1 Tangent Space Vectors and Tensors

1.1 Representations

At each point $P$ of a manifold $M$, there is a tangent space $T_P$ of vectors. Choosing a set of basis vectors $e_\alpha \in T_P$ provides a representation of each vector $u \in T_P$ in terms of components $u^\alpha$.

$$ u = u^\alpha e_\alpha = u^0 e_0 + u^1 e_1 + u^2 e_2 + ... = [u] [e] $$

where the last expression treats the basis vectors as a column matrix $[e]$ and the vector components as a row matrix $[u]$.

There is also a cotangent space $^T P$ of one-forms at each point. The basis one-forms which are dual to the basis vectors $e_\alpha$ are just the component-taking operators:

$$ \omega^\alpha (u) = \omega^\alpha \cdot u = u \cdot \omega^\alpha = u^\alpha $$

or, in terms of a row matrix $[\omega]$ of basis forms,

$$ u \cdot [\omega] = [u]. $$

When one of these operators acts on a basis vector, it produces one or zero according to:

$$ \omega^\alpha (e_\beta) = \delta^\alpha_\beta $$

or, in matrix notation,

$$ [e] \cdot [\omega] = [1] $$

An arbitrary one-form $\varphi$ can be represented by its components $\varphi_\beta$ in the form:

$$ \varphi = \varphi_\beta \omega^\beta = [\omega] [\varphi] $$

where the components are now arranged as a column matrix.

Here are various ways to represent a form acting on a vector to give a number:

$$ \varphi (u) = u (\varphi) = \varphi \cdot u = u \cdot \varphi = u^\alpha \varphi_\alpha $$

$$ = u^0 \varphi_0 + u^1 \varphi_1 + u^2 \varphi_2 + u^3 \varphi_3 + ... $$

$$ = [\varphi] [u] $$

1.2 Holonomic Basis

Given a coordinate patch — a set of independent functions $x^\alpha$ which can be used to label points in some part of $M$ — a special set of basis one-forms can be defined: $\omega^\alpha = dx^\alpha$

$$ dx^\alpha (u) = u (x^\alpha) $$

Recall that each tangent vector $u$ is really a directional derivative.

$$ u = \left. \frac{d}{d\lambda} \right|_C $$
where $C$ is a curve whose tangent vector at $P$ is $u$. Rewrite the definition of the $dx^{\alpha}$ in various ways:

$$dx^{\alpha} \cdot u = u \cdot dx^{\alpha} = dx^{\alpha}(u) = \frac{dx^{\alpha}(C(\lambda))}{d\lambda}$$

The basis vectors dual to the one-forms $dx^{\alpha}$ are just our old friends the partial derivatives:

$$\partial_{\alpha} = \frac{\partial}{\partial x^{\alpha}}$$

This type of basis is called a holonomic basis. For a vector $u$, here are various expressions in such a basis:

$$u = u^{\alpha} \partial_{\alpha}, \quad u^{\alpha} = dx^{\alpha} \cdot u$$

$$u^{\alpha} = u(x^{\alpha}) = \frac{dx^{\alpha}}{dx^\lambda} = \dot{x}^{\alpha}$$

$$u = \dot{x}^{\alpha} \partial_{\alpha} = \frac{dx^{\alpha}}{d\lambda} \frac{\partial}{\partial x^{\alpha}}$$

For a one-form $\varphi$ here are various expressions in a holonomic basis:

$$\varphi = \varphi_{\beta} dx^{\beta}, \quad \varphi_{\beta} = \varphi \cdot \partial_{\beta}$$

$$df \cdot \partial_{\beta} = \frac{\partial f}{\partial x^{\beta}}, \quad df = \frac{\partial f}{\partial x^{\beta}} dx^{\beta}$$

Some more expressions in a holonomic basis:

$$u \cdot \varphi = (u^{\alpha} \partial_{\alpha}) \cdot (\varphi_{\beta} dx^{\beta}) = u^{\alpha} \varphi_{\alpha}$$

$$u \cdot df = u^{\alpha} (df)_{\alpha} = \frac{dx^{\alpha}}{d\lambda} \frac{\partial f}{\partial x^{\alpha}}$$

$$u \cdot df = u(f) = \frac{df}{d\lambda}$$

1.3 n-beins, Vierbeins, Tetrads, etc.

Holonomic bases are not always appropriate. For example, local observers always want to use orthonormal basis vectors but it is not usually possible for a holonomic basis system to be orthonormal everywhere. However, if one has a curve $C$ which is defined in terms of coordinates and wants its tangent vector, then the holonomic basis is needed. Thus, non-holonomic basis systems are usually related to holonomic ones like this:

$$e_{\alpha} = b_{\alpha}^{\beta} \partial_{\beta}, \quad \omega^{\alpha} = dx^{\beta} b_{\beta}^{\alpha}$$

with the inverse relations:

$$\partial_{\alpha} = b_{\alpha}^{\beta} e_{\beta}, \quad dx^{\alpha} = \omega^{\alpha} b_{\beta}$$
Here, I am using Latin indexes for the non-holonomic basis and Greek indexes for the holonomic basis. Sometimes the two kinds of indexes are distinguished in other ways:

\[ e^{(\alpha)} = b^{(\alpha)} \partial \beta, \quad \omega^{(\alpha)} = dx^{\beta} b^{(\alpha)} \partial \beta \]

Underlines, primes, carets, and other doodads are used to mark one set of indexes as distinct from the other. Notice that the index positions are absolutely critical. If you are using Scientific Notebook or Scientific Workplace, you need to prevent superscripts and subscripts from lining up one over the other by inserting a zero space between them.

In terms of matrices, the relation between holonomic and non-holonomic representations would look like this:

\[
\begin{bmatrix}
| e \\
| \partial \\
| \omega \\
| dx
\end{bmatrix} =
\begin{bmatrix}
| b \\
| [\partial] \\
| [dx] \\
| [\omega]
\end{bmatrix}^{-1}
\]

Here are the non-holonomic components of the velocity of a particle which is located at coordinates \( x^i (t) \) at time \( t \):

\[
v = v^a \partial_a = v^a b_a^{(a)} e^{(a)} = v^{(a)} e^{(a)}
\]

\[
v^{(a)} = v^a b_a^{(a)}
\]

\[
v^a = \frac{dx^a}{dt} = \dot{x}^a
\]

\[
v^{(a)} = \dot{x}^a b_a^{(a)}
\]

In a four-dimensional spacetime, a non-holonomic basis system is often called a tetrad or a vierbein or a moving frame. The two types of indexes are not always integrated into a single formalism connected by tetrad coefficients \( b_a^{(a)} \). Sometimes we just perform a change of basis when it is convenient but use only one type of basis at a time.

### 1.4 Higher rank tensors

Vectors and one-forms are special cases of tensors. Higher rank tensors are multilinear functions of vectors and one-forms. They can be built up from tensor products defined like this:

\[ u \otimes v (\alpha, \beta) = u (\alpha) v (\beta), \quad \alpha \otimes \beta (u, v) = \alpha (u) \beta (v) \]

\[ u \otimes \alpha \otimes \varphi (\theta, u, v) = u (\theta) \alpha (u) \varphi (v) \]

The space of tensors that act on pairs of vectors can be built from the tensor products

\[ e_{\alpha \beta} = e_\alpha \otimes e_\beta \]

and is designated \( T_\alpha \otimes T_\beta \). Other tensors are built up similarly. An example of such a tensor would be

\[ T = T^{\alpha \beta} e_\alpha \otimes e_\beta \]
2 Connections

2.1 Derivatives

A connection is a way of differentiating vector and tensor fields. The essence of any type of “derivative” is that it obeys Leibniz’s product rule:

\[ D(AB) = D(A)B + AD(B) \]

as well as simple linearity

\[ D(B + C) = D(B) + D(C) \]

Evidently derivatives can be defined on any algebra which permits multiplications and additions. In more exotic cases, such as the algebra of square matrices for example, the linear Leibnizian operators are referred to as “derivations”. Consider the operator \( \partial_D \) defined on square matrices by

\[ \partial_D(A) = [A, D] = AD - DA \]

and show that it is a derivation.

