

1 Local Linearity

1.1 Tensor Fields

When we were dealing with objects constructed from a single vector space, the concept of linearity was simple: Linear combinations of arguments gave corresponding linear combinations of results as in $F(au + bv, k) = aF(u, k) + bF(v, k)$ where $a$ and $b$ are constants. The requirement for a function to be a tensor was just that it be linear in all of its arguments. Now we have a vector space $T_P$ of tangent vectors at each point $P$ of a manifold and wish to discuss vector fields such as $v$ that assign a vector $v(P)$ in the space $T_P$ to each point $P$ in a manifold. Two such vector fields $u, v$ can be combined to make a new vector field $w = fu + gw$

that assigns the vector $w(P) = f(P)u(P) + g(P)w(P)$ to each point $P$. The coefficients $f$ and $g$ can now be functions on the manifold. An operation that yields $F(q)$ when it acts on a vector field $q$ can be linear in the old sense and involve derivatives or even integrals over the manifold. We need a much more restrictive definition of linearity for this situation.

An operation that assigns a vector field $F(q)$ to a vector field $q$ on a manifold is said to be **locally linear** if it is linear and also has the property

$$F(fu)|_P = f(P)F(u)|_P.$$

Similarly, we can introduce an internal vector space $I_P$ at each point of a manifold and define fields with values in those spaces and locally linear operations on those spaces as well. In words, an operation is locally linear if it commutes with multiplication by a function of position. An operation that is locally linear in all of its vector arguments is called a tensor field.

An example of a simple operation that is linear, but not locally linear, is the Lie derivative of one vector field $u$ with respect to another vector field $w$. Each vector field is thought of as a directional derivative, so there is an obvious way for such a field to differentiate a scalar function:

$$\mathcal{L}_w f = wf.$$

There is also a simple way for one vector field to act on another and give a vector field as the result,

$$\mathcal{L}_wu = wu - uw = [w, u].$$

Clearly a derivative such as $\mathcal{L}_w$ cannot be locally linear in the thing that it acts on. However, if we let it act on a scalar function, we see that it is locally linear in the vector field $w$. For any functions $f, g$,

$$\mathcal{L}_{gw}f = g\mathcal{L}_wf.$$

However, this local linearity does not work when the derivative acts on a vector field.
A simple way to state the entire program of differential geometry is that it consists of finding all of the possible operations on a manifold that are either locally linear or Leibnizian in all of their arguments.

1.2 Dot Notation and Local Linearity

A subtle shift in properties accompanies the apparently innocent change in notation that occurs from one side of the following equation to the other.

\[ v f = v \cdot df \]

On the left side of this equation, \( v \) is regarded as an operator that acts on everything to its right. Thus, \( v f \) is certainly not locally linear in the function \( f \) because \( v \) will differentiate anything that \( f \) is multiplied by. On the right side of the equation, \( v \) is regarded as the argument of the linear form \( df \) and is not an operator at all. Thus, for a function, \( g \) we have

\[ v \cdot gdf = gv \cdot df \]

so that the right side of the equation is locally linear in the form \( df \).

2 Connections

2.1 Locally Linear Derivatives

Given any type of field \( w \) on a manifold that can be multiplied by a scalar field, consider operations that give back the same kind of field and obey the algebraic rules

\[
D_v (f w) = (D_v f) w + f D_v w
\]

and

\[
D_{gv} w = g D_v w
\]

for any functions \( f \) and \( g \). This type of operation is called a connection. By letting us differentiate a field \( w \) it introduces a way to compare or connect the values of \( w \) at neighboring points of the manifold.

When the operation \( D_v \) acts on a function, it is supposed to give back a function. If we multiply the field \( w \) by two different functions \( f_1 \) and \( f_2 \) in succession and compare the relations that must hold,

\[
D_v (f_1 f_2 w) = (D_v f_1 f_2) w + f_1 f_2 D_v w
\]
\[
D_v (f_1 f_2 w) = (D_v f_1) f_2 w + f_1 D_v (f_2 w)
\]
\[
= (D_v f_1) f_2 w + (f_1 D_v (f_2)) w + f_1 f_2 D_v w
\]

we see that Leibnis's rule must hold for \( D_v \) acting on functions also

\[
D_v f_1 f_2 = (D_v f_1) f_2 + f_1 D_v f_2.
\]
Since the vector field \( v \) is already identified with a derivative that acts on functions, we can always identify \( D_v f \) as \( v f \). Thus, we have, for any vector field \( v \) and function \( f \)

\[ D_v f = v f = L_v f. \]

It is quite common to distinguish the special case of a covariant derivative acting on either a scalar field or a vector field by using the special notation

\[ \nabla_v w = D_v w \]

whenever \( w (P) \) is either a real number or an element of \( V_P \). Thus, the operator \( D_v \) acts on any sort of tensor while \( \nabla_v \) is restricted to vectors and scalars. One purpose of this notation is to distinguish the covariant derivative operations that are defined directly from those that must be constructed.

### 2.2 The Connection Coefficients

Start with a set of basis vectors \( E_K (P) \) at each point of a manifold. These basis vectors span a vector space \( V_P \) at each point which could be the tangent space at that point, or it might be something else. The algebraic properties of the covariant derivative let us evaluate the derivative of an arbitrary field \( u \) by just expanding it in terms of the basis vectors.

\[ u = u^K E_K. \]

Since the derivative depends on a vector field, we also need the actual tangent space basis vectors \( e_k (P) \) that span the tangent space \( T_P \) at each point.

\[ \nabla_v u = \nabla_{(v^d e_d)} \left( u^K E_K \right). \]

From local linearity we can pull out the functions \( v^d \).

\[ \nabla_v u = v^d \nabla_{e_d} \left( u^K E_K \right). \]

Now use Leibniz’s product rule to get

\[
\begin{align*}
\nabla_v u & = v^d \nabla_{e_d} \left( u^K E_K \right) \\
& = v^d \left( \nabla_{e_d} (u^K) \right) E_K + v^d u^K \nabla_{e_d} (E_K) \\
& = \left( v^d e_d (u^K) \right) E_K + v^d u^K \nabla_{e_d} (E_K).
\end{align*}
\]

The only things that we cannot evaluate in this expression are the derivatives of the basis vectors

\[ \Gamma_{Kd} = \nabla_{e_d} (E_K). \]

These objects are supposed to be in the vector space \( V_P \), so we can expand them in the basis vectors

\[ \Gamma_{Kd} = E_J \Gamma^J_{Kd}. \]
or, putting the definition together,
\[ \nabla_{e_d} (E_K) = E_J \Gamma_{Kd}^J, \quad \Gamma_{Kd}^J = W^J (\nabla_{e_d} (E_K)) \]
where the \( W^J \) are the forms dual to the basis vectors \( E_K \).

