

- *Action on scalar fields*: acting on a scalar field ϕ, the covariant derivative is simply the gradient of the scalar field, i.e.,

$$\nabla_a \phi = \frac{\partial \phi}{\partial x^a}\,. \tag{4}$$

- *Linearity*: for tensors $T^{a_1\ldots a_k}{}_{b_1\ldots b_l}$ and $S^{a_1\ldots a_k}{}_{b_1\ldots b_l}$ of the same type, and for constant scalars α and β, the covariant derivative of a linear combination is the linear combination of the covariant derivatives, i.e.,

$$\begin{aligned}\nabla_c\left(\alpha T^{a_1\ldots a_k}{}_{b_1\ldots b_l} + \beta S^{a_1\ldots a_k}{}_{b_1\ldots b_l}\right) &= \alpha \nabla_c T^{a_1\ldots a_k}{}_{b_1\ldots b_l} \\ &\quad + \beta \nabla_c S^{a_1\ldots a_k}{}_{b_1\ldots b_l}\,. \end{aligned} \tag{5}$$

- *Leibnitz rule*: for arbitrary tensors $T^{a_1\ldots a_k}{}_{b_1\ldots b_l}$ and $S^{c_1\ldots c_m}{}_{d_1\ldots d_n}$, the covariant derivative of the outer product satisfies the product rule

$$\begin{aligned}\nabla_f\left(T^{a_1\ldots a_k}{}_{b_1\ldots b_l}S^{c_1\ldots c_m}{}_{d_1\ldots d_n}\right) = \\ \left(\nabla_f T^{a_1\ldots a_k}{}_{b_1\ldots b_l}\right) S^{c_1\ldots c_m}{}_{d_1\ldots d_n} \\ + T^{a_1\ldots a_k}{}_{b_1\ldots b_l}\left(\nabla_f S^{c_1\ldots c_m}{}_{d_1\ldots d_n}\right) .\end{aligned} \tag{6}$$

1.3 The connection

We shall now try and construct an appropriate covariant derivative, starting with a vector field $v^a(x)$.

Recalling Eq. (3), our strategy is to combine $\partial v^b/\partial x^a$ with an additional piece, linear in v^a (and with v^a undifferentiated), designed to cancel the unwanted final term on the right.

We write

$$\boxed{\nabla_a v^b = \frac{\partial v^b}{\partial x^a} + \Gamma^b_{ac} v^c\,,} \tag{7}$$

where the Γ^b_{ac} are called *connection coefficients* or sometimes simply *the connection.*

Note how in the final term of Eq. (7), the b index has moved onto the connection coefficient from the vector $\boldsymbol{v}$, and a new (dummy) index c is summed over.

Although the connection coefficients have indices, they are *not* the components of a tensor.

Instead, they must transform under a change of coordinates (to Γ'^b_{ac}) in such a way that $\nabla_a v^b$ transforms as a type-$(1,1)$ tensor.

Forming the covariant derivative in the new coordinates, we have

$$\begin{aligned}
\nabla'_a v'^b &= \frac{\partial v'^b}{\partial x'^a} + \Gamma'^b_{ac} v'^c \\
&= \frac{\partial x^d}{\partial x'^a}\frac{\partial x'^b}{\partial x^c}\frac{\partial v^c}{\partial x^d} + \frac{\partial x^d}{\partial x'^a}\frac{\partial^2 x'^b}{\partial x^d \partial x^c} v^c + \Gamma'^b_{ac}\frac{\partial x'^c}{\partial x^d} v^d \\
&= \frac{\partial x^d}{\partial x'^a}\frac{\partial x'^b}{\partial x^c}\nabla_d v^c - \frac{\partial x^d}{\partial x'^a}\frac{\partial x'^b}{\partial x^c}\Gamma^c_{de} v^e \\
&\quad + \frac{\partial x^d}{\partial x'^a}\frac{\partial^2 x'^b}{\partial x^d \partial x^c} v^c + \Gamma'^b_{ac}\frac{\partial x'^c}{\partial x^d} v^d \, .
\end{aligned} \tag{8}$$

If $\nabla_a v^b$ are the components of a tensor, the final three terms on the right here must vanish for arbitrary $\boldsymbol{v}$, which requires

$$\Gamma'^a_{bc} = \frac{\partial x'^a}{\partial x^d}\frac{\partial x^e}{\partial x'^b}\frac{\partial x^f}{\partial x'^c}\Gamma^d_{ef} - \frac{\partial x^d}{\partial x'^b}\frac{\partial x^e}{\partial x'^c}\frac{\partial^2 x'^a}{\partial x^d \partial x^e} \, . \tag{9}$$

Note that the presence of the final (inhomogeneous) term on the right means that the connection coefficients do not transform as the components of a tensor.

The connection is not unique: any coefficients that satisfy Eq. (9) will give a valid covariant derivative.