2.2 Directional Derivatives

The simplest kind of derivative takes a tensor field into another tensor field of the same rank. Suppose that \( K_{\text{op}} \) is this type of derivative and work out how it must act on a vector field that has been expressed in terms of a particular moving frame \( \{e_\alpha\} \):

\[ K_{\text{op}}(w) = K_{\text{op}}(w^\alpha e_\alpha) \]

\[ = K_{\text{op}}(w^\alpha) e_\alpha + w^\alpha K_{\text{op}}(e_\alpha) \]

The problem is reduced to two parts (1) knowing how the operator acts on functions such as the components \( w^\alpha \) (2) knowing the vector that the operator assigns to each basis vector.

The first part tells us a lot about the nature of this operator because there is only one derivative operator on functions, the directional derivative, and each such operator corresponds to a vector field. Thus, our operator should be labeled by the vector field which it turns into when it acts on functions. If \( K_{\text{op}}(f) = \nu f \) then write \( K_{\text{op}} \) as \( D_v \). Now we have

\[ D_v w = v(w^\alpha) e_\alpha + w^\alpha D_v e_\alpha \]

That leaves part (2). We need to specify a vector-valued function \( \Gamma_\alpha (v) \) of a vector which maps \( v \) into \( D_v e_\alpha \) for each value of \( \alpha \). Then, (changing the last dummy index with malice aforethought)

\[ D_v w = v(w^\alpha) e_\alpha + w^\beta \Gamma_\beta (v) \]
Expanding the vector-valued function in terms of its components,
\[
\Gamma^\alpha_{\beta} (v) = e_\alpha \Gamma^\alpha_{\beta} (v)
\]
\[
D_vw = v(w^\alpha) e_\alpha + w^\beta e_\alpha \Gamma^\alpha_{\beta} (v)
\]
Collecting the results, we have
\[
D_vw = [v(w^\alpha) + w^\beta \Gamma^\alpha_{\beta} (v)] e_\alpha
\]
where the connection functions \( \Gamma^\alpha_{\beta} : T_p \rightarrow \mathbb{R} \) represent the action of the derivatives on basis vector fields
\[
D_v e_\beta = e_\alpha \Gamma^\alpha_{\beta} (v)
\]
or equivalently,
\[
\Gamma^\alpha_{\beta} (v) = (D_v e_\beta) \cdot \omega^\alpha.
\]
If you wish to push the matrix notation a bit farther, you can write this definition as
\[
[D_v w] = \{ v[w] + [w] \Gamma (v) \}.
\]

### 2.3 Affine Connection

The directional derivative operator \( D_v \) has a very nice property when it acts on functions: For any functions \( h \) and \( f \)
\[
D_h v f = hvf = hD_v f
\]
and, for any vectors \( u \) and \( v \),
\[
D_{u+v} f = (u+v)f = uf + vf = D_u f + D_v f
\]
Evidently we have the option of extending this nice property to the action of \( D_v \) on vector fields so that, in general, for any function \( h \) and vector field \( v \),
\[
D_{hv} = hD_v
\]
and, for any vectors \( u \) and \( v \),
\[
D_{u+v} = D_u + D_v
\]
When this choice is made, the result is called an affine directional derivative. For an affine directional derivative, the connection functions obey
\[
\Gamma^\alpha_{\beta} (hv) = \omega^\alpha \cdot D_{hv} e_\beta = h\omega^\alpha \cdot D_v e_\beta = h\Gamma^\alpha_{\beta} (v)
\]
as well as
\[ \Gamma^\alpha_\beta (u + v) = \Gamma^\alpha_\beta (u) + \Gamma^\alpha_\beta (v) . \]

For affine directional derivatives, the connection functions are locally linear functions on the vector space \( T_P \) — they are one-forms called the connection forms.

\[ \Gamma^\alpha_\beta (v) = \omega^\alpha_\beta \cdot v \]

In terms of components,
\[ \Gamma^\alpha_\beta (v) = \Gamma^\alpha_\beta \delta \]

or
\[ \omega^\alpha_\beta = \Gamma^\alpha_\beta \delta \omega^\delta \]

and some various ways of expressing the directional derivative of an arbitrary vector field \( w \) are:

\[ D_v w = \left[ v (w^\alpha) + w^\beta (\omega^\alpha_\beta \cdot v) \right] e_\alpha \]
\[ = \left[ v (w^\alpha) + w^\beta \Gamma^\alpha_\beta \delta \right] e_\alpha \]
\[ = \left[ v^\delta e_\delta (w^\alpha) + w^\beta \Gamma^\alpha_\beta \delta \right] e_\alpha \]
\[ = v^\delta \left[ e_\delta (w^\alpha) + w^\beta \Gamma^\alpha_\beta \delta \right] e_\alpha \]

By specifying the connection coefficients \( \Gamma^\alpha_\beta \delta \) in each of the overlapping coordinate charts of a manifold so that the resulting directional derivatives agree in the overlap regions, one defines an affine connection on the manifold.

### 2.4 Extending the definition using Leibniz’s rule

Given an affine connection defined on vector fields, one can extend it by insisting that it obey the product rule for every conceivable type of product: For a form-field \( \varphi \) and a vector-field \( w \), insist that

\[ D_v (\varphi \cdot w) = (D_v \varphi) \cdot w + \varphi \cdot D_v w \]

and find that

\[ (D_v \omega^\alpha) \cdot e_\beta = -\omega^\alpha_\beta \cdot v \]

and

\[ D_v \varphi = v^\delta \left[ e_\delta (\varphi) - \varphi e_\delta \Gamma^\alpha_\beta \right] \omega^\beta \]

In terms of matrices, the column of components of \( D_v \varphi \) would be

\[ [D_v \varphi] = \left\{ v [\varphi] - \Gamma^T (v) [\varphi] \right\} \]

For any tensor product, \( u \otimes w \), just do

\[ D_v (u \otimes w) = (D_v u) \otimes w + u \otimes D_v w. \]

Since any tensor in \( T_P \otimes T_P \) can be expanded in terms of such tensor products, we can now find the directional derivative of any such tensor. Try it on the tensor \( S = S^\alpha_\beta e_\alpha \otimes e_\beta \):

\[ D_v S = D_v \left( S^\alpha_\beta e_\alpha \otimes e_\beta \right) \]
\[ = (D_v S^\alpha_\beta) e_\alpha \otimes e_\beta + S^\alpha_\beta D_v (e_\rho) \otimes e_\beta + S^\alpha_\beta e_\alpha \otimes D_v e_\rho \]
where I have changed two dummy indexes to \( \rho \) to make my life easier later. Continue:

\[
D_v S = v^\delta e_\delta \left( S^\alpha_\beta \right) e_\alpha \otimes e_\beta + v^\delta S^\rho^\gamma_\alpha \rho^\beta_\delta e_\alpha \otimes e_\beta + v^\delta S^\gamma^\rho_\rho^\beta_\delta e_\alpha \otimes e_\beta \\
D_v S = v^\delta \left[ e_\delta \left( S^\alpha_\beta \right) + S^\rho^\gamma_\alpha \rho^\beta_\delta + S^\alpha^\rho_\gamma \Gamma^\beta_\rho \delta \right] e_\alpha \otimes e_\beta
\]

Try something more involved, \( Q = Q^\alpha_\gamma e_\alpha \otimes e_\beta \otimes \omega^\gamma \). The result is

\[
D_v Q = v^\delta \left[ e_\delta \left( Q^\alpha_\beta \gamma \right) + Q^\rho^\beta_\gamma \Gamma^\alpha_\rho \delta + Q^\alpha^\rho_\gamma \Gamma^\beta_\rho \delta - Q^\gamma^\rho_\rho^\beta_\delta \Gamma^\rho_\gamma \delta \right] e_\alpha \otimes e_\beta \otimes \omega^\gamma
\]

The pattern should be fairly clear by now: Each tensor index is contracted with one of the first two indexes on a connection coefficient with plus signs for the terms in which a superscript tensor index is contracted and minus signs for the terms in which a subscript tensor index is contracted.