The expansion coefficients \( \Gamma_{Kd}^J (P) \) at each point are all we need to evaluate the covariant derivative \( \nabla_v w \) for any vector field \( v \) and any field \( u \) that assigns an object \( w (P) \) in \( V_P \) to each point \( P \).

\[
\nabla_v u = E_K v^d e_d (u^K) + E_J \Gamma_{Kd}^J v^d u^K
= E_K v^d (\mathcal{L}_v u^K + \Gamma^K_{Jd} u^J)
= E_K (\mathcal{L}_v u^K + \Gamma^K_{Jd} u^J v^d)
\]

These coefficients are called the connection coefficients. Notice the convention that the index associated with the differentiating vector field \( v \) is placed last.

The components of a covariant derivative are traditionally written using a semicolon to set off the ‘differentiating index’ in the following way:
\[
\nabla_v u = E_K u^K_{;d} v^d
\]
where, from the earlier version of the expression for \( D_v u \)
\[
u^K_{;d} = e_d (u^K) + \Gamma^K_{Jd} u^J.
\]
It is traditional to express the differentiating action of the basis vectors by commas as follows:
\[ e_d (u^K) = u^K_{,d} \]
The traditional expression for the components of the covariant derivative is then
\[ u^K_{,d} = u^K_{,d} + \Gamma^K_{Jd} u^J. \]

### 2.3 Matrix Notation for the Connection

Matrix notation is often used to take care of the indexes that are associated with the space \( V_P \). In that notation, we would define matrices as follows:

\[
[E] = \begin{bmatrix}
E_1 \\
E_2 \\
\vdots \\
E_N
\end{bmatrix}, \quad [u] = \begin{bmatrix} u & u^2 & \cdots & u^N \end{bmatrix}
\]

so that
\[ u = [u] [E] \]
and
\[
\nabla_v u = \nabla_v ([u] [E]) = (\nabla_v [u]) [E] + [u] \nabla_v [E]
= \mathcal{L}_v [u] + [u] v^d [\Gamma_d]
\]
Notice that the matrix $v^d [\Gamma_d]$ is actually a matrix-valued linear function of the vector $v$ while $\mathcal{L}_v [u]$ is just the same as $v \cdot d [u]$. Thus it is usual in this notation to define the matrix-valued one-form field

$$[\omega] = \omega^d [\Gamma_d]$$

where the $\omega^d$ are the spacetime basis forms at the point $P$. The covariant derivative of the vector $u$ then takes the form

$$\nabla_v u = v \cdot D [u]$$

and we obtain the wonderfully simple relation:

$$D [u] = d [u] + [u] [\omega] .$$

3 Covariant Derivatives of Tensors

3.1 Extending Leibniz’s Product Rule

Once a connection has been defined for the spaces $V_P$, it is also defined for all of the tensor spaces that are related to $V_P$ by imposing Leibniz’s product rule on every type of product. For example, to differentiate a field that assigns the second rank contravariant tensor $T(P)$ in $V_P \otimes V_P$ to each point $P$ expand the tensor as a sum of tensor products of basis vectors

$$T = T^{AB} E_B \otimes E_A$$

and use the product rule on this expression:

$$D_v T = (D_v T^{AB}) E_B \otimes E_A + T^{AB} (D_v E_B) \otimes E_A + T^{AB} E_B \otimes D_v E_A$$

$$= v^d (e_d T^{AB}) E_B \otimes E_A + T^{AB} (\nabla_v E_B) \otimes E_A + T^{AB} E_B \otimes \nabla_v E_A$$

$$= v^d (e_d T^{AB}) E_B \otimes E_A + T^{AJ} (v^d \Gamma^K_{jd}) E_K \otimes E_A + T^{JB} E_B \otimes E_K v^d \Gamma^K_{jd}$$

$$= v^d (e_d T^{AB} + T^{AJ} \Gamma^B_{jd} + T^{JB} \Gamma^A_{jd})$$

or, in index notation,

$$D_v T = T^{AB} :_d v^d$$

where

$$T^{AB} :_d = e_d T^{AB} + T^{AJ} \Gamma^B_{jd} + T^{JB} \Gamma^A_{jd} .$$

3.2 Differentiating Form Fields

To differentiate a field that assigns to each point $P$ a form $\alpha (P)$ in the space $\tilde{V}_P$ dual to $V_P$, assume that Leibniz’s rule applies to the dot product of a form-field with a vector-field:

$$D_v (\alpha \cdot u) = D_v (\alpha) \cdot u + \alpha \cdot D_v (u) .$$
Since \(\alpha \cdot u\) is simply a scalar function on the manifold, we know how to differentiate it.

\[
v (\alpha \cdot u) = D_v (\alpha) \cdot u + \alpha \cdot \nabla_v (u) .
\]

Once this version of the product rule has been imposed, apply it to the definition of the dual basis forms

\[
W^A \cdot E_B = \delta ^A _B
\]

and obtain

\[
D_v (W^A) \cdot E_B + W^A \cdot \nabla_v E_B = 0
\]

or

\[
D_v (W^A) \cdot E_B = -W^A \cdot \nabla_v E_B = -W^A \cdot v^d \Gamma ^J _{Bd} E_J = -v^d \Gamma ^A _{Bd}
\]

and thus,

\[
D_v (W^A) = -v^d \Gamma ^A _{Bd} W^B .
\]

