We now need to develop a little bit more math before we can discuss space-time curvature and Einstein equations.

So far we have discussed vectors, e.g.

\[ a^x = a^a \mathbf{e}_a \]

where \( a^a \) are the components of vector \( a \) in the basis described by the \( \{ \mathbf{e}_a \} \).

We also described the metric tensor, which is essential to calculate proper times or proper distances and generally gives the scalar product of two vectors, i.e.

\[ a \cdot b = g_{\alpha \beta} a^\alpha b^\beta \]

We can think of the metric tensor \( g \) as a linear map of two vectors into the number given by their scalar product, i.e.

\[ g(a, b) = a^\alpha b^\beta g_{\alpha \beta} = g(b, a) \]

by linear map we mean the usual linearity property,

\[ g(\alpha a + \beta b, c) = \alpha g(a, c) + \beta g(b, c) = (\alpha a^\alpha + \beta b^\beta) c^\beta g_{\alpha \beta} \]

which is obviously captured by component summation, i.e.

\[ a^\alpha b^\beta g_{\alpha \beta} + + \beta b^\beta c^\beta g_{\alpha \beta} \]

Now, let us be a bit more precise on the definition of vectors, and discuss their dual objects, dual vectors or 1-forms, and we'll see that the metric is, in fact, a dual object. This will make clear as well why we use upper indices to denote components of a vector, and lower indices for the metric.
Vectors (again!)

The simplest definition of a vector that carries over to curved space is to define vectors by directional derivatives (rather than directed line segments, for which one has to introduce tangent space in curved space).

Consider a function \( f(x^a) \) and a curve \( x^a(\sigma) \) labelled by parameter \( \sigma \). The directional derivative along the curve at point labelled by \( \sigma \) is

\[
\frac{df}{d\sigma} = \lim_{\varepsilon \to 0} \frac{f[x^a(\sigma+\varepsilon)] - f[x^a(\sigma)]}{\varepsilon} = \frac{df}{dx^a} \frac{dx^a}{d\sigma}
\]

but the vector \( \mathbf{t} \) with (coordinate) basis components \( t^a = dx^a/d\sigma \) is the tangent vector to the curve, so we can write

\[
\frac{df}{d\sigma} = \frac{dx^a}{d\sigma} \frac{\partial f}{\partial x^a} = t^a \frac{\partial f}{\partial x^a}
\]

so given a vector \( \mathbf{v} \) (with components \( v^a \)) we can define a directional derivative given as above, and as we just saw, the directional derivative, one can identify a vector. So there is a one-to-one correspondence between directional derivative & vectors - thus we may define vectors as directional derivative, i.e.

\[
\mathbf{v} \equiv v^a \frac{\partial}{\partial x^a}
\]

Thus one can think of \( \frac{\partial}{\partial x^a} \) as another notation for the coordinate basis vectors. This notation is very useful when you want to find components of a vector in different basis, e.g.

\[
\mathbf{a} = a^\beta \frac{\partial}{\partial x^\beta} = a^\beta \frac{\partial}{\partial x^a} \frac{\partial x^a}{\partial x^\beta} \equiv a^\beta \frac{\partial}{\partial x^\beta}
\]
So, after changing coordinates $x^a \rightarrow x'^a$ the new components of a vector $a$ are given by

$$a'_a = \frac{\partial x'^a}{\partial x^b} a^b$$

Thus the partial derivative notation for coordinate basis vectors is very handy.

**Dual Vectors (1-forms)**

A dual vector is simply a linear map that maps one vector to real numbers (also called a one-form).

Again linear map means that dual vector $w$ of dual vector $w(a\alpha + \beta b) = \alpha w(a) + \beta w(b)$

and we can write components for dual vector as usual

$$w(a) = w_\alpha a^\alpha$$

The simplest example of a dual vector is the gradient. Yes, that's right, the gradient is a dual vector, not a vector! (as it is usually introduced in calculus). Indeed the gradient maps a vector (tangent to a curve) to a number (the directional derivative). Along the curve

$$\frac{df}{dx} = \frac{\partial f}{\partial x^a} \cdot a^a$$

A set of 4 linearly independent dual vectors $\{\tilde{e}^a\}$ constitutes a basis.
for all dual vectors, and any dual vector \( \tilde{e} \) can be written in such basis as
\[
\tilde{e} = e^\alpha \tilde{e}_\alpha \quad \text{basis of dual vector} \quad \text{components}
\]
(note we use \( - \) for dual vectors as well as for vectors, why this notation makes sense will become clear shortly)

What's the relation between basis vectors and basis dual vectors? The action of \( \tilde{e}^\alpha \) dual vector on a vector (e.g., \( e_\beta \) basis vector) produces a number, in this case
\[
e^\alpha (\tilde{e}_{\beta}) = \delta^\alpha_{\beta}
\]
that's what mean for the two bases to be dual to each other.