### 2.5 Covariant Derivatives

When an affine directional derivative \( D_v \) acts on a vector-field \( w \), it produces another vector field \( D_v w \). Thus,

\[
D_v |_{P} \in T_P
\]

In terms of components, the vector \( D_v w \) has components

\[
(D_v w)^\alpha = v^\delta \left( e_\delta w^\alpha + w^\beta \Gamma^\alpha_\beta \delta \right).
\]

But, looking at this expression makes it obvious that its dependence on the vector-field \( v \) is completely local (the vector field is never differentiated) and linear. Thus, we can define a second-rank tensor field \( Dw \) according to

\[
Dw (\varphi; v) = \varphi \cdot D_v w.
\]

The components of this tensor field are traditionally denoted by \( w^{\alpha, \beta} \) where

\[
Dw = w^{\alpha, \rho} \delta e_\alpha \otimes \omega^\delta \\
D_v w = v^\delta w^{\alpha, \beta} e_\alpha
\]

and evidently

\[
w^{\alpha, \beta} = e_\delta w^\alpha + w^\beta \Gamma^\alpha_\beta \delta
\]

This tensor-field is called the **covariant derivative** of the vector-field \( w \).

For any tensor-field, one has a covariant derivative tensor that is one rank higher. The tensor field with components \( Q^{\alpha, \beta, \gamma} \) considered earlier has a covariant derivative with components

\[
Q^{\alpha, \beta, \gamma, \delta} = e_\delta \left( Q^{\alpha, \beta, \gamma} + Q^{\rho^\beta_\gamma} \Gamma^\alpha_\rho \delta + Q^{\alpha^\rho_\gamma} \Gamma^\beta_\rho \delta - Q^{\alpha^\rho_\gamma} \Gamma^\rho_\gamma \delta \right)
\]
The index-free notation for the covariant derivative $DQ$ has the disadvantage of hiding the slot for the differentiating vector field. The directional derivative notation displays the differentiating vector field but at the expense of not showing off the full tensor nature of the result. The standard index notation in the previous equation shows the dependence on the differentiating vector field by showing the corresponding index — call it the differentiating index. For example, in

$$D_\delta w = Dw(\cdot, v) = v^\delta w^{\alpha \beta} e_\alpha (\cdot)$$

the differentiating index is $\delta$. However the standard index notation obscures the operator nature of the covariant derivative.

A modification of the index notation is sometimes used to fix this deficiency. Instead of denoting the components of $Dw$ by $w^{\alpha \beta}$ just denote them by $D_\delta w^{\alpha \beta}$. In this spirit,

$$D_\delta Q^{\alpha \beta} = e_\delta (Q^{\alpha \beta} + Q^{\beta \gamma} \Gamma_{\rho \delta}^{\alpha} + Q^{\alpha \rho} \Gamma_{\rho \delta}^{\beta} - Q^{\alpha \beta} \Gamma_{\rho \delta}^{\rho}. $$

I will not be using this notation. If you encounter it, be very careful to distinguish between the directional derivative $D_v$ and the operator notation $D_\delta$ for the components of the covariant derivative.

### 2.6 The Perils of second derivatives

In the notation being used here, the derivative $D_v w$ of a vector field $w$ in the direction of a vector field $v$ is interpreted as a second rank tensor field with $v$ inserted as one of its arguments:

$$D_v w = Dw(\cdot; v).$$

Thus, the second derivative $D_u D_v w$ is evaluated as the derivative of this second rank tensor field.

$$D_u D_v w = DDw(\cdot; v, u).$$

An anthropomorphic way of saying this is that the operator $D_u$ "sees" the differentiating vector field $v$ as a tensor argument. The expression for $D_u D_v w$ must then be locally linear in that argument so that, for any function $f$ on the manifold,

$$D_u D_f v w = f D_u D_v w.$$  

Often, however, we do not want to think of a differentiating vector field as a tensor argument. Here, we handle this situation by using a different symbol for the derivative. The derivative $\nabla_v w$ is defined as identical to $D_v w$ except that $v$ is not regarded as a tensor argument. Thus, the expression $\nabla_u \nabla_v w$ is simply the derivative of the vector field $\nabla_v w$. One evaluates it just like the derivative of any other vector field. For a function $f$, that means

$$\nabla_u \nabla_f v w = \nabla_u (\nabla_f v w) = \nabla_u (f \nabla v w) = (\nabla_u f) w + f \nabla_u (\nabla v w)$$
so this expression is \textit{not} tensorial in $v$. It is easy to check that the tensorial derivative $D_u D_v w$ can be expressed as

$$D_u D_v w = \nabla_u \nabla_v w - \nabla_{\nabla_u v} w.$$ 

### 2.7 Overloaded Operators

In order to avoid total confusion, it is very important to notice that the operators we are using act in different ways on different kinds of objects. Computer scientists refer to this type of operator as ‘overloaded’. Thus, if $f$ is a scalar function and $v$ a vector field, then one can see $v$ acting in two very different ways in the equation

$$vf = v(f) = v(df) = v \cdot df.$$ 

On a scalar field, it acts as a derivative. On a one-form field it is a local linear operator. Similarly, the dot product between two objects depends on the nature of the objects. For vectors $u, v$ and a one-form $\alpha$,

$$u \cdot v = g(u, v) \quad u \cdot \alpha = \alpha(v).$$

Suppose that $k$ and $f$ are functions on a manifold and $v$ is a vector field. Compare the following two relations:

$$v(kf) = v(k)f + kv(f)$$

$$v(kdf) = kv(df).$$

This type of notation places a heavy burden on the reader to remember exactly what type of object each symbol stands for. Physicists sometimes suffer a certain amount of culture shock because they are used to typographical clues to the nature of each symbol: vector signs or bold face type for vectors, gothic letters for tensors, and so on. Although it is customary to adopt some sort of typographical conventions such as using greek letters for forms and latin letters for vectors, these conventions are not absolute. Thus, you might be told that $f$ is a one-form field while $k$ is a scalar and $v$ is a vector. In that case, you would then know that $v(kf) = kv(f)$ rather than the relation given above.

### 3 Geodesics

#### 3.1 Definition

A geodesic is the natural generalization of a straight line in ordinary space. The basic definition is that such a path \textit{does not turn}. Given the path $C : \mathbb{R} \to M$ and the corresponding tangent vector field $u(\lambda) \in T_{C(\lambda)}$ defined along it we can calculate the directional derivative of this vector field along the curve.

$$\nabla_u u = u^\delta (e_\delta u^\alpha \rho + \omega^\rho_{\alpha \beta}) e_\alpha$$

$$= (uu^\alpha + u^\delta w^\rho \Gamma^\alpha_{\rho \beta}) e_\alpha$$

9
A geodesic is a curve for which the tangent vector $u$ does not turn. In other words, its rate of change must be in the same direction as $u$ so that the tangent vector simply gets longer or shorter. Thus, we require
\[ \nabla_u u = ku \]
for some scalar function $k$ along the curve.