Now we can differentiate any form-field by using Leibniz’s rule.

\[
D_v \alpha = D_v (\alpha_A W^A) = \nabla_v (\alpha_A) W^A + \alpha_A D_v W^A
\]

\[
= v^d e^d \alpha_A W^A - \alpha_A v^d \Gamma ^A _{Jd} W^J
\]

\[
= v^d (e^d \alpha_A - \alpha_K \Gamma ^K _{Ad}) W^A
\]

### 3.3 Matrix Notation for Form-Fields

In terms of matrix notation, a one-form field \(\alpha\) would have a column-matrix of components \([\alpha]\) while a vector field \(u\) would have a row matrix of components \([u]\) with the relations

\[
\alpha = [W] [\alpha], \quad v = [u] [E]
\]

\[
u \cdot \alpha = [u] [E] \cdot [W] [\alpha] = [u] ([E] \cdot [W]) [\alpha] = [u] [1] [\alpha] = [u] [\alpha]
\]

so that

\[
D_v (\alpha \cdot u) = v (\alpha \cdot u) = v \cdot d (\alpha \cdot u)
\]

becomes

\[
D_v ([u] [\alpha]) = v \cdot d ([u] [\alpha]) .
\]

Now impose Leibniz’s rule on the matrix product and obtain

\[
(D_v [u]) [\alpha] + [u] D_v [\alpha] = (v \cdot d [u]) [\alpha] + [u] (v \cdot d [\alpha])
\]

Since we know that

\[
D_v [u] = v \cdot (d [u] + [u] [\omega])
\]

the requirement becomes

\[
(v \cdot (d [u] + [u] [\omega])) [\alpha] + [u] D_v [\alpha] = (v \cdot d [u]) [\alpha] + [u] (v \cdot d [\alpha])
\]
\[ (v \cdot d[u]) [\alpha] + (v \cdot [u] [\omega]) [\alpha] + [u] D_v [\alpha] = v \cdot d[u] [\alpha] + [u] (v \cdot d[\alpha]). \]

and some terms cancel yielding
\[ [u] D_v [\alpha] = v \cdot d[\alpha] [u] - v \cdot [u] [\omega] [\alpha] = [u] v \cdot (d[\alpha] - [\omega] [\alpha]) \]
or, with the same sort of definition that we used for vectors,
\[ D_v [\alpha] = v \cdot D[\alpha] \]
we obtain another wonderfully simple matrix expression:
\[ D[\alpha] = d[\alpha] - [\omega] [\alpha]. \]

3.4 Extending Leibniz’s Product Rule Again

Now that we know how to differentiate the basis forms \( W^B \), we can again use Leibniz’s rule to take derivatives of mixed-rank tensor fields. For example, we find that for \( t = t^A_B E_A \otimes W^B \)
the derivative has components
\[ t^A_{B;d} = e_d \left( t^A_B \right) + t^J_B \Gamma^A_{Jd} - t^A_J \Gamma^J_{Bd}. \]
With this procedure, we could go on to differentiate any tensor field at all. However, it is better to pause and consider what we are doing.

It is important to understand that we are not, in any sense, proving that Leibniz’s rule works for dot products and tensor products. Instead, we are imposing the rule and using it to extend the definition of the covariant derivative. This procedure is allowed only if it leads to a consistent result. We are on mathematically shaky ground here and should back up and secure the foundations before proceeding any further.

3.5 Extending Locally Linear Derivatives Directly

A procedure that leads to the same result as using Leibniz’s rule for tensor products and dot products has the advantage of constructing covariant derivatives directly without any need to worry about consistency. Given a tensor \( T \) that assigns the real number \( T(P, a_1(P), a_2(P), ..., a_q(P)) \) to the tensors \( a_1(P), a_2(P), ..., a_q(P) \) at the point \( P \), we define a new tensor \( D_v T \) that assigns the number
\[
(D_v T)(P, a_1(P), a_2(P), ..., a_q(P)) = D_v (T(P, a_1(P), a_2(P), ..., a_q(P))) \\
- T(P, D_v a_1(P), a_2(P), ..., a_q(P)) \\
- T(P, a_1(P), D_v a_2(P), ..., a_q(P)) \\
:: \\
- T(P, a_1(P), a_2(P), ..., D_v a_q(P))
\]
In words, we subtract a counter-term for each argument of the tensor. Each counter-term consists of the derivative operation acting on just one of the arguments. It is a straightforward matter to show that this direct construction procedure leads to an expression that is locally linear in each of its tensor arguments.

With this procedure for constructing covariant derivatives, we can start with $D_v$ acting on a vector field and define its action on a 1-form-field $\alpha$ as

$$(D_v \alpha)(u) = \nabla_v (\alpha(u)) - \alpha(\nabla_v u).$$

Here, $\alpha(u)$ is a scalar function on the manifold, so $\nabla_v (\alpha(u)) = v(\alpha(u))$ and we already know how to evaluate $D_v u$. This construction will give the same result we got before.

### 3.6 Higher Covariant Derivatives

Suppose that one wants to differentiate a vector field $w$ first along the vector field $v$ and then along the vector field $u$. Confusingly, there are two common ways to do this. The simplest procedure is to obtain a vector field $D_v w$ and then take the covariant derivative of that vector field, $D_u (D_v w)$. When this procedure is intended, we use the specialized notation $\nabla$ for $D$ in order to emphasize that all of the actions are on vector fields. The difficulty with the resulting operation is that it is not locally linear in the differentiating field $v$ in the way that $D_v w$ is.