However, we shall see shortly how on a manifold with a metric, the metric naturally picks out a unique connection.

Given two connections, Γ and $\tilde{\Gamma}$, which satisfy Eq. (9), their difference does transform as a type-$(1,2)$ tensor since the last term on the right of Eq. (9) cancels.

This means that, generally, the connection is unique up to a type-$(1,2)$ tensor.

1.3.1 Extension to other tensor fields

We now construct the covariant derivative for more general tensor fields.

First, consider a type-$(2,0)$ tensor T^{ab}, which we can always decompose into a sum of outer products of vectors.

We can consider these terms separately because of linearity of the covariant derivative, so consider $T^{ab} = u^a v^b$ for some vectors u^a and v^b.

The Leibnitz rule gives

$$\begin{aligned} \nabla_a \left(u^b v^c\right) &= \left(\nabla_a u^b\right) v^c + u^b \left(\nabla_a v^c\right) \\ &= \left(\frac{\partial u^b}{\partial x^a} + \Gamma^b_{ad} u^d\right) v^c + u^b \left(\frac{\partial v^c}{\partial x^a} + \Gamma^c_{ad} v^d\right) \\ &= \frac{\partial}{\partial x^a}\left(u^b v^c\right) + \Gamma^b_{ad} u^d v^c + \Gamma^c_{ad} u^b v^d \,, \end{aligned} \tag{10}$$

so that, generally,

$$\nabla_a T^{bc} = \frac{\partial T^{bc}}{\partial x^a} + \Gamma^b_{ad} T^{dc} + \Gamma^c_{ad} T^{bd} \,. \tag{11}$$

For dual vector fields, $X_a(x)$, the covariant derivative is inherited from that for vector fields if we impose the further requirement that *the covariant derivative commutes with contraction.*

Let us think about what this means for the scalar formed from the contraction of a dual vector X_a and a vector v^a.

Using, in addition, the Leibnitz rule, we have

$$\nabla_a \left(X_b v^b\right) = \left(\nabla_a X_b\right) v^b + X_b \left(\nabla_a v^b\right) \,. \tag{12}$$

However, we have already specified that the covariant derivative of a scalar is the gradient, so

$$\nabla_a \left(X_b v^b\right) = \frac{\partial X_b}{\partial x^a} v^b + X_b \frac{\partial v^b}{\partial x^a} \,. \tag{13}$$

Comparing with Eq. (12), and using the expansion of the covariant derivative of a vector in terms of the connection, we are left with

$$\boxed{\nabla_a X_b = \frac{\partial X_b}{\partial x^a} - \Gamma^c_{ab} X_c \,.} \tag{14}$$

Note, in particular, the minus sign and the placement of indices on the connection term.

We can build up the covariant derivative of more general tensors now as outer products of vectors and dual vectors as needed.

For example, for rank-2 tensors we have

$$\boxed{\begin{aligned} \nabla_c T^{ab} &= \partial_c T^{ab} + \Gamma^a_{cd} T^{db} + \Gamma^b_{cd} T^{ad} \\ \nabla_c T^a{}_b &= \partial_c T^a{}_b + \Gamma^a_{cd} T^d{}_b - \Gamma^d_{cb} T^a{}_d \\ \nabla_c T_{ab} &= \partial_c T_{ab} - \Gamma^d_{ca} T_{db} - \Gamma^d_{cb} T_{ad} \,. \end{aligned}} \tag{15}$$

Here, we have introduced a very convenient shorthand notation writing $\partial/\partial x^a$ as ∂_a; we shall use this extensively from now on.

Finally, we note that the covariant derivative of the mixed metric tensor $g^a{}_b$ vanishes since

$$\begin{aligned} \nabla_c g^a{}_b &= \partial_c \delta^a_b + \Gamma^a_{cd} \delta^d_b - \Gamma^d_{cb} \delta^a_d \\ &= \Gamma^a_{cb} - \Gamma^a_{cb} = 0 \,, \end{aligned} \tag{16}$$

where we used $g^a{}_b = \delta^a_b$.

This is equivalent to requiring that the covariant derivative commutes with contraction.

1.4 The metric connection

On a manifold equipped with a metric, such as the spacetime of general relativity, there is a natural connection that is singled out by the following two further conditions.

- *Metric compatibility*, where we enforce that the covariant derivative of the metric vanishes:

$$\boxed{\nabla_a g_{bc} = 0 \,.} \tag{17}$$

- *Commutative action on scalar fields*, so that

$$\boxed{\nabla_a \nabla_b \phi = \nabla_b \nabla_a \phi\,.} \tag{18}$$

We shall see shortly why it is reasonable to impose these conditions.

However, for the moment let us just explore their consequences.