### 3.2 Affine parameters

This geodesic definition is awkward. If we change the curve parameter from $\lambda$ to some parameter $\tilde{\lambda}$ then the quantity $\nabla_u u$ changes accordingly since $u \to \tilde{u} = \frac{\partial}{\partial \tilde{\lambda}}$

\[ \tilde{u} = \frac{d\lambda}{d\tilde{\lambda}} u, \quad u = \frac{d\lambda}{d\tilde{\lambda}} \tilde{u} \]

Define $\phi = \frac{d\lambda}{d\tilde{\lambda}}$ so that $u = \phi \tilde{u}$ and the geodesic condition becomes
\[ \nabla_{\phi \tilde{u}} (\phi \tilde{u}) = k \phi \tilde{u} \]
or
\[ \phi \nabla_{\tilde{u}} (\phi \tilde{u}) = k \phi \tilde{u} \]
which becomes
\[ \nabla_{\tilde{u}} (\phi \tilde{u}) = k \tilde{u} \]
or
\[ \nabla_{\tilde{u}} (\phi) \tilde{u} + \phi \nabla_{\tilde{u}} \tilde{u} = k \tilde{u} \]

Cast this expression in the form of the definition of a geodesic
\[ \nabla_{\tilde{u}} \tilde{u} = \frac{k - \nabla_{\tilde{u}} (\phi)}{\phi} \tilde{u} \]
or

\[ \nabla_{\tilde{u}} \tilde{u} = \tilde{k} \tilde{u}, \quad \tilde{k} = \frac{k - \nabla_{\tilde{u}} (\phi)}{\phi} \]

Thus, the arbitrary function $k$ or $\tilde{k}$ is linked to the arbitrary choice of curve parameter along the geodesic.

Now eliminate the arbitrariness by choosing the new curve parameter so that
\[ \nabla_{\tilde{u}} (\phi) = k. \]
so that the equation for a geodesic with this particular curve parameter is just
\[ \nabla_{\tilde{u}} \tilde{u} = 0. \]

We need to show that this new curve parameter can always be found. Rewrite its defining condition as
\[ \frac{d\phi}{d\lambda} = k \quad \rightarrow \quad \frac{d\phi}{d\lambda} \frac{d\lambda}{d\lambda} = k \]
but
\[ \phi = \frac{d\lambda}{d\tilde{\lambda}} \]
so that
\[ \frac{d\phi}{d\lambda} = k\phi \quad \rightarrow \quad \phi = \phi_0 e^{\int_0^\lambda k(\lambda') d\lambda'} \]
or
\[ \frac{d\tilde{\lambda}}{d\lambda} = \phi_0 e^{\int_0^\lambda k(\lambda') d\lambda'} \quad \rightarrow \quad \tilde{\lambda} = \tilde{\lambda}_0 + \phi_0 \int_0^\lambda e^{\int_0^{\lambda'} k(\lambda'') d\lambda''} d\lambda' \]

This expression exhibits the new curve parameter explicitly in terms of the old one. Notice that choosing the new parameter \( \tilde{\lambda} \) to be an increasing function of the old parameter at just one point — \( \frac{d\lambda}{d\lambda} \bigg|_{0} = \phi_0 > 0 \) — guarantees that \( \frac{d\lambda}{d\lambda} \) is always positive so that the new parameter is an increasing function of the old one everywhere and is guaranteed to be a good curve parameter. In terms of this new parameter,
\[ \nabla_{\tilde{\lambda}} \tilde{u} = 0. \]
The resulting curve is called a geodesic with an affine parameter along it.

Suppose that the parameters \( \lambda \) and \( \tilde{\lambda} \) are both affine parameters. Just set \( k = 0 \) in the above discussion and notice that the solution degenerates to
\[ \tilde{\lambda} = \tilde{\lambda}_0 + \phi_0 \lambda \]
where \( \tilde{\lambda}_0 \) and \( \phi_0 \) are both constants. Thus, different affine parameters along a given curve are all linearly related to one another.

### 3.3 Riemannian Coordinates

A geodesic curve with an affine parameter is completely determined once its tangent vector at one point is given. Use this fact to produce a special coordinate patch around a chosen point \( P_0 \). The simplest way to start is by working backwards — given a set of real number coordinates \( \{x^\mu\} \) specify the point \( P \) which is labeled by those points. In other words, define a map
\[ \exp : \mathbb{R}^n \rightarrow U \subset M \]
by specifying the point \( \exp(x^0, x^1, x^2, \ldots) \). To specify this point, choose a set of basis vectors \( \{e_\alpha\} \) for the tangent space \( T_{P_0} \) and start an affinely parameterized geodesic curve \( C_x \) from \( P_0 \) with the initial value of the parameter set to zero and the initial tangent vector set to \( x^\mu e_\mu \).

\[ C_x (0) = P_0, \quad \frac{\partial}{\partial \lambda} \bigg|_{C_x,\lambda=0} = x^\mu e_\mu \]

The image point is then defined to be the point with parameter value equal to 1.
\[ \exp (x) = C_x (1). \]
The map exp will produce a well-behaved coordinate system near the reference point $P_0$. It can break down if the geodesics start crossing each other — as they would on a sphere for example. These coordinates are usually called Riemannian coordinates and the map is often called the exponential map of the tangent space into the manifold. It has some special properties that we need to work out.

For each set of coordinates $x$, there is a geodesic $C_x$ such that $C_x(0) = P_0$ is the reference point and $C_x(1) = P$ is a point which corresponds to the coordinate set $x$. So long as we stay in a subset $U$ which is close enough to the reference point, the map from coordinate space $\mathbb{R}^n$ to $U$ is one-to-one and can be inverted to provide coordinate functions $X^\mu$ on $U$. Thus, the coordinates assigned to the point $P$ that was reached by the geodesic with initial tangent vector $x^\mu e_\mu$ are just the numbers $x^\mu$ again. Formally, the coordinate functions $X^\mu$ work like this:

$$X^\mu (C_x (1)) = x^\mu$$

What are the coordinates assigned to the point $C_x (\alpha)$ for some number $\alpha < 1$? This point will be reached by a geodesic $C_y$ which differs from $C_x$ only by having a different affine curve parameter so that

$$C_x (\alpha) = C_y (1)$$

The coordinates $y^\mu$ assigned to this point correspond to the initial tangent vector to the curve $C_y$, so that tangent vector is what we must find. The tangent vector is just $\frac{d}{d\lambda'}$ where $\lambda'$ is the affine parameter along $C_y$. Different affine parameters for geodesics which go through the same set of points must be linearly related. In this case, all of the affine parameters start at zero, so they must be related by constant factors. The affine parameter $\lambda'$ on $C_y$ must be some constant $k$ times the affine parameter $\lambda$ on $C_x$.

$$\lambda' = k \lambda$$

But the parameter values $\lambda' = 1$ and $\lambda = \alpha$ correspond to the same point so that

$$1 = k\alpha$$

or

$$k = \alpha^{-1}$$
and the affine parameter on $C_y$ is just $\lambda' = \alpha^{-1}\lambda$. Now find the tangent vector to $C_y$.

$$\frac{d}{d\lambda'} = \alpha \frac{d}{d\lambda}$$

or, in components

$$y^\alpha e_\alpha = \alpha x^\alpha e_\alpha$$

which yields

$$y^\alpha = \alpha x^\alpha$$

In terms of the coordinate functions and the curve $C_x$

$$X^\mu (C_x(\alpha)) = \alpha x^\mu = \alpha X^\mu (C_x(1))$$

For any geodesic through the reference point $P_0$, the coordinate functions are given by

$$X^\mu (C(\alpha)) = \alpha X^\mu (C(1))$$

In terms of the simpler notation $x^\mu (t) = X^\mu (C(t))$, this result implies that the coordinates along any geodesic through $P_0$ obey the relation

$$x^\mu (t) = tx^\mu (1) = tu^\mu$$

where $u^\mu$ is the initial tangent vector.

