Derivatives in General Relativity

One of the problems with curved space is in dealing with vectors – how do you add a vector at one point in the surface of a sphere to a vector at a different point, and end up with a result that is still in the sphere? Even worse, how do you take the derivative of a vector in a spherical surface? That's easy in Euclidean space, because a vector at a point P can be naturally compared by simple translation to the vector at a nearby point Q. But in curved spacetime, we are trying to compare vectors at different points that may not even have the same *tangent basis vectors*! This document addresses this issue. In these sections, all the examples use 2-D or 3-D spaces (as opposed to a 4-D spacetime) to keep things simple and easier to visualize.

In 2-D spaces or higher, all derivatives require that you specify a direction; you can either take the direction along an axis (which gives you the partial derivative), or along a line, etc. The concept of transport is used for derivatives on curved surfaces : take a scalar, vector, or tensor (SVT) at one point and (somehow) transport it to the same place as another SVT. If the two SVTs are the same, the derivative between the two points is zero. Otherwise, by comparing the SVTs, you can calculate the derivative.

There are two basic types of derivatives for neighboring *tangent spaces* in general relativity : covariant derivatives and Lie derivatives. The **Lie derivative** evaluates the change of one SVT field along the direction of a separate vector field. A **covariant derivative** introduces an extra geometric structure (the affine connection, which requires a *metric tensor*) on the surface and evaluates the change in the SVT field using the connection.

Path Tangent Vectors

"Tangent vector" in general refers to any vector that is tangent to a curved surface at a point. However, when following along a worldline in spacetime, its **path tangent vector** points along the direction of travel.

The path tangent vector **T** to a parameterized curve $x^a(\lambda)$ is given by

$$T^a(\lambda) = dx^a / d\lambda$$

(where x^a represents position as a function of λ), so that (as usual) $\mathbf{T} = T^a\,\mathbf{e}_a$

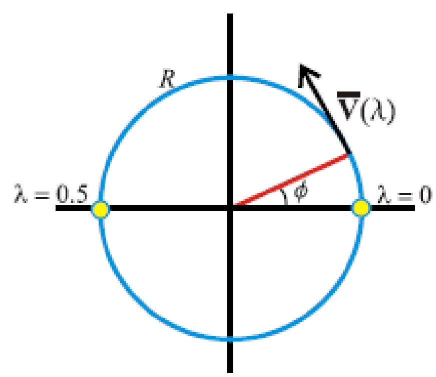
For example, let the path be a circle where R=1, parameterized by λ →
Then

$$x^1(\lambda) = x(\lambda) = \cos\lambda$$
$$x^2(\lambda) = y(\lambda) = \sin\lambda$$

And the path tangent vector is :

$$T^x = dx / d\lambda = -\sin\lambda$$
$$T^y = dy / d\lambda = \cos\lambda$$

For the figure to the right, $\lambda = 25°$, so **T** = (-0.4,0.9) as indicated by the black arrow.



φ → λ

Derivative of a Scalar Field

If $\mathbf{e}_i$ is the i^{th} basis vector and x^i is the i^{th} coordinate, then the gradient of a scalar field S=f(x,y,z) is a vector whose components are the derivative with respect to the coordinates :

$$\partial S/\partial x^i = \mathbf{e}_i\,\partial S/\partial x^i$$

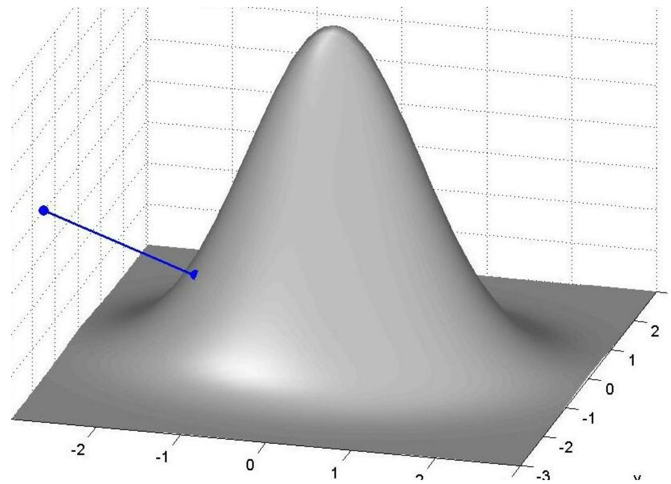
$= \mathbf{e}_1\, \partial S/\partial x^1 + \mathbf{e}_2\, \partial S/\partial x^2 + \mathbf{e}_3\, \partial S/\partial x^3$