We begin with the commutative action on scalar fields; this implies that the connection must be symmetric in its lower indices,

$$\Gamma^a_{bc} = \Gamma^a_{cb}\,. \tag{19}$$

To see this, we expand $\nabla_a \nabla_b \phi$ as

$$\nabla_a \nabla_b \phi = \partial_a \partial_b \phi - \Gamma^c_{ab} \partial_c \phi\,. \tag{20}$$

The first term on the right is symmetric in a and b, so if $\nabla_{[a} \nabla_{b]} \phi = 0$ for all ϕ, we must have $\Gamma^c_{[ab]} = 0$.

More generally, the antisymmetric part of the connection transforms as a tensor (this follows from Eq. 9), which is called the *torsion tensor*.

However, in general relativity we shall only be concerned with a symmetric, or torsion-free, connection so that $\nabla_{[a} \nabla_{b]} \phi = 0$.

We now turn to metric compatibility:

$$0 = \nabla_c g_{ab} = \partial_c g_{ab} - \Gamma^d_{ca} g_{db} - \Gamma^d_{cb} g_{ad}\,. \tag{21}$$

If we write down the other two cyclic permutations of the indices a, b and c, we have

$$0 = \partial_b g_{ca} - \Gamma^d_{bc} g_{da} - \Gamma^d_{ba} g_{cd}\,, \tag{22}$$

$$0 = \partial_a g_{bc} - \Gamma^d_{ab} g_{dc} - \Gamma^d_{ac} g_{bd}\,. \tag{23}$$

Adding Eqs (21) and (23) and subtracting Eq. (22), and using the symmetry of the connection gives

$$2\Gamma^d_{ca} g_{db} = \partial_c g_{ab} + \partial_a g_{bc} - \partial_b g_{ca} \,. \tag{24}$$

Solving for Γ by contracting with the inverse metric, we find an explicit and unique expression for the connection coefficients[1]:

$$\boxed{\Gamma^a_{bc} = \frac{1}{2} g^{ad} \left(\partial_b g_{dc} + \partial_c g_{db} - \partial_d g_{bc} \right) \,.} \tag{25}$$

This expression allows computation of the connection coefficients in an arbitrary coordinate system.

The covariant derivative of the inverse metric also has vanishing covariant derivative,

$$\nabla_a g^{bc} = 0 \,, \tag{26}$$

which follows from taking the covariant derivative of $g_{ab} g^{bc} = \delta^c_a$.

1.4.1 Other useful properties of the metric connection

Since $\nabla_a g_{bc} = 0$, we can interchange the order of raising/lowering indices and covariant differentiation, e.g.,

$$\begin{aligned} \nabla_c T^{ab} &= \nabla_c \left(g^{bd} T^a{}_d \right) \\ &= \left(\nabla_c g^{bd} \right) T^a{}_d + g^{bd} \left(\nabla_c T^a{}_d \right) \\ &= g^{bd} \left(\nabla_c T^a{}_d \right) \,. \end{aligned} \tag{27}$$

Note that the (downstairs) index associated with the covariant derivative is a genuine tensor index and so can be raised with the inverse metric in the usual way, e.g.,

$$\nabla^a v^b = g^{ac} \nabla_c v^b \,. \tag{28}$$

Finally, we sometimes require the connection coefficients summed over the upper and a lower index, which we denote by Γ^a_{ab}.

[1] The coefficients of the metric connection are sometimes called *Christoffel symbols.*

We can relate this to the derivative of the (coordinate-dependent) determinant of the metric functions as follows.

Since $\nabla_c g_{ab} = 0$, we have

$$\partial_c g_{ab} = \Gamma^d_{ca} g_{db} + \Gamma^d_{cb} g_{ad} \,, \tag{29}$$

which implies

$$\begin{aligned} g^{ab}\partial_c g_{ab} &= g^{ab}\left(\Gamma^d_{ca} g_{db} + \Gamma^d_{cb} g_{ad}\right) \\ &= 2 g^{ab} g_{db} \Gamma^d_{ca} \\ &= 2\Gamma^a_{ac} \,. \end{aligned} \tag{30}$$

The contraction on the left can be written as $g^{-1}\partial_c g$, where g is the determinant of the matrix with elements given by the metric functions.

This follows from the general result (known as Jacobi's formula) for an invertible matrix $\mathbf{M}$:

$$(\det\mathbf{M})^{-1}\,\partial_c \det\mathbf{M} = \mathrm{Tr}\left(\mathbf{M}^{-1}\partial_c\mathbf{M}\right) \,. \tag{31}$$

(A simple proof for the case of a symmetric matrix follows from taking the derivative of the result[2]

$$\ln(\det\mathbf{M}) = \mathrm{Tr}\,(\ln\mathbf{M}) \;, \tag{32}$$

but Eq. (31) holds generally.)