Now consider what the geodesic equation will look like in this coordinate system. Let the functions $x^\nu (t)$ describe the coordinate image of a curve passing through the reference point. The tangent vector to this curve is given by

$$u = \frac{d}{dt} = \frac{dx^\mu}{dt} \partial_\mu = \dot{x}^\mu e_\mu$$

Now calculate for a bit:

$$\nabla_u u = \nabla_u (u^\mu \partial_\mu) = (\nabla_u u^\mu) \partial_\mu + u^\mu \nabla_u \partial_\mu$$

$$\nabla_u u = u (u^\mu) \partial_\mu + u^\mu u^\nu \nabla_{e_\rho} \partial_\mu$$

$$= \frac{d}{dt} u^\mu \partial_\mu + u^\mu u^\nu \partial_\rho \Gamma^\mu_{\rho\nu}$$

$$= \dot{u}^\mu \partial_\mu + u^\mu u^\nu \partial_\rho \Gamma^\mu_{\rho\nu}$$

$$= (\ddot{x}^\rho + u^\mu u^\nu \Gamma^\mu_{\rho\nu}) \partial_\rho$$

and find that the geodesic equation becomes

$$\ddot{x}^\rho + u^\mu u^\nu \Gamma^\mu_{\rho\nu} = 0$$

But, in these coordinates,

$$x^\rho (t) = tx^\rho (1)$$

so that

$$\ddot{x}^\rho = 0$$
Putting this result back into the geodesic equation, we get the constraints
\[ u^\mu u^\nu \Gamma^\rho_{\mu\nu} (P_0) = 0 \]
for any values of the components \( u^\mu \).

Split the connection into symmetric and antisymmetric parts:
\[ \Gamma^\rho_{\mu\nu} = \Gamma^\rho_{<\mu\nu>} + \Gamma^\rho_{[\mu\nu]} \]
and notice that, by antisymmetry
\[ u^\mu u^\nu \Gamma^\rho_{[\mu\nu]} = -u^\mu u^\nu \Gamma^\rho_{[\nu\mu]} \]
and, by renaming dummy indexes,
\[ u^\mu u^\nu \Gamma^\rho_{[\mu\nu]} = -u^\nu u^\mu \Gamma^\rho_{[\mu\nu]} = -u^\mu u^\nu \Gamma^\rho_{[\mu\nu]} \]
so that
\[ u^\mu u^\nu \Gamma^\rho_{[\mu\nu]} = 0. \]
The antisymmetric part of the connection drops out of the constraints, leaving
\[ u^\mu u^\nu \Gamma^\rho_{(\mu\nu)} (P_0) = 0 \]
which can be satisfied for all values of \( u^\mu \) only if
\[ \Gamma^\rho_{(\mu\nu)} (P_0) = 0. \]
Thus, the symmetric part of the connection coefficient array vanishes at the reference point of a Riemannian coordinate system.

4 Connection Tensors

4.1 Torsion

Consider how covariant derivatives commute when they act on a function.
\[ [D_u, D_v] f = D_u D_v f - D_v D_u f \]
But
\[ D_u D_v f = \nabla_u \nabla_v f - \nabla_{[uv]} f \]
and these derivatives of functions are just the vectors themselves acting as operators
\[ D_u D_v f = u v f - (\nabla_u v) f = \{uv - \nabla_u v\} f \]
The commutator is then
\[ [D_u, D_v] f = \{uv - \nabla_u v\} f - \{vu - \nabla_v u\} f = \{[u,v] - \nabla_u v + \nabla_v u\} f \]
or, since \( f \) is an arbitrary function, the commutator is just a vector field.

\[
[D_u, D_v] = [u, v] - \nabla_u v + \nabla_v u.
\]

The torsion is defined to be a third rank tensor field that yields the vector field

\[
T(u, v) = -[D_u, D_v] = -[u, v] + \nabla_u v - \nabla_v u
\]

for any vector fields \( u, v \).

For future reference, note that the definition implies that the arguments \( u, v \) of the torsion tensor are both differentiating vector fields. Thus, a derivative operation such as \( \nabla_w T(u, v) \) will treat \( T(u, v) \) as a simple vector field and will not include counter terms for its arguments.

In a holonomic frame, the components of the torsion tensor are

\[
T^\rho_{\alpha\beta} = dx^\rho \cdot T(e_\alpha, e_\beta) = dx^\rho \cdot \left( \nabla_{e_\alpha} e_\beta - \nabla_{e_\beta} e_\alpha \right)
\]

so the torsion is essentially the antisymmetric part of the connection coefficient array.

The torsion is a tensor — it has to be because we constructed it from covariant derivatives — and no coordinate transformation can ever make it vanish. As we have seen, the rest of the connection, the symmetric part, can be made to vanish at a point by choosing Riemannian coordinates there.

### 4.2 Definition of Curvature

Now consider how covariant derivatives commute when they act on a vector field. As we saw earlier, the object with the simplest properties is

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

Here, \( \mathcal{R} \) is regarded as a linear operator on the vector field \( w \). This definition uses derivatives that do not treat the differentiating vector fields \( u, v \) as tensor arguments. Nevertheless, it produces a fourth-rank tensor. With all of its arguments made explicit, this Riemannian Curvature tensor is

\[
\text{Riem}(\alpha, w, u, v) = \alpha \cdot \mathcal{R}(u, v) w = \alpha \cdot [\nabla_u, \nabla_v] w - \alpha \cdot \nabla_{[u,v]} w
\]

where \( \alpha \) is a form and \( w, u, v \) are vectors. The components of this tensor can be expressed in terms of the connection and commutation coefficients as

\[
R^\alpha_{\beta\mu\nu} = \nabla_{e_\mu} \Gamma^\alpha_{\beta\nu} - \nabla_{e_\nu} \Gamma^\alpha_{\beta\mu} + \Gamma^\alpha_{\sigma\mu} \Gamma^\sigma_{\beta\nu} - \Gamma^\alpha_{\sigma\nu} \Gamma^\sigma_{\beta\mu} - c^\alpha_{\mu\nu} \Gamma^\alpha_{\beta\sigma}.
\]

where the derivatives are regarded as acting on ordinary scalar fields.
In an environment such as Scientific Notebook, where matrix operations are carried out automatically, it is useful to organize the curvature and connection in terms of matrices:

$$[\Gamma]_{\mu} = \begin{bmatrix}
\Gamma^0_{0\mu} & \Gamma^0_{1\mu} & \Gamma^0_{2\mu} & \Gamma^0_{3\mu} \\
\Gamma^1_{0\mu} & \Gamma^1_{1\mu} & \Gamma^1_{2\mu} & \Gamma^1_{3\mu} \\
\Gamma^2_{0\mu} & \Gamma^2_{1\mu} & \Gamma^2_{2\mu} & \Gamma^2_{3\mu} \\
\Gamma^3_{0\mu} & \Gamma^3_{1\mu} & \Gamma^3_{2\mu} & \Gamma^3_{3\mu}
\end{bmatrix}$$

$$[R]_{\mu\nu} = \begin{bmatrix}
R^0_{0\mu\nu} & R^0_{1\mu\nu} & R^0_{2\mu\nu} & R^0_{3\mu\nu} \\
R^1_{0\mu\nu} & R^1_{1\mu\nu} & R^1_{2\mu\nu} & R^1_{3\mu\nu} \\
R^2_{0\mu\nu} & R^2_{1\mu\nu} & R^2_{2\mu\nu} & R^2_{3\mu\nu} \\
R^3_{0\mu\nu} & R^3_{1\mu\nu} & R^3_{2\mu\nu} & R^3_{3\mu\nu}
\end{bmatrix}$$

so that the curvature expression becomes

$$[R]_{\mu\nu} = \nabla_{e_\mu} [\Gamma]_{\nu} - \nabla_{e_\nu} [\Gamma]_{\mu} + [\Gamma]_{\mu} [\Gamma]_{\nu} - [\Gamma]_{\nu} [\Gamma]_{\mu} - c^\tau_{\mu\nu} [\Gamma]_{\tau}$$

Because this expression is manifestly antisymmetric in the indexes \(\mu, \nu\) there are just six independent matrix expressions to evaluate.