When \( \delta^\alpha_{\beta} \) is the Kronecker delta (\( = 1 \) if \( \alpha = \beta \), otherwise zero).

This makes sense, since then
\[
\omega(\tilde{a}) = \sum_{\alpha} \omega_{\alpha} e^\alpha (\tilde{e}_{\beta}) = \sum_{\alpha} \omega_{\alpha} \delta^\alpha_{\beta} e^\alpha (\tilde{e}_{\beta}) = \sum_{\alpha} \omega_{\alpha} \delta^\alpha_{\beta} e^\alpha (\tilde{e}_{\beta})
\]

\[
= \sum_{\alpha} \omega_{\alpha} e^\alpha
\]

Let's look at the gradient again: How does its components change when we change coordinates \( x^\alpha \to x'^\alpha \)? That's pretty easy.

\[
\tilde{e}^\alpha \frac{\partial f}{\partial x^\alpha} = e^\alpha \frac{\partial f}{\partial x'^\alpha} \frac{\partial x'^\alpha}{\partial x^\alpha} = \tilde{e}^\beta \frac{\partial f}{\partial x'^\beta}
\]

So we have:
\[
\frac{\partial f}{\partial x'^\beta} \frac{\partial x'^\beta}{\partial x^\alpha} = \frac{\partial f}{\partial x^\alpha}
\]

new components
\[
\frac{\partial x'^\beta}{\partial x^\alpha}
\]
old components

This is the opposite transformation law as for vectors (see above).

We sometimes represent vectors as arrows, is there a
Simple graphical representation for dual vectors? Indeed, there is.
The picture must reflect that maps vectors into real numbers: the
one typically used by mathematicians is that the dual vectors
correspond to a series of surfaces:

The number produced when a dual vector acts on a vector is the
number of surfaces that the arrow of the vector goes through, e.g.

this will give $\approx 3.5$ - the closer the spacing of surfaces, the larger
the number will be (for a fixed vector). Note dual vectors do not
define a unique direction (since it's not a vector!), but rather a
way to slice the space.

This makes sense for the gradient! A large gradient means the
function is changing faster, so its contour levels are more closely spaced

Contours of $f$:

Usually in calculus people call the normal to $f = \text{const}$. the direction of
the gradient, $\mathrm{grad}$ and associate a vector with it. But we see
that in order to do that we have to say the gradient is the vector
that crosses the greatest number of contours per unit length. The
key here is that to define length one needs the metric!
Thus, as a stand-alone geometrical object (maps vectors into numbers, note no metric is needed here) the gradient is not a vector. To make the gradient a vector one must introduce the metric.

yet again!

Now, what is a vector anyway? The dualism discussed above is in fact complete. We can now regard a vector as a linear map from dual vectors into numbers!

\[
a \cdot (\mathbf{\omega}) = a^\alpha \omega_\alpha
\]

This should look familiar. We do these mappings through scalar products all the time: e.g. in linear algebra we have column vectors and the dual row vectors so the product is

\[(a \ b \ \cdots) \begin{pmatrix} P_1 & \cdots & P_n \end{pmatrix} = a P_1 + b P_2 + \cdots\]

Also in quantum mechanics, we have the inner product

\[\int \phi^*(x) \psi(x) \, dx\]

where \(\psi\) is the wave function, and \(\phi^*\) is complex conjugate (dual).

We can now bring the metric into the conversation. Since the metric maps two vectors into a number, the following object

\[g(a, \mathbf{\omega})\]

is a covector, or one-form: indeed it maps a vector (that has to be fed into the 2nd slot) into a number.
So we say the metric connects a vector $\vec{a}$ with a corresponding one-form or covector $\tilde{a}$ - since this connection is one-to-one, we can think of vectors and covectors as different kinds of components of the same object, i.e., we write for the one-form $\tilde{a}$:

$$\tilde{a}^\beta = g(\tilde{a}^\alpha \tilde{a}^\beta) = g_{\alpha \beta} = \frac{\partial}{\partial a^\alpha} a^\beta$$

by one-form component definition.

Since this is true for any $\tilde{a}^\beta$, we obtain the covector or one-form components in terms of the vector components through the metric:

$$a^\beta = g_{\alpha \beta} a^\alpha$$

"lowering an index"

Note we distinguish type of components by position of index: $a^\beta$ are components of covector or one-form $g(\tilde{a}^\beta)$, $a_\alpha$ are components of vector $\tilde{a}^\alpha$; but since they are one-to-one (through the metric), we use the same letter to denote them. Sometimes, we can be "loose" and refer to $a^\beta$ and $a_\alpha$ as the "lower and upper components" of vector $\tilde{a}^\alpha$.