$= (\partial S/\partial x^1, \partial S/\partial x^2, \partial S/\partial x^3)$ same thing in vector notation

$= (\partial S/\partial x, \partial S/\partial y, \partial S/\partial z)$ in Cartesian coordinates (for example)

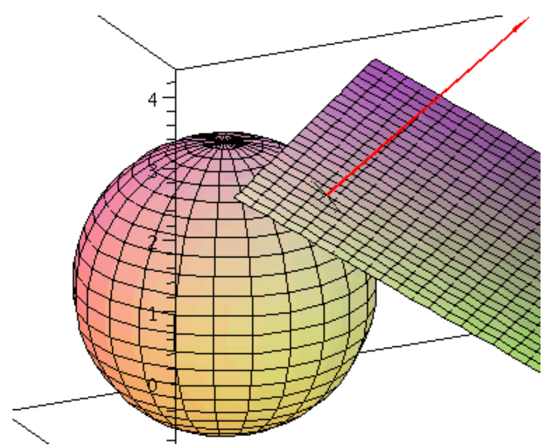
Note that "∂/∂_i" is always accompanied by an $\mathbf{e}_i$ whether it is shown or not!

In Euclidean space, this is the gradient. The gradient of a scalar field at point P is a vector pointing perpendicular to the surface at P. In the figure to the right, the surface S=f(x,y) and the blue line is the gradient $(\partial S/\partial x, \partial S/\partial y)$ at a specific point (x,y).



There are several notations used by various authors, and sometimes the $\mathbf{e}_i$ is assumed :

$$\partial_i S = S_{,i} = \nabla_i S = \partial S/\partial x^i = \mathbf{e}_i\, \partial S/\partial x^i$$



For example, the surface of a sphere with a radius of 4 can be described by $z = \pm(4 - x^2 - y^2)^{1/2}$, or in its more usual format, $S : x^2+y^2+z^2 = 16$. The normal to this function at the point P=(–1, 1, 1.41) is given by $\nabla_i S = (\partial S/\partial x, \partial S/\partial y, \partial S/\partial z) = (2x, 2y, 2z)$ which at P is the vector (–2, 2, 2.82)

In 4-D spacetime, due to the negative sign on the spatial coordinates, and the (sometimes implied) presence of *c* on the time coordinate, the correct definitions of the four-gradient are :

$$\partial_\alpha = \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right) \qquad \partial^\alpha = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right)$$

Where (as usual) ∂^α could also be calculated by $g^{\mu\alpha}\partial_\mu$

The divergence of the gradient is called the **Laplacian**. The Laplacian of a scalar field is a scalar :

$$\nabla\cdot\nabla S = \mathbf{e}_i \cdot \partial(\mathbf{e}_j\, \partial S/\partial x^j)/\partial x^i = \mathbf{e}_i \cdot \mathbf{e}_j\, \partial^2 S/\partial x^j \partial x^i = g^{ij} S_{,ij}$$

And also goes by several different notations : $\nabla\cdot\nabla = \nabla^2 = \Delta = \nabla_i \nabla^i = \partial_i \partial^i = \partial^2$

The Laplacian of a vector is a vector :

$$\nabla^2 \mathbf{V} = g^{jk} V^i_{,jk}\, \mathbf{e}_i$$

Directional Scalar Derivative

The derivative along the coordinate axes can be generalized to a **directional derivative of a scalar field** in the direction of any unit vector **u**. The directional derivative of a scalar field S is :

$$(\nabla_i S) \cdot \mathbf{u} = u_i (\partial S/\partial x^i)$$
$$= u_x \partial S/\partial x + u_y \partial S/\partial y + u_z \partial S/\partial z \quad \text{in Cartesian coordinates}$$

This result is a scalar, which is the magnitude of the rate of change of the scalar field in the direction of **u**.