Putting these pieces together, we get the useful result

$$\Gamma^a_{ac} = \frac{1}{2} g^{-1}\partial_c g = |g|^{-1/2}\partial_c |g|^{1/2} \,. \tag{33}$$

1.5 Relation to local Cartesian coordinates

The covariant derivative constructed with the metric connection has the nice property that it reduces to partial differentiation in local Cartesian coordinates.

[2]The log of a symmetric matrix is defined by symmetrising with an orthogonal matrix $\mathbf{O}$, taking the log of the diagonal elements of the resultant, and rotating back with $\mathbf{O}^T$.

Recall that at any point P, we can find local Cartesian coordinates such that

$$g_{ab}(P) = \mathrm{diag}(\pm 1, \pm 1, \ldots, \pm 1)\,, \quad \left.\frac{\partial g_{ab}}{\partial x^c}\right|_P = 0\,. \quad (34)$$

Since the derivative of the metric vanishes at P, the metric connection also vanishes there and, in these coordinates, the components of the covariant derivative of a tensor reduce at P to the partial derivatives of the components of the tensor.

This is very important for enforcing the equivalence principle as it is straightforward to check that some law of physics, written as a tensor equation, reduces to its usual special-relativistic form in local Cartesian coordinates.

Moreover, in Euclidean space, we see that the metric-compatible covariant derivative is equivalent *everywhere* to the usual derivative employed in Euclidean tensor calculus.

Indeed, in this case, one can *define* the covariant derivative of a tensor by specifying that its form in global Cartesian coordinates is simply the partial derivatives of the Cartesian components; the form in some general coordinate system then follows from the appropriate coordinate transformation of these components.

This is exactly what we do when constructing expressions for derivative operations on tensors in curvilinear coordinates in Euclidean space.

Example: covariant derivative in Euclidean space

Let x^a be a global Cartesian coordinate system in Euclidean space, and x'^a some other general coordinate system.

Given a vector field $\boldsymbol{v}$, with Cartesian components v^a, let us define the covariant derivative of $\boldsymbol{v}$ to be that tensor whose Cartesian components are $\partial_a v^b$.

The components in the x'^a coordinates are then given by the usual transformation law for the components of a type-$(1,1)$ tensor, so

$$\nabla'_a v'^b = \frac{\partial x^c}{\partial x'^a}\frac{\partial x'^b}{\partial x^d}\frac{\partial v^d}{\partial x^c}\,. \tag{35}$$

Let us express this in terms of derivatives of the components $\boldsymbol{v}$ in the x'^a coordinates, using

$$\begin{aligned}\frac{\partial v^d}{\partial x^c} &= \frac{\partial}{\partial x^c}\left(\frac{\partial x^d}{\partial x'^e}v'^e\right) \\ &= \frac{\partial}{\partial x^c}\left(\frac{\partial x^d}{\partial x'^e}\right)v'^e + \frac{\partial x^d}{\partial x'^e}\frac{\partial v'^e}{\partial x^c}\,,\end{aligned} \tag{36}$$

to give

$$\nabla'_a v'^b = \frac{\partial v'^b}{\partial x'^a} + \underbrace{\frac{\partial^2 x^d}{\partial x'^a \partial x'^e}\frac{\partial x'^b}{\partial x^d}}_{\Gamma'^b_{ae}} v'^e\,. \tag{37}$$

We see that a connection-like term (with the connection being symmetric) naturally arises as a consequence of a non-linear coordinate transformation or, equivalently, as the basis vectors $\partial/\partial x'^a$ of the primed coordinate system having Cartesian components $\partial x^b/\partial x'^a$ that depend on position.

Indeed, we can verify that the connection coefficients that appear in Eq. (37) are exactly the metric connection by noting that the metric in the primed coordinates is

$$g'_{ab} = \frac{\partial x^c}{\partial x'^a}\frac{\partial x^d}{\partial x'^b}\delta_{cd}\,, \tag{38}$$

and for the inverse,

$$g'^{ab} = \frac{\partial x'^a}{\partial x^c}\frac{\partial x'^b}{\partial x^d}\delta^{cd}\,. \tag{39}$$

Forming the connection coefficients from

$$\Gamma'^{b}_{ae} = \frac{1}{2} g'^{bf} \left(\frac{\partial g'_{ef}}{\partial x'^{a}} + \frac{\partial g'_{af}}{\partial x'^{e}} - \frac{\partial g'_{ae}}{\partial x'^{f}} \right) \tag{40}$$

gives (exercise!)

$$\Gamma'^{b}_{ae} = \frac{\partial^2 x^d}{\partial x'^{a} \partial x'^{e}} \frac{\partial x'^{b}}{\partial x^d} , \tag{41}$$

consistent with Eq. (37).

1.6 Divergence, curl and the Laplacian

The familiar operations of taking the divergence and curl of a vector field, and the Laplacian, generalise to tensor calculus on manifolds.

The *divergence* of a vector field $\boldsymbol{v}$ is the scalar field $\nabla_a v^a$.