For an arbitrary function, $f$ we have

$$\nabla_u (\nabla_{fv} w) = \nabla_u (f \nabla_v w) = (\nabla_u f)(\nabla_v w) + f \nabla_u (\nabla_v w)$$

If we want to have a second derivative that is locally linear in both of the differentiating vector fields, then the procedure introduced for generating such derivatives suggests that we need to regard the vector field $v$ as an argument of a tensor

$$Dw(v) = D_v w$$

and then construct the covariant derivative of this tensor field.

$$(D_u Dw)(v) = D_u (Dw(v)) - Dw (D_u v)$$

$$= \nabla_u \nabla_v w - \nabla_{\nabla_v w} w$$

Notice that this construction requires the covariant derivative of a vector field $v$ that assigns an element $v(P)$ of the tangent space $T_P$ to each point $P$. That covariant derivative is specified by the tangent-space connection coefficients

$$\Gamma^j_{kd} = \omega^j (\nabla_e \epsilon_k)$$

and may be quite independent of the connection that links the vectors spaces $V_P$. 

8
The resulting second covariant derivative is again locally linear since, for any scalar function \( f \)
\[
\nabla^{\nabla_{u} f}w = \nabla f \nabla_{u} w + \nabla_{(\nabla_{u} f)}w = f \nabla\nabla_{u} w + (\nabla_{u} f) \nabla_{v} w
\]
so that

\[
(D_{u}D_{w})(fv) = f \nabla u \nabla_{v} w - \nabla_{u}v \nabla_{w}
\]

In order to continue on to still higher degree covariant derivatives, regard the differentiating vector field \( u \) in \( D_{u}D_{w} \) as itself a tensor argument and define

\[
(D^{2}w)(v, u) = \nabla_{u} \nabla_{v} w - \nabla_{u}v \nabla_{w}
\]

and the covariant derivative of this tensor field with respect to a vector field \( t \) would be

\[
(D_{t}D^{2}w)(v, u) = (D^{3}w)(v, u, t) = \nabla_{t} ((D^{2}w)(v, u)) - (D^{2}w)(\nabla_{t}v, u) - (D^{2}w)(v, \nabla_{t}u)
\]

Here only derivatives of a vector field \( w \) have been discussed. In a similar way, one can define higher order covariant derivatives of tensors of arbitrary rank.

### 3.7 The Hessian

A useful example of a higher derivative is the second covariant derivative operator

\[
\text{Hess}(v, u) = D_{u}(D_{v}f) = (D^{2}f)(v, u)
\]

which is sometimes called the *Hessian tensor* of the function \( f \). Here, \( D_{v}f = vf = v \cdot df \). The covariant derivative \( D_{u} \) is acting on a form-field that has the tangent vector field \( v \) as its argument. Thus, the tangent vector connection enters the expression when we correct for the behavior of \( v \).

\[
\text{Hess}(v, u) = D_{u}(D_{v}f) = u (\nabla_{v}f) - D_{v} \nabla_{w}f = uvf - \nabla_{w}vf = uvf - (\nabla_{w}v) \cdot df
\]

The components of the Hessian in a frame-field \( \{e_{a}\} \) would then be

\[
\text{Hess}(e_{a}, e_{b}) = e_{b}(e_{a}f) - (\nabla_{e_{a}}e_{b}) \cdot df = e_{b}(e_{a}f) - (\nabla_{e_{a}}e_{b}) \cdot \omega^{j}(e_{j}f) = e_{b}(e_{a}f) - \Gamma^{j}_{ab}(e_{j}f).
\]
Example 1 Consider a geometrical statement of Newton’s Theory of Gravitation. The path of an object is described by a parameterized curve in $\mathbb{R}^3$ with time as the parameter. The tangent vector $\frac{d}{dt}$ to the curve is just the velocity vector $v$. The acceleration vector is the covariant derivative along the curve:

$$ a = \frac{dv}{dt} = D_v v. $$

In this case, the covariant derivative is very simple to evaluate in cartesian coordinates because all of the connection coefficients are zero and the components of $D_v v$ are just the time derivatives of the components of the velocity vector $v$.

$$ a^i = \frac{dv^i}{dt}. $$

If the acceleration is due to a gravitational field, then there is a Newtonian gravitational potential $\phi$ – a function on the manifold whose gradient yields the desired acceleration vector:

$$ a = -\nabla \phi = -g^{-1}(d\phi) \quad \text{or} \quad D_v v = -g^{-1}(d\phi) $$

$$ q \cdot D_v v = g(D_v v, q) = -q \cdot d\phi = -D_q \phi $$

The acceleration in the direction of a unit vector $\hat{n}$ is then given by the dot product

$$ \hat{n} \cdot D_v v = g(D_v v, \hat{n}) = -D_{\hat{n}} \phi $$

$$ \hat{n} \cdot a = -D_{\hat{n}} \phi $$

Now suppose that there is a family of objects, each labeled by a parameter $s$. The velocity vector is then $v = \frac{\partial}{\partial t}$ and we have a new vector $w = \frac{\partial}{\partial s}$ which is proportional to the separation vector between nearby particles. The velocity of a particle labeled $s$ relative to a neighbor labeled $s + \delta s$ is then $\delta v(w) = \delta s D_w v$ and the acceleration of each particle relative to its neighbors is

$$ \frac{d}{dt} \delta v(w) = \delta s \frac{d}{dt} D_w v = \delta s D_v D_w v. $$

The relative acceleration in the direction of a unit vector $\hat{n}$ is then given by the dot product

$$ \hat{n} \cdot \frac{d}{dt} \delta v(w) = \delta s \hat{n} \cdot \frac{d}{dt} D_w v = \delta s g(D_v D_w v, \hat{n}). $$

Because the connection coefficients all vanish in Cartesian coordinates, we can switch the order of differentiation

$$ D_v D_w v = D_w D_v v. $$

Because the metric tensor components are all constants in Cartesian coordinates, we can move derivatives outside the metric function:

$$ g(D_v D_w v, \hat{n}) = D_w g(D_v v, \hat{n}) = -D_{\hat{n}} D_w \phi = -\text{Hess}\phi(\hat{n}, w) $$
so the relative acceleration of nearby particles in the direction \( \hat{n} \) is given by the Hessian of the gravitational potential

\[
\hat{n} \cdot \frac{d}{dt} D_w v = -\text{Hess}\phi (\hat{n}, w).
\]

\[
\frac{d}{dt} (\delta v (w)) = -\delta s \text{Hess}\phi (\hat{n}, w).
\]