If the vector is the path tangent vector $\mathbf{T}(\lambda)$ of a path $x^i(\lambda)$, then the derivative of S along the path is :

$$\nabla_T S = T^i \nabla_i S = T^i (\partial S/\partial x^i) = (\partial S/\partial x^i)(\partial x^i/\partial\lambda) = \partial S/\partial\lambda$$

which is just the rate of change of the scalar field as you travel along the path.

Derivative of a Vector Field

The derivative of a vector field $\mathbf{V} = V^\alpha \mathbf{e}_\alpha$ with respect to coordinate x^β in *Euclidean space with Cartesian coordinates* is :

$$\partial\mathbf{V}/\partial x^\beta = \partial(V^\alpha \mathbf{e}_\alpha)/\partial x^\beta = \mathbf{e}_\alpha \partial V^\alpha/\partial x^\beta$$
$$= (\partial V^1/\partial x^\beta, \partial V^2/\partial x^\beta, \partial V^3/\partial x^\beta) \qquad \text{for all } \beta$$

since the basis vectors do not vary and so $\partial\mathbf{e}_\alpha/\partial x^\beta = 0$. This is the **ordinary partial derivative**, but virtually all authors leave out the $\mathbf{e}^\beta$ that goes along with the ∂/∂_β :

$$\partial\mathbf{V}/\partial x^\beta = \partial(V^\alpha \mathbf{e}_\alpha)/\partial x^\beta \otimes \mathbf{e}^\beta = \partial V^\alpha/\partial x^\beta\, (\mathbf{e}_\alpha \otimes \mathbf{e}^\beta)$$

Which means that $\partial\mathbf{V}/\partial x^\beta$ is really a **tensor!** For example, if $\mathbf{V} = (V^x(x,y,z), V^y(x,y,z), V^z(x,y,z))$:

$$\partial\mathbf{V}/\partial x = \begin{array}{ll} \left|\begin{array}{ccc} \partial V^x/\partial x & \partial V^y/\partial x & \partial V^z/\partial x \\ \partial V^x/\partial y & \partial V^y/\partial y & \partial V^z/\partial y \\ \partial V^x/\partial z & \partial V^y/\partial z & \partial V^z/\partial z \end{array}\right| & \begin{array}{l} \leftarrow \partial\mathbf{V}/\partial x^1 \\ \leftarrow \partial\mathbf{V}/\partial x^2 \\ \leftarrow \partial\mathbf{V}/\partial x^3 \end{array} \end{array}$$

For the remaining sections, the $\mathbf{e}^\beta$ term will also be left out for simplicity, but it should always be kept in mind that a term like $\partial V^\alpha/\partial x^\beta$ implies a tensor, because there are **two** non-summed indexes.

There are several different notations for the ordinary partial derivative as well. And since index notation shows a vector **V** by its components V^α, the notation "$\partial\mathbf{V}/\partial x^\beta$" is not actually used. Again, sometimes the $\mathbf{e}_\alpha$ is merely implied :

$$V^\alpha{}_{,\beta} = \partial_\beta V^\alpha = \partial V^\alpha/\partial x^\beta = \mathbf{e}_\alpha \partial V^\alpha/\partial x^\beta$$

Affine Connection

Informally, an **affine space** is a vector space without an origin, so points in an affine space cannot be added together. However, a vector **V** may be added to a point P by placing the initial point of the vector at P and then *transporting* P to the end of the vector.

An **affine connection** is a geometrical object which connects nearby *tangent spaces* in a surface (like taping two maps together along a common edge), and so permits *tangent vector* fields to be differentiated. The choice of an affine connection is equivalent to describing a way of differentiating vector fields which satisfies several reasonable properties (such as linearity and the product rule). An affine connection also defines the idea of *parallel transport* and *geodesics*. Conversely, specifying a method of parallel transport defines the connection. So affine connections are intimately tied together with all the concepts we have been discussing so far. The main invariant measures of an affine connection are its torsion and its curvature.