We can find the inverse relation by introducing the inverse of the metric, $g_{\alpha \beta} g^\beta \gamma = \delta^\alpha \gamma$:

$$\Rightarrow \quad a^\alpha = g^\alpha \beta a_\beta$$

"raising an index"

We can play similar games with raising or lowering indices of any tensor using the metric. Clearly raising index on metric...
tensor itself is particularly easy,

\[ g^{\alpha \beta} = g^{\alpha \gamma} g_{\gamma \beta} = g^{\alpha \beta} \]

\[ g_{\alpha \beta} = g_{\beta \alpha} g^{\alpha \gamma} = g_{\beta \alpha} g^{\alpha \gamma} = g^{\beta \gamma} \] (as it should)

In general, e.g.

\[ t_{\alpha \gamma} = g_{\gamma \delta} t_{\alpha \delta} \], etc.

The Covariant derivative

We are interested in taking derivatives of vector fields in curved spacetime, and would expect such derivative of a vector to be a tensor with components \( \nabla_{\alpha} \mathbf{v}^\beta \), i.e. one index up for the vector, one index down for the derivative.

The tricky part, however, is that defining a derivative of a vector one needs to compare vectors at different spacetime points, and operation between vectors is only defined at the same point (where tangent space is unique). To define derivative of vectors, therefore, one must transport vectors from one spacetime point to another.

The covariant derivative of a vector field \( \mathbf{v} \) in the direction \( \partial \) in curved spacetime is defined as

\[ \nabla_{\alpha} \mathbf{v} = \lim_{\epsilon \to 0} \left[ \frac{\mathbf{v}(x^\alpha + \epsilon^\alpha \mathbf{e}) - \mathbf{v}(x^\alpha)}{\epsilon} \right] - \nabla(x^\alpha) \]

where the vector \( \mathbf{v}(x^\alpha + \epsilon^\alpha \mathbf{e}) \) is parallel transported to the point \( x^\alpha \), as defined in a local inertial frame (LIF), where parallel-transported
can be carried out as in flat space-time. Pictorially:

\[ \nabla_\pm \vec{V} = \mp \frac{2 \vec{V}}{dx^\pm} \quad \text{(LIF)} \]

In such a LIF, the components of \( \vec{V} \) do not change as they are parallel transported, and thus

\[ (\nabla_\pm \vec{V})^\alpha = \mp \beta \frac{2 \vec{V}^\alpha}{dx^\pm} \quad \text{(LIF)} \]

or in components

\[ \nabla_\beta V^\alpha = \frac{\partial V^\alpha}{\partial x^\beta} \quad \text{(LIF)} \]

which is the usual notion of derivative. In generic coordinate systems, however, components of \( \vec{V} \) will not stay constant as they are parallel transported. To see this consider a simple example, the 2D polar coordinates (where metric is non-trivial)

We can write the piece induced by the change in basis vectors as parallel transport is performed as linear in \( \vec{V} \) and \( \vec{\varepsilon} \), i.e.

\[ V^\alpha_{\parallel} (x^\delta) = V^\alpha (x^\delta + \varepsilon + \delta \varepsilon) + \Lambda^\alpha_{\beta \delta} (x^\delta) \vec{V}^\beta (x^\delta) \varepsilon + \delta \varepsilon \]

or in components, we have for the covariant derivative

\[ \nabla_\beta V^\alpha = \frac{\partial V^\alpha}{\partial x^\beta} + \Lambda^\alpha_{\beta \delta} \nabla^\delta \]

the point is that each term is basis dependent, but the total is not. So
should be clear from the definition of covariant derivative.

How do we obtain the \( \nabla \phi \)? We already know how, actually, since the concept of parallel transport is encoded in the
geodesic equation! A geodesic is a curve that has its tangent vector propagated parallel to itself (i.e., the closest thing to a straight line) — that means its covariant derivative in its own direction vanishes,

\[
(\nabla_u \bar{u})^\alpha = u^\beta \left( \frac{\partial \bar{u}^\alpha}{\partial x^\beta} + \Gamma^\alpha_{\beta\gamma} \bar{u}^\gamma \right) = 0
\]

where \( \bar{u}^\alpha = dx^\alpha/d\tau \). But the geodesic equation can be written as

\[
u^{\beta}(\frac{\partial \bar{u}^{\alpha}}{\partial x^{\beta}} + \Gamma^{\alpha}_{\beta\gamma} \bar{u}^{\gamma}) = 0
\]

since \( \frac{\partial u^{\alpha}}{\partial x^{\beta}} \frac{dx^{\beta}}{d\tau} = \frac{du^{\alpha}}{d\tau} \).