In this same language, the inverse square law of Newton’s Theory is contained in the equation satisfied by the potential \( \phi \)

\[
g^{ab} D_{e^a} D_{e^b} \phi = 0
g'^{ab} D_{e^a} D_{e^b} \phi = 0
D_{g^{-1}(\omega^a)} D_{e^a} \phi = 0
\]

\[\text{Hess}\phi (g^{-1}(\omega^a), e_a) = 0\]

which implies a restriction on the relative accelerations of nearby particles:

\[
g^{-1}(\omega^a) \cdot \frac{d}{dt} D_{e^a} v = 0
\omega^a \cdot \frac{d}{dt} D_{e^a} v = 0
\omega^a \cdot \delta v (e_a) = 0.
\]

4 The Structure Tensors

4.1 The Basic Idea of Differential Geometry

Covariant derivatives do not commute. In general, for a function \( f \), \( D_u (D_v f) \) is not the same as \( D_v (D_u f) \). Similarly, for a vector-field \( w \), \( \nabla_u \nabla_v w \) is not the same as \( \nabla_v \nabla_u w \). Thus, it is natural to ask for the commutators \([D_u, D_v] f\) and \([\nabla_u, \nabla_v] w\). The primary result of differential geometry is that these commutators can be expressed in terms of tensor fields – locally linear functions. Since the action of covariant derivatives on tensor fields is defined, the result is a completely closed algebra.

4.2 The Torsion Tensor

The commutator of covariant derivatives acting on a scalar is evidently just the antisymmetric part of the Hessian operator discussed in the preceding section.

\[
[D_u, D_v] f = \text{Hess} f (v, u) - \text{Hess} f (u, v)
= uvf - (\nabla_u v) \cdot df - vu f + (\nabla_v u) \cdot df
= [u, v] f + (\nabla_v u - \nabla_u v) \cdot df
= [u, v] f + (\nabla_v u - \nabla_u v) f
\]
From this expression, for each choice of vectors \( u, v \) the operator
\[
T(u, v) = -[D_u, D_v] = \nabla_u v - \nabla_v u - [u, v]
\]
is actually a directional derivative or vector. Since we have
\[
T(u, v) f = \text{Hess}(v, u) - \text{Hess}(u, v)
\]
the quantity is manifestly a tensor that assigns a vector to each pair of vectors \( u, v \). This tensor is a property of the tangent-space connection alone and is called the torsion. Its components in a given frame-field are:
\[
T^c_{\ ab} = \omega^c \cdot T(e_a, e_b) = \omega^c \cdot \nabla_{e_a} e_b - \omega^c \cdot \nabla_{e_b} e_a - \omega^c \cdot [e_a, e_b]
\]
\[
T^c_{\ ba} = \Gamma^c_{\ ba} - \Gamma^c_{\ ab} - c^c_{\ ab}.
\]
Thus, the torsion is essentially the antisymmetric part of the connection.

One way to see what the torsion tensor means is to see what it prevents one from doing. Because the torsion is a tensor field, there is no way to get rid of it unless it happens to be equal to zero everywhere. In physics, we would like to be able to construct coordinates that act like Minkowski coordinates near a particular spacetime point. We would also like the covariant derivative of a vector or tensor field near that particular point to become just the ordinary partial derivative of the field components. A non-zero torsion tensor field at that spacetime point would mean that we cannot succeed because partial derivatives commute and the commutator of covariant derivatives is the torsion, so covariant derivatives could not commute. In Einstein’s Theory of General Relativity, one uses this argument as a justification for setting the torsion tensor equal to zero everywhere.

### 4.3 The Curvature Tensor

The commutator of covariant derivatives acting on a vector field \( w \) is just \([\nabla_u, \nabla_v] w\). However, it is easy to see that this quantity in not locally linear in any of the vector fields \( u, v, w \). Thus, it is not a tensor field. However, the quantity
\[
\mathcal{R}(u, v) w = [\nabla_u, \nabla_v] w - \nabla_{[u,v]} w
\]
is locally linear as one can easily check. For any function, \( f \)
\[
\mathcal{R}(u, v)(f w) = [\nabla_u, \nabla_v] (f w) - \nabla_{[u,v]} f (w)
\]
\[
= ([u, v] f) w + f [\nabla_u, \nabla_v] w - (\nabla_{[u,v]} f) w - f \nabla_{[u,v]} (w)
\]
\[
= f \mathcal{R}(u, v)(w)
\]
\[
\mathcal{R}(f u, v)(w) = [\nabla_{fu}, \nabla_v] w - \nabla_{[fu,v]} w
\]
\[
= [f \nabla_u, \nabla_v] w - f \nabla_{[u,v]} w
\]
\[
= f \nabla_u \nabla_v w - \nabla_v f \nabla_u w - f \nabla_{[u,v]} w
\]
\[
= f \mathcal{R}(u, v)(w) - (\nabla_v f) \nabla_u w + (vf) \nabla_u w
\]
\[
= f \mathcal{R}(u, v)(w) + (vf) \nabla_u w
\]
\[
= f \mathcal{R}(u, v)(w).
\]
Thus, we have a tensor $\mathcal{R}$ which assigns to any pair of tangent-space vectors $u, v$ a tensor $\mathcal{R}(u, v)$ in $V_P \otimes V_P$ which, in turn, assigns to each vector $w$ in $V_P$, the vector $\mathcal{R}(u, v)w$ in $V_P$. This elaborate structure is called the curvature tensor.