There are infinitely many affine connections on a surface, but if the surface has a *metric* then there is a natural choice of affine connection, called the **Levi-Civita connection**, which has zero torsion. Then the value of the parameter used to describe a worldline is proportional to the worldline's path length (and also happens to represent its *proper time* if the worldline is *timelike*). [any worldline or only geodesics?]

In local coordinates, the components of the Levi-Civita connection are called Christoffel symbols (next section), and because the Levi-Civita connection is unique, they can be calculated directly from the metric. For this reason, this type of connection is often called a **metric connection**.

It should be noted that the zero-torsion condition is an assumption of general relativity, and there are extensions to general relativity which allow for torsion.

Christoffel Symbols

In any *curvilinear coordinate system* **or** *non-Euclidean space*, the basis vectors change orientation with position, so the derivative of a vector field $\mathbf{V} = V^\alpha \mathbf{e}_\alpha$ becomes :

$$\partial \mathbf{V}/\partial x^\beta = \partial(V^\alpha \mathbf{e}_\alpha)/\partial x^\beta = (\partial V^\alpha/\partial x^\beta)\mathbf{e}_\alpha + V^\alpha(\partial \mathbf{e}_\alpha/\partial x^\beta)$$

Since $\partial \mathbf{e}_\alpha/\partial x^\beta$ is itself a vector, it can be written as some linear combination of the basis vectors :

$$\partial \mathbf{e}_\alpha/\partial x^\beta = \Gamma^\mu{}_{\alpha\beta}\, \mathbf{e}_\mu$$

$$= \Gamma^x{}_{\alpha\beta}\mathbf{e}_x + \Gamma^y{}_{\alpha\beta}\mathbf{e}_y + \Gamma^z{}_{\alpha\beta}\mathbf{e}_z \qquad \text{in Cartesian coordinates, for example}$$

Γ is called a **Christoffel symbol** (of the second kind), **affine connection**, or **metric connection**. It provides the "weights" or coefficients to create $\partial \mathbf{e}_\alpha/\partial x_\beta$ from the basis vectors $\mathbf{e}_i$. It is sometimes written as $\left\{ {\sigma \atop \mu\nu} \right\}$

In terms of the covariant basis vectors $\mathbf{e}_i$ and contravariant basis vectors $\mathbf{e}^i$ the Christoffel symbol is defined as :

$$\Gamma^i{}_{jk} = \mathbf{e}^i \cdot (\partial \mathbf{e}_j/\partial x^k) \qquad \text{("the } i^{th} \text{ component of the } j^{th} \text{ basis wrt the } k^{th} \text{ axis")}$$

where "$\mathbf{e}^i \cdot$ " means to take the i^{th} component of what follows (= dot product along the $\mathbf{e}^i$ axis).

The Christoffel symbol measures how the *basis vectors* change as we move in a particular direction. Once we know how the basis vectors change, then we can use this information to correct the ordinary derivative which we obtained using that basis, in order to obtain the “covariant” derivative in a curved surface (next section).

The Christoffel symbol can also be defined in terms of the metric tensor :

$$\Gamma^a{}_{bc} = g^{ai} (\partial_b g_{ci} + \partial_c g_{ib} - \partial_i g_{bc}) / 2$$

Note that there can be **many** variations on this formula due to the nature of the Einstein notation, the multiple formats for the derivatives, and the symmetry of the metric tensor!

In addition, the Riemann tensor described in GR1c can now be given in terms of the Christoffel symbol :

$$R^\ell{}_{ijk} = \frac{\partial}{\partial x^j}\Gamma^\ell_{ik} - \frac{\partial}{\partial x^k}\Gamma^\ell_{ij} + \Gamma^\ell_{js}\Gamma^s_{ik} - \Gamma^\ell_{ks}\Gamma^s_{ij}$$

And there are many variations on this formula as well, for all the same reasons, and the symmetry of the bottom indexes in Γ (see below), as well as the choice of contravariant/covariant indexes for R.