It follows from Eq. (33) that

$$\nabla_a v^a = \partial_a v^a + \Gamma^{a}_{ab} v^b = |g|^{-1/2} \partial_a \left(|g|^{1/2} v^a \right) , \tag{42}$$

which is often convenient.

The *curl* of a dual-vector field $\boldsymbol{X}$ is defined to be the antisymmetric part of its covariant derivative; it is the type-$(0,2)$ tensor

$$(\mathrm{curl}\boldsymbol{X})_{ab} \equiv \nabla_a X_b - \nabla_b X_a \,. \tag{43}$$

The curl is actually independent of the connection (for a symmetric connection) since

$$\begin{aligned} \nabla_a X_b - \nabla_b X_a &= \partial_a X_b - \Gamma^{c}_{ab} X_c - \partial_b X_a + \Gamma^{c}_{ba} X_c \\ &= \partial_a X_b - \partial_b X_a \,. \end{aligned} \tag{44}$$

The curl of a gradient vanishes by construction for a symmetric connection: $\nabla_{[a} \nabla_{b]} \phi = 0$.

You are used to thinking of the curl as a vector, obtained by contracting $(\mathrm{curl}\boldsymbol{X})_{ab}$ with the Levi-Civita (alternating) symbol, but this does not generalise to beyond three dimensions.

Finally, we generalise the Laplacian operator. Acting on a scalar field ϕ, we have

$$\nabla^2\phi \equiv \nabla_a\left(g^{ab}\nabla_b\phi\right) = |g|^{-1/2}\partial_a\left(|g|^{1/2}g^{ab}\partial_b\phi\right) . \quad (45)$$

The Laplacian generalises to tensor fields, e.g.,

$$\nabla^2 T^{ab} = g^{cd}\nabla_c\nabla_d T^{ab} . \quad (46)$$

2 Intrinsic derivative of vectors along a curve

We often need to take derivatives of tensors defined along a curve, for example, the derivative of some tensor-valued property of a particle with respect to proper time for the particle.

Consider a vector $\boldsymbol{v}(u)$ defined along a curve $x^a(u)$.

The *intrinsic derivative* of $\boldsymbol{v}$ along the curve $x^a(u)$ is the vector [defined along $x^a(u)$] obtained by contracting the tangent vector to the curve, dx^a/du, with the covariant derivative of $\boldsymbol{v}$; we write

$$\frac{Dv^a}{Du} \equiv \frac{dx^b}{du}\nabla_b v^a = \frac{dx^b}{du}\left(\partial_b v^a + \Gamma^a_{bc}v^c\right) . \quad (47)$$

Note that, since

$$\frac{dx^b}{du}\frac{\partial v^a}{\partial x^b} = \frac{dv^a}{du} , \quad (48)$$

we only require knowledge of $\boldsymbol{v}$ along the curve $x^a(u)$ to compute the intrinsic derivative:

$$\frac{Dv^a}{Du} = \frac{dv^a}{du} + \frac{dx^b}{du}\Gamma^a_{bc}v^c . \quad (49)$$

Note carefully the distinction between dv^a/du and Dv^a/Du:

- dv^a/du are the usual ordinary derivatives of the components of $\boldsymbol{v}$ with respect to u, and do not form the components of a vector;
- Dv^a/Du include the connection term and do form the components of a vector.

The intrinsic derivative can be extended to other tensor-valued objects.

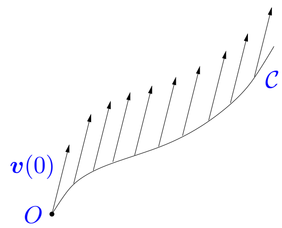
For example, for a type-$(1,1)$ tensor $T^a{}_b(u)$, we define the intrinsic derivative as

$$\frac{DT^a{}_b}{Du} = \frac{dx^c}{du}\nabla_c T^a{}_b = \frac{dT^a{}_b}{du} + \frac{dx^c}{du}\left(\Gamma^a_{cd}T^d{}_b - \Gamma^d_{cb}T^a{}_d\right) . \tag{50}$$

3 Parallel transport

Consider a curve $\mathcal{C}$ defined in 2D Euclidean space in Cartesian coordinates by $x^a(u)$.

At some initial point O, where $u = 0$, take a vector $\boldsymbol{v}(0)$ and transport it along $\mathcal{C}$ keeping its Cartesian components constant, so preserving its length and direction (see figure to the right).



The resulting vector field $\boldsymbol{v}(u)$, defined along $x^a(u)$, is said to be *parallel transported* along $x^a(u)$.

In this Euclidean example, in Cartesian coordinates we have $dv^a/du = 0$.

This is equivalent to the tensor equation, $Dv^a/Du = 0$, when written in Cartesian coordinates, but the tensor equation now gives a coordinate-independent notion of parallel transport in Euclidean space.