The components of the curvature tensor can be found by inserting the appropriate sort of basis vectors, and basis forms. The Ricci Identity

$R^A_{Bab} = \omega^A \cdot \mathcal{R}(e_a, e_b) E_B$

$\mathcal{R}(e_a, e_b)E_B = [\nabla e_a, \nabla e_b] 0 = D_wT(v, u) f + D_vT(u, w) f + D_aT(w, v) f E_B - \nabla_{e_a e_b} E_B$

$= \nabla e_a (\nabla e_b E_B) - \nabla e_b (\nabla e_a E_B) - c_{ab} e_k \nabla e_k E_B$

$= \nabla e_a (E_K \Gamma^K_{Bb}) - \nabla e_b (E_K \Gamma^K_{Ba}) - c_{ab} \nabla e_k E_B$

$= E_K \nabla e_a \Gamma^K_{Bb} + (\nabla e_a E_K) \Gamma^K_{Bb}$

Here, connection coefficients such as $e_a \Gamma^K_{Bb}$ are being regarded as just functions, so the derivatives $\nabla e_a$ are really just the operators $e_a$. It is useful to absorb some of these indexes into a matrix notation. In particular, the indexes associated with the space $V_P$ can be absorbed to give the expression

$[\mathcal{R}](e_a, e_b) = e_a [\Gamma_b] - e_b [\Gamma_a] + [\Gamma_a][\Gamma_b] - [\Gamma_b][\Gamma_a] - c^{\cdot \cdot}_{ab} [\Gamma_k]$

or equivalently

$[\mathcal{R}](e_a, e_b) = e_a \cdot d[\Gamma_b] - e_b \cdot d[\Gamma_a] + [\Gamma_a][\Gamma_b] - [\Gamma_b][\Gamma_a] - [e_a, e_b] \cdot \omega^k [\Gamma_k]$. 

5 Curvature Identities

5.1 The Ricci Identity

Why is the curvature tensor defined in terms of the non-tensorial second derivative $\nabla_u \nabla_v w$? It would seem to be better to start with the tensorial second derivative of a vector field $D_u D_v w$. In that case one would not need an extra term to satisfy local linearity and the resulting curvature definition would be $\mathcal{R}'(u, v) w = [D_u, D_v] w$. 

13
The difference between the two approaches is that this one includes compensating terms in which the derivative acts on one of the differentiating vector fields.

\[ D_u D_v w = \nabla_u \nabla_v w - \nabla_{\nabla_v} w \]

so that

\[ \mathcal{R}'(u, v) w = [\nabla_u, \nabla_v] w - \nabla_{\nabla_v} w \]

Thus, the two definitions differ by a term that involves the torsion tensor

\[ \mathcal{R}'(u, v) - \mathcal{R}(u, v) = \nabla_{[u,v]} - (\nabla_{\nabla_v} - \nabla_{\nabla_u}) = -\nabla_{T(u,v)}. \]

The commutator of covariant derivatives is related to the curvature tensor by

\[ [D_u, D_v] w = \mathcal{R}(u, v) w - \nabla_{T(u,v)} w. \]

This expression is usually called the Ricci Identity. The definition of \( \mathcal{R}(u, v) \) depends only on the connection on the vector spaces \( V_P \) at each point. If those vector spaces are not the tangent spaces, then this definition of the curvature is completely independent of the tangent-space connection. The commutator of covariant derivatives, however, does depend on the tangent-space connection through the torsion tensor.

The Ricci identity can be used to find the commutator of covariant derivatives acting on all sorts of tensor fields. For a form-field \( \alpha \) the form-field \([D_u, D_v] \alpha \) evaluated for particular vector field \( w \) is obtained from the definition of the covariant derivative of a form-field. One way to proceed is to start with the definition of the torsion acting on a scalar field \( \alpha \cdot w \)

\[ [D_u, D_v] (\alpha \cdot w) = T(v, u) (\alpha \cdot w) \]

and then use Leibniz’s product rule for the dot product.

\[
D_u D_v (\alpha \cdot w) = D_u ((D_v \alpha) \cdot w) + D_u (\alpha \cdot D_v w) \\
= (D_u D_v \alpha) \cdot w + (D_v \alpha) \cdot (D_u w) + (D_u \alpha) \cdot D_v w + \alpha \cdot D_u D_v w
\]

\[
[D_u, D_v] (\alpha \cdot w) = D_u D_v (\alpha \cdot w) - D_v D_u (\alpha \cdot w) \\
= (D_u D_v \alpha) \cdot w + (D_v \alpha) \cdot (D_u w) + (D_u \alpha) \cdot D_v w + \alpha \cdot D_u D_v w - (D_v D_u \alpha) \cdot w - (D_u \alpha) \cdot (D_v w) - (D_v \alpha) \cdot D_u w - \alpha \cdot D_u D_v w \\
= ([D_u, D_v] \alpha) \cdot w + \alpha \cdot [D_u, D_v] w
\]

The expression that we want is then

\[
([D_u, D_v] \alpha) \cdot w = T(v, u) (\alpha \cdot w) - \alpha \cdot [D_u, D_v] w \\
= T(v, u) (\alpha \cdot w) - \alpha \cdot \mathcal{R}(u, v) w + \alpha \cdot \nabla_{T(u,v)} w \\
([D_u, D_v] \alpha) \cdot w = -\alpha (\mathcal{R}(u, v) w - \nabla_{T(u,v)} w) + T(v, u) (\alpha \cdot w)
\]
When the torsion vanishes, the Ricci identities become far simpler, taking the forms

\[ [D_u, D_v] w = \mathcal{R}(u, v) w \]
\[ ([D_u, D_v] \alpha)(w) = -\alpha (\mathcal{R}(u, v) w) \]

or, in components,

\[ w^A ;_b - w^A ;_a = R^A_{Sab} w^S \]
\[ \alpha_{B;b} - \alpha_{B;a} = -\alpha_S R^S_{Bab}. \]

For a general tensor \( M \) with both form and vector arguments, the expression for the commutator is simplest in terms of components. For example,

\[ M_{AB}^{CD} ;_b - M_{AB}^{CD} ;_a = R^A_{Sab} M_{SB}^{CD} + R^B_{Sab} M_{AS}^{CD} - M_{SD}^{AB} R^S_{Cab} - M_{CS}^{AB} R^S_{Dab} \]

### 5.2 Torsion Bianchi Identities

The commutators of any associative algebra obey Jacobi’s identity

\[ [A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0. \]