Things to know about Christoffel symbols :

Christoffel symbols are “in” the surface. Despite the notation used to describe them, Christoffel symbols are not actually tensors.

Christoffel symbols appear when either the space is curved, or when non-Cartesian coordinates are used in a Euclidean space.

If torsion is zero (which is assumed in general relativity), the Christoffel symbol is symmetric in its bottom indexes, so $\Gamma^i{}_{jk} = \Gamma^i{}_{kj}$ which means, for example :

$\Gamma^\alpha{}_{\gamma\beta}V^\gamma X^\beta$ is the exact thing same as $\Gamma^\alpha{}_{\beta\gamma}V^\beta X^\gamma$ is exactly the same as $\Gamma^\alpha{}_{\beta\gamma}V^\gamma X^\beta$

In Euclidean space with Cartesian coordinates, $\Gamma^i{}_{jk} = 0$ for all i,j,k because the basis vectors never change. Christoffel symbols are also all zero in the *tangent plane* (which is a local inertial frame and so a Minkowski spacetime) because the basis vectors do not change.

Example

In Euclidean 2-D polar coordinates (r,θ) the non-zero Christoffel symbols are :

$\Gamma^\theta{}_{r\theta} = 1/r$ $\Gamma^\theta{}_{\theta r} = 1/r$ $\Gamma^r{}_{\theta\theta} = -r$

And so the coordinate derivatives $\partial \mathbf{e}_\alpha/\partial x^\beta = \Gamma^\mu{}_{\alpha\beta}\, \mathbf{e}_\mu$ are :

$\partial\mathbf{e}_r/\partial r = \Gamma^r{}_{rr}\mathbf{e}_r + \Gamma^\theta{}_{rr}\mathbf{e}_\theta$ $\partial\mathbf{e}_r/\partial\theta = \Gamma^r{}_{r\theta}\mathbf{e}_r + \Gamma^\theta{}_{r\theta}\mathbf{e}_\theta$

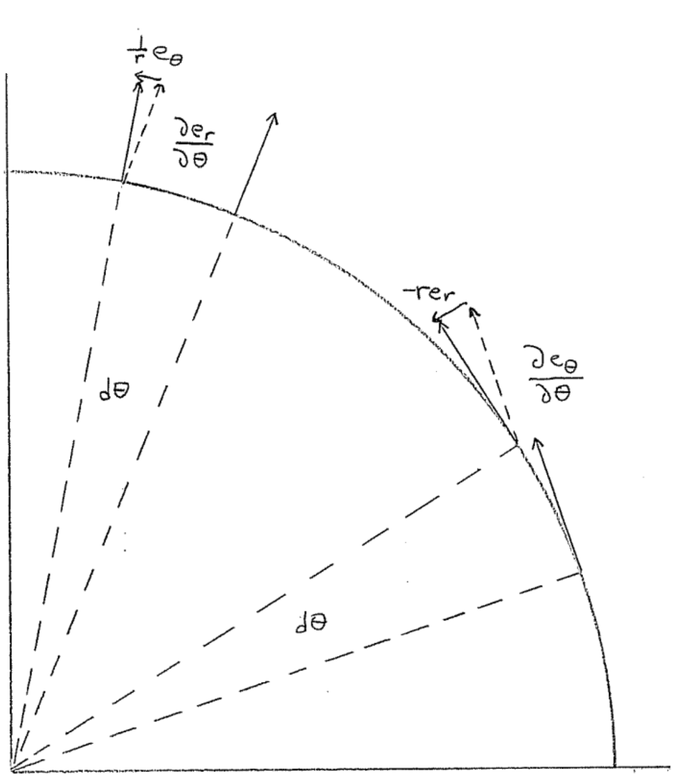
$\partial\mathbf{e}_\theta/\partial r = \Gamma^r{}_{\theta r}\mathbf{e}_r + \Gamma^\theta{}_{\theta r}\mathbf{e}_\theta$ $\partial\mathbf{e}_\theta/\partial\theta = \Gamma^r{}_{\theta\theta}\mathbf{e}_r + \Gamma^\theta{}_{\theta\theta}\mathbf{e}_\theta$