Then we see that \( \Gamma^\alpha_{\beta\gamma} = \Gamma^\alpha_{\gamma\beta} \) are the Christoffel symbols — thus the covariant derivative of a vector reads,

\[
\nabla_{\alpha} V^{\beta} = \frac{\partial V^{\beta}}{\partial x^{\alpha}} + \Gamma^{\beta}_{\alpha\gamma} V^{\gamma}
\]

The covariant derivative of a scalar is just the partial derivative, as a scalar does not change as it is parallel transported:

\[
\nabla_{\alpha} f = \frac{\partial f}{\partial x^{\alpha}} \quad \text{or} \quad \nabla_{\alpha} f = u^{\alpha} \frac{\partial f}{\partial x^{\alpha}}
\]

Similarly, for higher ranked tensors we have

\[
\nabla_{\alpha} t^{\alpha\beta} = \frac{\partial t^{\alpha\beta}}{\partial x^{\alpha}} + \Gamma^{\alpha}_{\lambda\beta} t^{\lambda\gamma} + \Gamma^{\alpha}_{\gamma\beta} t^{\alpha\gamma}
\]

i.e., one \( \Gamma \) factor for each index.

How about covectors or one forms?
Since $\gamma^a \delta^b$ is a scalar we can use std derivative
\[ \nabla_\xi (\gamma^a \delta^b) = \frac{\partial}{\partial x^\xi} (\gamma^a \delta^b) = \frac{\partial \gamma^a}{\partial x^\xi} \delta^b + \gamma^a \frac{\partial \delta^b}{\partial x^\xi} \]
Imposing Leibnitz rule we must have
\[ \nabla_\xi (\gamma^a w^b) = (\nabla^\xi \gamma^a) w^b + \gamma^a \nabla^\xi w^b \]
\[ \text{want:} \frac{\partial w^b}{\partial x^\xi} + \Gamma^b_{\alpha \beta} \delta^\alpha w^\beta \]
\[ \Rightarrow \frac{\partial x^\alpha}{\partial x^\xi} w^\alpha + \gamma^a \delta^b \nabla^\xi \gamma^a = w^\alpha \nabla^\xi \delta^\alpha + \gamma^a (\frac{\partial \delta^\alpha}{\partial x^\xi} + \Gamma^\alpha_{\beta \delta} \delta^\beta) \]
\[ \Rightarrow w^\alpha \nabla^\xi \delta^\alpha = w^\alpha \partial x^\alpha - \gamma^a \Gamma^\alpha_{\beta \delta} \delta^\beta \]
\[ \Rightarrow \nabla^\xi \delta^\alpha = \partial x^\alpha - \Gamma^\alpha_{\beta \delta} \delta^\beta \]
This generalises to tensors in the expected way
\[ \nabla_\xi \delta^\alpha \beta = \frac{\partial \delta^\alpha \beta}{\partial x^\xi} + \Gamma^\alpha_{\sigma \beta} \delta^\sigma \beta - \Gamma^\beta_{\sigma \gamma} \delta^\alpha \gamma \]
Other important properties of the covariant derivative,

1) $\nabla^\xi g_{\alpha \beta} = 0$ , i.e. cov der of metric vanishes -

This is clearly true in LT where $\delta x^\alpha = \delta x^\beta = 0$ and $\delta x^\sigma = 0$

so it's true in any other coordinate system. (you can check!)

2) Since the cov. der is constructed using parallel transported
tensor to vector, a vector is parallel propagated along
a curve if its cov der vanishes along curve

\[ \nabla_\xi \vec{V} = 0 \Rightarrow \vec{V} \text{ is parallel transported} \]
along $x^\alpha (\xi)$ with $x^\alpha = \frac{dx^\alpha}{d\xi}$
Construction of Freely Falling Frames

Inertial frames in Newtonian Mechanics are constructed by parallel transporting an initial choice for the 3 coordinate axes along the straight path of a free particle in flat space.

We can now do the same in GR for freely falling frames, i.e. a LIF all along a geodesic of an observer (the closest possible to global inertial frames in GR). The idea is very simple, we choose a set of orthonormal basis vectors \( \{ e_\alpha(\tau) \}_3 \) at each point along geodesic \( \gamma(\tau) \) as follows:

- The four velocity \( \dot{\gamma}(\tau) \) is the basis vector \( e_0 \) defining timelike direction.

- 3 mutually orthogonal basis vectors orthogonal to \( \dot{\gamma}(\tau) \) are picked at one point along geodesic to define the 3 spatial directions. The axes at other points along the geodesic is found by parallel propagating these vectors along \( \dot{\gamma}(\tau) \):

\[
\nabla_{\dot{\gamma}} e_\alpha = 0
\]

Note for \( \alpha = 1 \) this is automatically satisfied, \( \nabla_{\dot{\gamma}} \dot{\gamma} = 0 \) because of geodesic eqn., so only \( \alpha = 1, 2, 3 \) are non-trivial.