Thus, the commutators of covariant derivatives obey this identity. For commutators of covariant derivatives acting on a scalar field one has

\[ 0 = [D_w, [D_u, D_v]] f + [D_v, [D_w, D_u]] f + [D_u, [D_v, D_w]] f \]

which becomes

\[ 0 = D_w [D_u, D_v] f + D_v [D_w, D_u] f + D_u [D_v, D_w] f - [D_u, D_v] D_w f - [D_w, D_u] D_v f - [D_v, D_w] D_u f \]

The positive terms in this expression are derivatives of the torsion tensor. The terms with negative signs are commutators acting on forms and require the Ricci identity.

\[ [D_u, D_v] D_w f = w \cdot [D_u, D_v] df \]
\[ = -df (\mathcal{R}(u, v) w - \nabla_{T(u,v)} w) + T(v, u) (df \cdot w) \]
\[ = - (\mathcal{R}(u, v) w - \nabla_{T(u,v)} w) f + T(v, u) (wf) \]
\[ = -\mathcal{R}(u, v) w f + \nabla_{T(u,v)}wf + \nabla_{T(v,u)}wf \]
\[ = -\mathcal{R}(u, v) w f \]

Thus, the identity becomes

\[ 0 = D_u T(v, u) f + D_v T(u, w) f + D_w T(w, v) f + \mathcal{R}(u, v) wf + \mathcal{R}(w, u) vf + \mathcal{R}(v, w) uf \]
This identity can be called the torsion Bianchi Identity. Notice that, when the torsion is zero, it simply makes a statement about the symmetry of the curvature tensor. Because the three vectors \( u, v, w \) are all in the tangent space, this identity applies only to the tangent-space curvature tensor.

### 5.3 Curvature Bianchi Identities

Now apply Jacobi’s identity to covariant derivatives of vector fields. It is convenient this time to use the vectorial derivative

\[
0 = D_t T(v, u) + D_u T(u, w) + D_w T(w, v) + \mathcal{R}(u, v) w + \mathcal{R}(w, u) v + \mathcal{R}(v, w) u
\]

or just

\[
0 = D_u T(v, u) + D_v T(u, w) + D_w T(w, v) + \mathcal{R}(u, v) w + \mathcal{R}(w, u) v + \mathcal{R}(v, w) u
\]

For the first terms we get

\[
\nabla_t [\nabla_u, \nabla_v] w = \nabla_t (\nabla_u [\nabla_v, \nabla_v] w)
\]

For the remaining terms,

\[
[\nabla_u, \nabla_v] \nabla_t w = (\mathcal{R}(u, v) + \nabla_{[u,v]} w) \nabla_t w
\]

and the terms combine in pairs to make

\[
\nabla_t [\nabla_u, \nabla_v] w - [\nabla_u, \nabla_v] \nabla_t w = \nabla_t \mathcal{R}(u, v) w - \mathcal{R}(u, v) \nabla_t w + \nabla_t [\nabla_v, \nabla_{[u,v]}] w - [\nabla_u, \nabla_v] \nabla_t
\]

Now express the vectorial derivative in terms of the covariant derivative

\[
D_t \mathcal{R}(u, v) w = \nabla_t \mathcal{R}(u, v) w - \mathcal{R}(u, \nabla_t v) w - \mathcal{R}(\nabla_t u, v) w - \mathcal{R}(\nabla_t u, \nabla_t v) w
\]

or

\[
\nabla_t \mathcal{R}(u, v) w - \mathcal{R}(u, v) \nabla_t w = D_t \mathcal{R}(u, v) w + \mathcal{R}(u, \nabla_t v) w + \mathcal{R}(\nabla_t u, v) w.
\]
Each pair of terms then gives
\[
\nabla_t [\nabla_u, \nabla_v] w - [\nabla_u, \nabla_v] \nabla_t w
\]
\[
= D_t \mathcal{R}(u,v) w + \mathcal{R}(u, \nabla_t v) w + \mathcal{R}(\nabla_t u, v) w + \mathcal{R}(t, [u,v]) w + \nabla_{[t,[u,v]]} w
\]
\[
= D_t \mathcal{R}(u,v) w + \mathcal{R}(u, \nabla_t v) w - \mathcal{R}(v, \nabla_t u) w + \mathcal{R}(t, [u,v]) w + \nabla_{[t,[u,v]]} w
\]

Now add these pairs up, cyclically permuting the arguments \( t, u, v \).
\[
0 = D_t \mathcal{R}(u,v) w + D_u \mathcal{R}(t,u) w + D_v \mathcal{R}(v,t) w
\]
\[
+ \mathcal{R}(u, \nabla_t v) w + \mathcal{R}(t, \nabla_u v) w + \mathcal{R}(v, \nabla_a t) w
\]
\[
- \mathcal{R}(v, \nabla_t u) w - \mathcal{R}(t, \nabla_v u) w - \mathcal{R}(u, \nabla_v t) w
\]
\[
+ \mathcal{R}(t, [u,v]) w + \mathcal{R}(v, [t,u]) w + \mathcal{R}(u, [v,t]) w
\]
\[
+ \nabla_{[t,[u,v]]} w + \nabla_{[v,[t,u]]} w + \nabla_{[u,[v,t]]} w
\]

The terms can be matched up as follows:
\[
0 = D_t \mathcal{R}(u,v) w + D_u \mathcal{R}(t,u) w + D_v \mathcal{R}(v,t) w
\]
\[
+ \mathcal{R}(u, \nabla_t v) w - \mathcal{R}(u, \nabla_v t) w + \mathcal{R}(u, [v,t]) w
\]
\[
+ \mathcal{R}(t, \nabla_u v) w - \mathcal{R}(t, \nabla_v u) w + \mathcal{R}(t, [u,v]) w
\]
\[
+ \mathcal{R}(v, \nabla_a t) w - \mathcal{R}(v, \nabla_a u) w - \mathcal{R}(v, \nabla_t u) w
\]
\[
+ \nabla_{[t,[u,v]]} w + \nabla_{[v,[t,u]]} w + \nabla_{[u,[v,t]]} w
\]
or
\[
0 = D_t \mathcal{R}(u,v) w + D_u \mathcal{R}(t,u) w + D_v \mathcal{R}(v,t) w
\]
\[
+ \mathcal{R}(u, [v,t]) w - \nabla_v t + \nabla_t v w
\]
\[
+ \mathcal{R}(t, [u,v]) w - \nabla_u v + \nabla_v u w
\]
\[
+ \mathcal{R}(v, [t,u]) w - \nabla_t u + \nabla_u t w
\]