### 4.3 Ricci Identities

Calculate the commutator of covariant derivatives acting on a vector field as follows:

$$[D_u, D_v] w = D_u D_v w - D_v D_u w$$

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

$$= [\nabla_u, \nabla_v] w + \nabla_{\nabla_v w} - \nabla_{\nabla_u w}$$

$$= [\nabla_u, \nabla_v] w + \nabla_{\nabla_v w - \nabla_u w}$$

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

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

$$= \mathcal{R} (u, v) w - \nabla T(u, v) w$$

$$\{[D_u, D_v] + D_T(u, v)\} w = \mathcal{R} (u, v) w$$

In terms of components, this identity takes the form

$$w^\alpha_{;\nu\mu} - w^\alpha_{;\mu\nu} + T^\delta_{\mu\nu \alpha;\delta} = w^\rho R^\alpha_{\rho\mu\nu}$$

The operator \{[D_u, D_v] + D_T(u, v)\} obeys Leibniz’s rule so that its action on tensors of various rank follows a pattern similar to that of covariant derivatives, with each index in turn getting contracted with the curvature tensor.

$$\varphi_\beta;\nu\mu - \varphi_\beta;\mu\nu + T^\delta_{\mu\nu \varphi_\beta;\delta} = -\varphi_\mu R^\rho_{\beta\mu\nu}$$

$$K^\alpha_{\beta;\nu\mu} - K^\alpha_{\beta;\mu\nu} + T^\delta_{\mu\nu K^\alpha_{\beta;\delta}} = K^\rho_{\beta} R^\alpha_{\rho\mu\nu} - K^\alpha_{\rho} R^\rho_{\beta\mu\nu}$$

$$g_{\alpha\beta;\nu\mu} - g_{\alpha\beta;\mu\nu} + T^\delta_{\mu\nu g_{\alpha\beta;\delta}} = -g_{\rho\delta} R^\rho_{\alpha\mu\nu} - g_{\alpha\delta} R^\delta_{\beta\mu\nu}$$
4.4 Bianchi Identities

4.4.1 Jacobi Identity

The fundamental consistency relation for any system of commutators is the Jacobi identity

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

which can be verified just by writing out the expression in terms of products — everything cancels. It does not matter what sort of objects are being commuted. So long as they obey an associative algebra (where factors can be regrouped), then the Jacobi identity must hold. Any time that a set of commutation relations is given, the Jacobi identity needs to be true.

4.4.2 Torsion

Take the operators \( A, B, C \) in the Jacobi identity to be covariant derivatives and apply the identity to a scalar field — a function on the manifold.

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

From here, just write out the commutators, working from the outside in and recognize the torsion and curvature terms as they occur. The result is the Torsion Bianchi Identity:

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

In the case of zero torsion, this identity takes the form of a symmetry of the curvature tensor

\[
R(u, v) w + R(w, u) v + R(v, w) u = 0
\]

or, in terms of components

\[
R^\sigma_{\alpha\beta\gamma} + R^\sigma_{\gamma\alpha\beta} + R^\sigma_{\beta\gamma\alpha} = 0.
\]

4.4.3 Curvature

Now apply the identity to a vector field on the manifold

\[
0 = [\nabla_t, [\nabla_u, \nabla_v]] w + [\nabla_v, [\nabla_t, \nabla_u]] w + [\nabla_u, [\nabla_v, \nabla_t]] w
\]

and, after applying the definitions of the curvature and torsion tensors, obtain the curvature Bianchi Identity in the form

\[
0 = D_t R(u, v) + D_v R(t, u) + D_u R(v, t) - R(u, T(v, t)) - R(v, T(t, u)) - R(t, T(u, v)).
\]
In the case of zero torsion, this identity is a symmetry of the covariant derivative of the curvature tensor

\[ 0 = D_t R(u, v) + D_v R(t, u) + D_u R(v, t) \]

or, in terms of components

\[ R^\rho_{\rho\alpha\beta\gamma} + R^\rho_{\rho\gamma\alpha\beta} + R^\rho_{\rho\beta\gamma\alpha} = 0 \]

where the last three indexes are permuted cyclically from one term to the next.

### 4.5 Metric

In order to specify the lengths of vectors, and the angles between vectors, a tensor \( g \) which eats pairs of vectors is needed — a second rank covariant tensor. Thus, the length of the vector \( v \) is given by \( \sqrt{g(v, v)} \) while the angle \( \theta \) between spacelike vectors \( u \) and \( v \) is given by \( \cos \theta = \frac{g(u, v)}{\sqrt{g(u, u)g(v, v)}} \). In the spacetime of relativity, we insist that there should be a set of basis vectors such that these expressions take their special relativity forms. For such basis vectors \( e_\alpha \), the metric has components

\[
\begin{align*}
g_{00} &= g(e_0, e_0) = -1 \\
g_{11} &= g(e_1, e_1) = +1 \\
g_{22} &= g(e_2, e_2) = +1 \\
g_{33} &= g(e_3, e_3) = +1 \\
g_{\mu\nu} &= g(e_\mu, e_\nu) = 0 \text{ for } \mu \neq \nu
\end{align*}
\]

or, more briefly, \([g] = \text{diag}(-1, +1, +1, +1)\). Mathematically, we can guarantee the existence of such a basis set for either \( g \) or \(-g\) by requiring three things of the metric of components \([g]\):

- (1) \( \det [g] < 0 \)
- (2) \( [g] \) is symmetric — \( [g]^T = [g] \)
- (3) \( [g] \) is invertible — \( [g]^{-1} \) exists.

The tensor \( g \) can be regarded as a map \( g : T_P \rightarrow \hat{T}_P \) where the form \( g(u) \) which is assigned to a vector \( u \) acts on the vector \( v \) according to the definition

\[ g(u)(v) = g(u, v) \]

In terms of components,

\[ g(u)_\alpha = g_{\alpha\beta} v^\beta. \]

The inverse of this map, has the component representation:

\[ g^{-1}(\varphi)^\gamma = g^{\gamma\alpha} \varphi_\alpha. \]
Putting these last two expressions together, we find
\[ g^\gamma_\alpha g_\alpha_\beta = \delta^\gamma_\beta \]
or, in terms of matrix arrays of components,
\[ [g^{-1}] [g] = I, \quad [g^{-1}] = [g]^{-1}. \]
Now define a second rank tensor which acts on pairs of forms according to
\[ g^{-1} (\varphi, \psi) = g^{-1} (\varphi) \cdot \psi = \varphi \cdot g^{-1} (\psi). \]
This contravariant second rank tensor has the components
\[ g^{-1} (\omega^\beta, \omega^\alpha) = g^{-1} (\omega^\beta) \cdot \omega^\alpha = g^{\alpha\beta} \]

Historically, the metric tensor field has been regarded as the fundamental field variable of general relativity with everything else derived from it. That point of view has changed radically during the last few years. It now looks as if the connection is the fundamental gravitational variable while the metric tensor is a derived quantity which is sometimes needed in order to describe interactions between gravity and matter.