More generally, we define parallel transport on a Riemannian manifold by

$$\boxed{\frac{Dv^a}{Du} = 0\,.} \tag{51}$$

This definition easily extends to parallel transport of other tensors, e.g., $DT^{ab}/Du = 0$.

3.1 Properties of parallel transport

Note the following properties of parallel transport.

- The equation $Dv^a/Du = 0$ is an ordinary differential equation for the components v^a, and it has a unique solution if the v^a are specified at some initial point A.
- The vector obtained by parallel transporting from A to a second point B on the curve $x^a(u)$ is independent of the parameterisation used since, for an infinitesimal step, the change in the components are
$$\delta v^a = \delta u \frac{dv^a}{du} = -\delta u \Gamma^a_{bc}\frac{dx^b}{du}v^c = -\Gamma^a_{bc}\delta x^b v^c\,. \tag{52}$$
- The length of a vector is preserved under parallel transport since[3]
$$\frac{d|\boldsymbol{v}|^2}{du} = \frac{D}{Du}\left(g_{ab}v^a v^b\right) = 2g_{ab}v^a\frac{Dv^b}{Du} = 0\,. \tag{53}$$
- More generally, if two vectors are parallel transported along a curve, their scalar product is constant.

Note from Eq. (52) how the connection, through the operation of parallel transport, allows us to *connect* vectors

[3] The intrinsic derivative inherits the properties of the covariant derivative, such as commutativity with contraction and the Leibnitz property.

at neighbouring points separated by coordinate increments δx^a.

If we make such a step in local Cartesian coordinates, we keep the components of the vector constant.

However, we generally cannot find a global system of such coordinates and this leads to a major difference between parallel transport in Euclidean and non-Euclidean space: the latter is generally path dependent, and so the vector obtained by parallel transporting around a closed loop differs from the original vector (see Fig. 3.1).

This path dependence is a measure of the intrinsic curvature of the manifold (which we shall discuss in detail later in the course).

Finally, we note that on a surface embedded in Euclidean space, parallel transport from a point A to an infinitesimally-separated point B corresponds to parallel transport in the embedding space followed by projection into the surface at B (see *General Theory of Relativity* by Dirac).

4 Geodesic curves

Geodesic curves on a manifold are the generalisation of straight lines in Euclidean space.

They can be defined as curves of extremal distance between two points (except in the special case of null curves; see below).

Geodesics can equivalently be defined as curves $x^a(u)$ that parallel transport their tangent vector $t^a = dx^a/du$, generalising the usual notion of "straight" in Euclidean space.

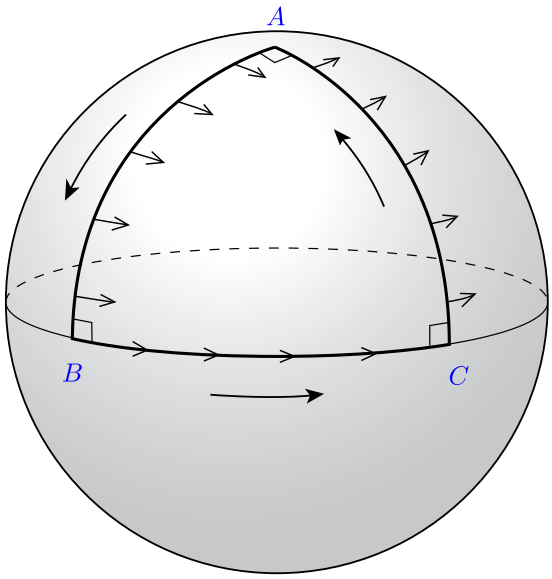


Figure 1: Parallel transport around a closed path on the surface of the 2-sphere. The path consists of a great circle through the north pole (A) down to the equator at B, a length of the equator from B to C, and the great circle through C and A. The vector indicated by the small arrows is parallel transported around this path and ends up back at A rotated by $\pi/2$. (You may wish to revisit this figure once you have covered Sec. 4 to understand why the indicated small arrows are parallel transported along the geodesic curves.)

Geodesics are important in general relativity because, as we shall argue later, free test particles[4], including massless particles, follow geodesic curves in spacetime.

4.1 Tangent vectors

We have already mentioned the idea of a tangent vector to a curve.

For a curve $x^a(u)$, the tangent vector is a vector $\boldsymbol{t}$ with coordinate components

$$t^a = \frac{dx^a}{du} \,. \tag{54}$$

Note that the tangent vector depends on the choice of

[4]In this context, a test particle is supposed to have sufficiently small mass that its motion does not affect the spacetime geometry.

parameterisation (although the tangent vectors in all parameterisations are parallel, of course).