Now recall the definition of the torsion tensor
\[
T(u,v) = \nabla_u v - \nabla_v u - [u,v]
\]
and obtain the final form of the curvature Bianchi Identities:
\[
0 = D_t \mathcal{R}(u,v) w + D_u \mathcal{R}(t,u) w + D_v \mathcal{R}(v,t) w
\]
\[
- \mathcal{R}(u, T(v,t)) w - \mathcal{R}(v, T(t,u)) w - \mathcal{R}(t, T(u,v)) w.
\]

In this case, the vectors \( u, v, t \) are all tangent-space vectors but the vector-field \( w \) that the identity is understood to be operating on can be a member of any sort of local vector space \( V_P \). Thus, the curvature Bianchi identities restrict all curvature tensors and not just the tangent-space curvature.
6 The Geodesic Deviation Equation

6.1 Geodesics

A curve is called a geodesic if its tangent vector $u$ obeys the equation

$$\nabla_u u = 0.$$ 

A geodesic is a curve whose tangent vector is covariantly constant. Thus it is a generalization of a straight line. Notice that only the connection is needed to define geodesics.

Now consider a one-parameter family of geodesics, each labeled by a parameter value $s$. A point on one of these curves can be located by giving the value of $s$ to specify which curve it is on and the value of the curve parameter $t$ along that curve. Assume that the family of curves is smooth so that the set of points on curves in the family is actually a two dimensional sub-manifold on which we can use $s$ and $t$ as coordinates. The vector-field $\frac{\partial}{\partial t}$ within this sub-manifold is just the tangent-vector $u$ and relates points at different parameter values along the same curve. The vector-field

$$w = \frac{\partial}{\partial s}$$

will be called the separation vector and relates points with the same parameter values on neighboring curves.
A measure of how neighboring geodesics differ from one another initially is the rate of change of their tangent vectors in the direction of the separation vector $\nabla_w u$. If all of the curves start off parallel, then this derivative will be zero. We would then like to know if the curves will remain parallel.

One measure of how rapidly neighboring geodesics depart from being parallel is just the derivative of the vector $\frac{du}{ds} = \nabla_w u$ in the direction of tangent vectors to the curves. Because $u$ and $w$ are just partial derivatives on a submanifold, their commutator vanishes and the definition of the curvature tensor yields

\[ [\nabla_u, \nabla_w] u = \mathcal{R}(u, w) u \]

or

\[ \nabla_u \nabla_w u = \nabla_w \nabla_u u + \mathcal{R}(u, w) u \]

Because the curve is a geodesic, $D_u u = 0$ and we get

\[ \frac{d}{dt} \left( \frac{du}{ds} \right) = \nabla_u \nabla_w u = \mathcal{R}(u, w) u. \]

This result is called the geodesic deviation equation. It relates the curvature components to the failure of initially parallel geodesics to remain parallel. Notice that the torsion does not enter into the equation at all. However, this form of the equation depends on identifying $\frac{du}{ds} = \nabla_w u$ as the object that measures the deviation.

### 6.2 Geodesic Deviation Equation Again

Use a local coordinate system to locate the points on a family of curves. The curve labeled by parameter value $s$ maps the curve-parameter value $t$ into a point with coordinates $x^a(s, t)$. These coordinates yield a set of coordinate basis vectors

\[ e_a = \frac{\partial}{\partial x^a} \]

and dual basis one-forms

\[ \omega^a = dx^a. \]

Now consider two nearby curves, one labeled $s$ and the other labeled $s + \delta s$. The separation between the curves can be described by the components

\[ \delta x^a = x^a(s + \delta s, t) - x^a(s, t) \approx \frac{\partial x^a}{\partial s} \delta s. \]

A vector can be constructed from these components by just using the basis vectors:

\[ \delta x = e_a \delta x^a e_a \approx \delta s \frac{\partial x^a}{\partial s} e_a \]

\[ \delta s \frac{\partial x^a}{\partial s} e_a = \delta s \frac{\partial x^a}{\partial s} \frac{\partial}{\partial x^s} = \delta s \frac{\partial}{\partial s} \]
Thus, the separation between neighboring curves can be expressed in terms of the vector

\[ w = \frac{\partial}{\partial s} \]

and is given by the separation vector:

\[ \delta x = (\delta s) w. \]

This time we will seek to find a system of differential equations that governs the evolution of this vector in the parameter \( t \).

The rate of change of the separation \( \delta x \) as \( t \) changes is given by the covariant derivative

\[
\frac{d}{dt} \delta x = D_u \delta x = D_u (\delta s) w
\]

\[
= (D_u (\delta s)) w + (\delta s) D_u w
\]

\[
= (\delta s) D_u w
\]

where \( u \) is the field of tangent vectors to the curves

\[ u = \frac{\partial}{\partial t} \]

and we are observing a pair of curves labeled by constant values of \( s \). Compare this quantity to the one that entered in to the geodesic deviation equation:

\[
\frac{d\delta x}{dt} = (\delta s) \nabla_w w
\]

\[
\frac{du}{ds} = \nabla_w u
\]

From the definition of torsion for the two commuting vector fields \( u, v \)

\[ \nabla_u w = \nabla_w u + T(u, w) \]

so we find the relationship

\[
\frac{d\delta x}{dt} = (\delta s) (\nabla_u w + T(u, w))
\]

or

\[
(\delta s) \frac{du}{ds} = \frac{d\delta x}{dt} + (\delta s) T(w, u)
\]

and the geodesic deviation equation becomes

\[
\frac{d}{dt} \left( \frac{d\delta x}{dt} + (\delta s) T(w, u) \right) = R(u, \delta x) u.
\]