### 4.6 Diagonalization of the Metric

For a basis transformation
\[ e_{\alpha'} = U_{\alpha'}^\rho e_\rho \]
to yield an orthonormal basis, the coefficients \( U_{\alpha'}^\beta \) must obey the conditions
\[ U_{\alpha'}^\rho e_\rho \cdot U_{\beta'}^\sigma e_\sigma = \eta_{\alpha\beta} = \begin{cases} -1 & \text{for } \alpha = \beta = 0 \\ 1 & \text{for } \alpha = \beta \neq 0 \\ 0 & \text{for } \alpha \neq \beta \end{cases} \]
or
\[ U_{\alpha'}^\rho U_{\beta'}^\sigma g_{\rho\sigma} = \eta_{\alpha\beta} \]
In matrix notation, these conditions take the form
\[ [U] [g] [U]^T = [\eta] \]
Notice that these are not the same as the conditions on a similarity transformation that diagonalizes the matrix \([g]\). A similarity transformation would take the form \( A [g] A^{-1} \).

If the original basis is already orthonormal, the condition on a basis change to a new orthonormal set is
\[ [U] [\eta] [U]^T = [\eta] \]
In this case, the transformation
\[ [e'] = [U] [e] \]
is a Lorentz Transformation.
4.7 Metricity

The link between the metric tensor and the connection is the covariant derivative of the metric tensor — the metricity tensor. For an arbitrary metric in an arbitrary connection, the metricity tensor is just

\[ Q(\alpha, \beta, v) = -D_v g^{-1}(\alpha, \beta) \]

or

\[ Q^{\alpha\beta}_{\gamma} = -D_v g^{\alpha\beta} = -\varepsilon_{\gamma} g^{\alpha\beta} - g^{\alpha\beta} \Gamma^\alpha_{\rho\gamma} - g^{\alpha\rho} \Gamma^\beta_{\rho\gamma}. \]

I am defining the metricity in terms of the inverse metric tensor because that turns out to be more useful in calculations of gravitational dynamics than the regular metric tensor. From matrix algebra, any derivative operator \( D \) acting on the inverse of a matrix yields

\[ D (m^{-1}) = -m^{-1} (Dm) m^{-1} \]

so that an equivalent definition of metricity is

\[ Q(a, b, v) = D_v g(a, b) \]

or

\[ Q^{\alpha\beta\gamma} = g_{\alpha\rho} Q^{\rho\gamma} g_{\beta\gamma} = D_v g_{\alpha\beta} = g_{\alpha\beta;\gamma} = \varepsilon_{\gamma} g_{\alpha\beta} - g_{\rho\beta} \Gamma^\rho_{\alpha\gamma} - g_{\alpha\rho} \Gamma^\beta_{\rho\gamma} \]

Ordinarily, one requires the metricity tensor to vanish so that the covariant derivatives of inner products obey Leibniz’s rule:

\[ D_v (u \cdot v) = D_v (u^a g_{\alpha\beta} v^\beta) = (D_v u^a) g_{\alpha\beta} v^\beta + u^a Q_{\alpha\beta\gamma} v^\beta + u^a g_{\alpha\beta} (D_v v^\beta) \]

The metricity amounts to the covariant derivative of the “dot”. A connection for which the metricity tensor vanishes is called a “metric compatible connection”. This condition is usually interpreted as a constraint on the connection coefficients:

\[ \varepsilon_{\gamma} g_{\alpha\beta} = g_{\rho\beta} \Gamma^\rho_{\alpha\gamma} + g_{\alpha\rho} \Gamma^\gamma_{\beta\rho} \]

I will follow the usual interpretation for a while and get expressions for the connection coefficients in terms of the metric. However, before we mess up this expression any further, notice that it is a beautifully simple system of linear first order differential equations in the metric components. Given the connection coefficients, we can recover the metric tensor by solving this system of linear equations.

The standard approach defines the object

\[ \Gamma_{\alpha\beta\gamma} = g_{\alpha\rho} \Gamma^\rho_{\beta\gamma} \]

and notices that we already know that the torsion tensor is given by

\[ T_{\alpha\beta\gamma} = g_{\alpha\rho} \xi^\rho_{\beta\gamma} - 2 \Gamma^\rho_{\alpha[\beta\gamma]} \]
so that
\[ \Gamma_{\alpha[\beta\gamma]} = \frac{1}{2} \left( g_{\alpha\rho} e^\rho_{\beta\gamma} - T_{\alpha\beta\gamma} \right) \]
The metric compatibility condition takes the form
\[ e_\gamma g_{\alpha\beta} = \Gamma_{\beta\alpha\gamma} + \Gamma_{\alpha\beta\gamma} \]
or
\[ \Gamma_{(\alpha\beta)\gamma} = \frac{1}{2} e_\gamma g_{\alpha\beta}. \]
This system of linear equations can be solved for the connection coefficients.

The system to be solved is
\[ g_{\alpha\beta} e^\rho_{\gamma\beta} = b_{\alpha\beta\gamma}; \quad C_{\alpha[\beta\gamma]} = a_{\alpha\beta\gamma} \]
or, equivalently
\[ \Gamma_{\alpha\beta\gamma} + \Gamma_{\beta\alpha\gamma} = b_{\alpha\beta\gamma} \]
\[ \Gamma_{\alpha\beta\gamma} - \Gamma_{\alpha\gamma\beta} = a_{\alpha\beta\gamma} \]
Cyclically permute the arguments of this last equation and write it as
\[ \Gamma_{\beta\alpha\gamma} = \Gamma_{\beta\gamma\alpha} + a_{\beta\alpha\gamma} \]
which can be substituted into the first equation of the pair
\[ \Gamma_{\alpha\beta\gamma} + \Gamma_{\beta\gamma\alpha} + a_{\beta\alpha\gamma} = b_{\alpha\beta\gamma} \]
and gives the key result:
\[ \Gamma_{\alpha\beta\gamma} + \Gamma_{\beta\gamma\alpha} = b_{\alpha\beta\gamma} - a_{\beta\alpha\gamma}. \]
Now denote cyclic index permutations by
\[ (CT)_{\alpha\beta\gamma} = \Gamma_{\beta\gamma\alpha} \]
and
\[ B_{\alpha\beta\gamma} = b_{\alpha\beta\gamma} - a_{\beta\alpha\gamma} \]
so that the equation that we have to solve becomes just
\[ \Gamma + CT = B. \]
Operate on it twice with the cyclic permutation operator
\[ CT + C^2 \Gamma = CB \]
\[ C^2 \Gamma + \Gamma = C^2 B. \]
Now we have three equations in the three unknowns \( \Gamma, CT, C^2 \Gamma \).
Eliminate $C^2\Gamma$ using the last equation
\[ C^2\Gamma = C^2B - \Gamma. \]
Next, eliminate $CT$ by using middle equation
\[ CT = CB - C^2\Gamma = CB - C^2B + \Gamma. \]
The remaining equation then becomes
\[ \Gamma + CB - C^2B + \Gamma = B \]
or
\[ 2\Gamma = C^2B - CB + B \]
\[ \Gamma = \frac{1}{2} (C^2B - CB + B) \]
\[ \Gamma_{\alpha\beta\gamma} = \frac{1}{2} (B_{\gamma\alpha\beta} - B_{\beta\gamma\alpha} + B_{\alpha\beta\gamma}) \]
Now unravel the definitions
\[ B_{\alpha\beta\gamma} = b_{\alpha\beta\gamma} - a_{\beta\alpha\gamma} \]
\[ = e\gamma g_{\alpha\beta} - g_{\beta\alpha} e\sigma\gamma + T_{\beta\alpha\gamma} \]
and get the final expression for the connection coefficients in terms of the metric and torsion tensors.
\[ \Gamma^\alpha_{\beta\gamma} = \frac{1}{2} g^{\alpha\sigma} (e\gamma g_{\sigma\beta} - g_{\beta\sigma} e\rho\sigma\gamma + T_{\beta\sigma\gamma} + e\beta g_{\gamma\sigma} - g_{\sigma\beta} e\rho\gamma\beta + T_{\sigma\beta\gamma} - e\sigma g_{\beta\gamma} + g_{\beta\sigma\gamma} e\sigma\beta - T_{\gamma\beta\sigma}) \]
In the case of a holonomic frame and zero torsion, this expression reduces to just a combination of derivatives of the metric tensor components.
\[ \Gamma^\alpha_{\beta\gamma} = \frac{1}{2} g^{\alpha\sigma} (e\gamma g_{\sigma\beta} + e\beta g_{\gamma\sigma} - e\sigma g_{\beta\gamma}) \]