In a pseudo-Riemannian manifold, the square of a vector, defined by $\boldsymbol{g}(\boldsymbol{t},\boldsymbol{t})$, is said to be timelike, spacelike or null according to

$$\boxed{\begin{array}{ll} \boldsymbol{g}(\boldsymbol{t},\boldsymbol{t}) > 0 & \text{timelike;} \\ \boldsymbol{g}(\boldsymbol{t},\boldsymbol{t}) < 0 & \text{spacelike;} \\ \boldsymbol{g}(\boldsymbol{t},\boldsymbol{t}) = 0 & \text{null.} \end{array}}$$

At a point, a curve is timelike, spacelike or null according to the character of its tangent vector there.

For a non-null curve, the length of the tangent vector is the derivative of the proper path length s along the curve with respect to the parameter u:

$$|\boldsymbol{t}| = \left| g_{ab} \frac{dx^a}{du}\frac{dx^b}{du} \right|^{1/2} = \left| \frac{ds}{du} \right| . \tag{55}$$

4.2 Stationary property of non-null geodesics

Consider a non-null curve $x^a(u)$ between points A and B, with $u = 0$ at A and $u = 1$ at B.

The length from A to B along the curve is

$$L = \int_A^B ds = \int_0^1 \underbrace{\left| g_{ab}\dot{x}^a\dot{x}^b \right|^{1/2}}_{F} du \,, \tag{56}$$

where $\dot{x}^a = dx^a/du$.

The form of the integrand F is invariant under reparameterisation: if we switch to some other parameter $\kappa(u)$, where $\kappa(u)$ is monotonic in the interval $0 \le u \le 1$, the length becomes

$$L = \int_{\kappa(0)}^{\kappa(1)} \left| g_{ab} \frac{dx^a}{d\kappa}\frac{dx^b}{d\kappa} \right|^{1/2} d\kappa \,. \tag{57}$$

If the curve is extremal, the length is unchanged to first order for arbitrary changes in the path, $x^a(u) \to x^a(u) + \delta x^a(u)$, which have fixed endpoints.

This is a standard problem in the calculus of variations, and extremal curves satisfy the Euler–Lagrange equations

$$\frac{\partial F}{\partial x^a} = \frac{d}{du}\left(\frac{\partial F}{\partial \dot{x}^a}\right) . \tag{58}$$

The derivatives here are

$$\begin{aligned} \frac{\partial F}{\partial x^c} &= \pm \frac{1}{2F} \partial_c g_{ab} \dot{x}^a \dot{x}^b \, , & (59) \\ \frac{\partial F}{\partial \dot{x}^c} &= \pm \frac{1}{F} g_{ac} \dot{x}^a \, , & (60) \end{aligned}$$

with the $+$ sign for timelike curves and the $-$ sign for spacelike.

The Euler–Lagrange equations become

$$\frac{d}{du}\left(\frac{1}{F} g_{ac}\dot{x}^a\right) = \frac{1}{2F}\partial_c g_{ab}\dot{x}^a\dot{x}^b . \tag{61}$$

The left-hand side is

$$\frac{d}{du}\left(\frac{1}{F} g_{ac}\dot{x}^a\right) = -\frac{1}{F^2}\frac{dF}{du} g_{ac}\dot{x}^a + \frac{1}{F} g_{ac}\ddot{x}^a + \frac{1}{F}\partial_b g_{ac}\dot{x}^a\dot{x}^b \, , \tag{62}$$

where we used $dg_{ac}/du = \partial_b g_{ac} \dot{x}^b$.

Moving terms around, we have

$$\begin{aligned} g_{ac}\ddot{x}^a &= \frac{1}{F}\frac{dF}{du} g_{ac}\dot{x}^a - \frac{1}{2}\left[2\partial_b g_{ac} - \partial_c g_{ab}\right]\dot{x}^a\dot{x}^b \\ \Rightarrow \quad \ddot{x}^d &= \frac{1}{F}\frac{dF}{du}\dot{x}^d - \frac{1}{2} g^{dc}\left[\partial_a g_{bc} + \partial_b g_{ac} - \partial_c g_{ab}\right]\dot{x}^a\dot{x}^b \\ &= \frac{1}{F}\frac{dF}{du}\dot{x}^d - \Gamma^d_{ab}\dot{x}^a\dot{x}^b . \qquad (63) \end{aligned}$$

We find that a non-null geodesic satisfies

$$\ddot{x}^a + \Gamma^a_{bc}\dot{x}^b\dot{x}^c = \left(\frac{\ddot{s}}{\dot{s}}\right)\dot{x}^a \, , \tag{64}$$

where we have used $F = ds/du$.

This is a tensor equation; this is more obvious if we write it in terms of the tangent vector $t^a = dx^a/du$ since then it becomes

$$\frac{Dt^a}{Du} = \left(\frac{\ddot{s}}{\dot{s}}\right) t^a \,. \tag{65}$$

There is a preferred class of parameters such that Eq. (64) simplifies to

$$\boxed{\ddot{x}^a + \Gamma^a_{bc}\dot{x}^b\dot{x}^c = 0\,.} \tag{66}$$

Such parameters have $\ddot{s} = 0$ and so are linearly related to the path length: $u = as + b$ for constants a and b.