4.8 Curvature-Metricity Identity
Let the curvature operator act on the metric tensor
\[ \{ [D_u, D_v] + D_{T(u,v)} \} g(a,b) = -g(\mathcal{R}(u,v)a,b) - g(a,\mathcal{R}(u,v)b) \]
or
\[ D_u D_v g(a,b) - D_v D_u g(a,b) + D_{T(u,v)} g(a,b) = -g(\mathcal{R}(u,v)a,b) - g(\mathcal{R}(u,v)b,a) \]
which, from the definition of metricity yields
\[ D_u Q(a,b,v) - D_v Q(a,b,u) + Q(a,b,T(u,v)) = -g(\mathcal{R}(u,v)a,b) - g(\mathcal{R}(u,v)b,a) \]
or, in terms of indexes,
\[ Q_{\alpha\beta\nu;\mu} - Q_{\alpha\beta\mu;\nu} + Q_{\alpha\beta\delta} T^\delta_{\mu\nu} = -R_{\beta\alpha\mu\nu} - R_{\alpha\beta\mu\nu} \]
which connects the metricity to the curvature tensor. When the metricity is zero, this identity becomes an additional index symmetry of the curvature tensor
\[ R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\mu\nu} \]
4.9 Index Symmetries of the Curvature Tensor

From its definition, the curvature tensor obeys

\[ R_{\alpha\beta\nu\mu} = -R_{\alpha\beta\mu\nu}. \]

In the most general case, this is the only symmetry which the tensor has. Count its components in a four dimensional space: Array its independent components in rows according to the last two indexes and columns according to the first two. That makes six independent rows and sixteen independent columns for a total of 96 independent components.

In the most common case, the connection is metric compatible, which, from the curvature-metricity identity, guarantees antisymmetry in the first two indexes. The array now has just six columns and six rows and the curvature has 36 independent components.

If, in addition to being metric compatible, the connection is torsion-free, then the torsion Bianchi identities put additional constraints on the curvature components:

\[ R^{\sigma}_{\alpha \beta \gamma} + R^{\sigma}_{\gamma \alpha \beta} + R^{\sigma}_{\beta \gamma \alpha} = 0 \]

These additional 16 equations reduce the number of components to just 20.

Lower an index in the torsion Bianchi identities and start manipulating them, along with the antisymmetries in the first two and second two indexes:

\[ R_{\sigma \alpha \beta \gamma} + R_{\sigma \gamma \alpha \beta} + R_{\sigma \beta \gamma \alpha} = 0 \]  \hspace{1cm} (1)

symmetrize on \( \alpha, \sigma \) to so that the first term vanishes and the remaining two yield

\[ R_{\sigma \gamma \alpha \beta} + R_{\alpha \gamma \sigma \beta} + R_{\sigma \beta \gamma \alpha} + R_{\alpha \beta \gamma \sigma} = 0 \]

or, re-grouping the first term with the last and the second with the third,

\[ (R_{\sigma \gamma \alpha \beta} + R_{\alpha \gamma \sigma \beta}) + (R_{\alpha \gamma \sigma \beta} + R_{\sigma \beta \gamma \alpha}) = 0 \]

and using the antisymmetry of the first and second pairs of curvature indexes

\[ (R_{\sigma \alpha \beta \gamma} - R_{\gamma \alpha \sigma \beta}) + (R_{\sigma \beta \gamma \alpha} - R_{\gamma \alpha \sigma \beta}) = 0 \]  \hspace{1cm} (2)

Now symmetrize equation (1) on \( \beta, \sigma \) so that the last term vanishes

\[ R_{\sigma \alpha \beta \gamma} + R_{\beta \sigma \alpha \gamma} + R_{\sigma \gamma \alpha \beta} + R_{\beta \gamma \alpha \sigma} = 0 \]

or

\[ (R_{\sigma \alpha \beta \gamma} - R_{\beta \gamma \sigma \alpha}) + (R_{\beta \sigma \alpha \gamma} - R_{\gamma \alpha \sigma \beta}) = 0 \]  \hspace{1cm} (3)

Then symmetrize equation (1) on \( \gamma, \sigma \) so that the middle term vanishes

\[ R_{\sigma \alpha \beta \gamma} + R_{\gamma \alpha \beta \sigma} + R_{\sigma \beta \gamma \alpha} + R_{\gamma \beta \sigma \alpha} = 0 \]

or

\[ (R_{\sigma \alpha \beta \gamma} - R_{\beta \gamma \sigma \alpha}) + (R_{\gamma \alpha \beta \sigma} - R_{\beta \sigma \gamma \alpha}) = 0 \]  \hspace{1cm} (4)
The three equations (2,3,4) have the form

\[
\begin{align*}
  z + x &= 0 \\
  y + z &= 0 \\
  x + y &= 0
\end{align*}
\]

where \( x, y, z \) are the tensor combinations in parentheses. The solution to this system is easily seen to be \( x = y = z = 0 \) so that we obtain the index symmetry

\[ R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta}. \]

Notice that this last index symmetry is derived from the torsion Bianchi identities for zero torsion. It does not contain quite all of the information which is in those identities, however. One way to see this is to count the independent components of the Riemann tensor again, using this index symmetry. Think of the array of components as having rows and columns labeled by antisymmetrized index pairs. That makes it a six by six array. The last index symmetry makes that array symmetric which means that it has \( \frac{6 \times 5 \times 2}{3} \) independent components — 21. Since we know from counting the torsion Bianchi identities that there are only 20 independent components, there is evidently one constraint that is not captured by any of the index symmetries.

To see what the remaining constraint is, notice that we symmetrized various indexes with the one fixed index in the torsion identity. That procedure would miss whatever constraint is gotten by antisymmetrizing those indexes. Since the torsion bianchi identities are already antisymmetric in three indexes and have the form

\[ R_{\rho[\alpha\beta\gamma]} = 0 \]

the only possible result of antisymmetrizing the remaining fixed index \( \rho \) with any other index is the totally antisymmetrized expression

\[ R_{[\alpha\rho\beta\gamma]} = 0. \]

The one component of the curvature tensor that is permitted by the index symmetries but forbidden by this remaining constraint is

\[ R_{\alpha\rho\beta\gamma} = \phi \varepsilon_{\alpha\rho\beta\gamma} \]

where \( \phi \) is an arbitrary function and \( \varepsilon_{\alpha\rho\beta\gamma} \) is the totally antisymmetric Levi-Civita symbol. The missing constraint is that this totally antisymmetric component of the Riemann tensor must be zero.

\[ \varepsilon^{\alpha\rho\beta\gamma} R_{\alpha\rho\beta\gamma} = 0. \]