These are called *affine parameters*.

4.3 Relation to parallel transport

For a non-null geodesic in an affine parameterisation, the tangent vector $t^a = dx^a/du$ is parallel transported

$$\boxed{\frac{Dt^a}{Du} = 0\,.} \tag{67}$$

Indeed, an equivalent definition of a non-null geodesic with affine parameterisation is that it is a curve whose tangent vector is parallel transported.[5]

This is all consistent with what we know in Euclidean space: there, a geodesic between two points is just the straight line connecting them, and the tangent vector is constant if we use a parameter linearly related to length along the line (i.e., an affine parameter).

Note that $Dt^a/Du = 0$ means that the length of the tangent vector is constant, which makes sense as $|\boldsymbol{t}| = ds/du$ and is constant for an affine parameter.

[5]For connections more general than the metric connection, such *auto-parallel curves* are generally non-geodesic.

For null curves, we cannot use the stationary property to define geodesics since the path length vanishes.

Instead, we define *null geodesics* as curves with null tangent vector satisfying Eq. (67).

In all cases, if we pick a vector at some starting point, and then solve $Dt^a/Du = 0$ and $t^a = dx^a/du$, we generate a unique geodesic curve in an affine parameterisation that is everywhere timelike, spacelike or null according to the character of the initial vector.

This follows since parallel transport preserves $g_{ab}t^a t^b$.

4.4 Alternative "Lagrangian" procedure

There is an alternative Lagrangian procedure to generate the equations for an affinely-parameterised geodesic.

Consider lowering the index on the equation of parallel transport for the tangent vector t^a of a geodesic in an affine parameterisation:

$$g_{ab}\frac{Dt^a}{Du} = 0 \quad \Rightarrow \quad \frac{Dt_a}{Du} = \frac{dt_a}{du} - \Gamma^c_{ba} t^b t_c = 0\,. \tag{68}$$

Using the explicit form for the metric connection gives

$$\frac{dt_a}{du} - \frac{1}{2} g^{cd} \left(\partial_b g_{ad} + \partial_a g_{bd} - \partial_d g_{ab}\right) t^b t_c = 0\,. \tag{69}$$

The first and third terms in brackets cancel to leave the following useful alternative form of the geodesic equation:

$$\boxed{\frac{dt_a}{du} = \frac{1}{2}\partial_a g_{bc} t^b t^c\,.} \tag{70}$$

As $t_a = g_{ab}dx^b/du$, we have

$$\frac{d}{du}\left(g_{ab}\frac{dx^b}{du}\right) = \frac{1}{2}\frac{\partial g_{bc}}{\partial x^a}\frac{dx^b}{du}\frac{dx^c}{du}\,. \tag{71}$$

This is exactly the Euler–Lagrange equation,

$$\frac{\partial L}{\partial x^a} = \frac{d}{du}\left(\frac{\partial L}{\partial \dot{x}^a}\right), \tag{72}$$

which would follow from the "Lagrangian"

$$L = g_{ab}\frac{dx^a}{du}\frac{dx^b}{du}. \tag{73}$$

This route through to the geodesic equations in an affine parameterisation is often very convenient as it avoids us having to compute the metric connection directly.

We shall make use of this later in the course when discussing motion around spherical masses.

4.5 Conserved quantities along geodesics

For an affinely-parameterised geodesic, the tangent vector $\boldsymbol{t}$ is parallel transported so $|\boldsymbol{t}|$ is constant.

For a non-null geodesic, we can always take $|\boldsymbol{t}| = 1$ by taking $u = s$, where, recall, s is path length along the curve.

For a null geodesic, we have $|\boldsymbol{t}| = 0$.

The constraint

$$g_{ab}\frac{dx^a}{du}\frac{dx^b}{du} = \text{const.} \tag{74}$$

is a very useful first-integral of the geodesic equation.

This first integral follows directly from the alternative Lagrangian approach by noting that L does not depend explicitly on u (in the same way that energy conservation arises in classical mechanics when the Lagrangian has no explicit time dependence).

Further conserved quantities arise when the manifold has special symmetries.

In particular, from Eq. (70) we see that

$$\boxed{\partial_c g_{ab} = 0 \quad \Rightarrow \quad t_c = \text{const.}} \tag{75}$$

In words, if the metric does not depend on a coordinate x^c, then the cth component of the tangent (dual) vector is conserved along an affinely-parameterised geodesic.

This also follows directly from the alternative Lagrangian route as conservation of the *conjugate momentum*, $\pi_c = \partial L/\partial \dot{x}^c$, if the Lagrangian does not depend on x^c, i.e., $\partial_c g_{ab} = 0$.