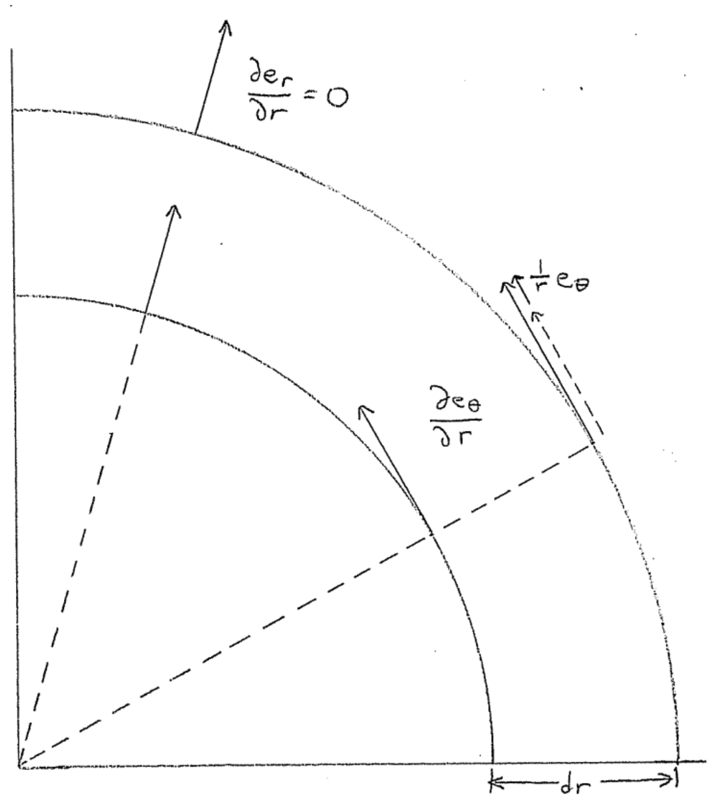
Which simplifies to :

$\partial\mathbf{e}_r/\partial r = 0$ $\partial\mathbf{e}_r/\partial\theta = \mathbf{e}_\theta / r$

$\partial\mathbf{e}_\theta/\partial r = \mathbf{e}_\theta / r$ $\partial\mathbf{e}_\theta/\partial\theta = -r\ \mathbf{e}_r$

The figure below shows what each coordinate derivative means. The vector from the smaller value of r or θ is copied (as a dotted-line) over to the vector at the larger value. The difference between them (from the head of the dotted-line vector to the head of the solid-line vector) is $\partial\mathbf{e}_\alpha/\partial x^\beta$.





For reference, in Euclidean 3-D spherical coordinates [r,φ,θ], the non-zero Christoffel symbols are :

$$\Gamma^r{}_{\theta\theta} = -r \qquad \Gamma^r{}_{\varphi\varphi} = -r \sin^2\varphi \qquad \Gamma^\theta{}_{\varphi\varphi} = -\sin\theta \cos\varphi$$
$$\Gamma^\theta{}_{r\theta} = \Gamma^\theta{}_{\theta r} = 1/r \qquad \Gamma^\varphi{}_{r\varphi} = \Gamma^\varphi{}_{\varphi r} = 1/r \qquad \Gamma^\varphi{}_{\varphi\theta} = \Gamma^\varphi{}_{\theta\varphi} = \cot\varphi$$

Covariant Derivative

In curved space, the **covariant derivative** is the change in the vector itself, plus the changes in the basis vectors. If M is a surface with *curvilinear coordinates* x^i, and **V** is a *tangent* vector field to M at a point P, then (from the previous section) the derivative of **V** along the x^β axis is :

$$\partial \mathbf{V}/\partial x^\beta = \partial(V^\alpha \mathbf{e}_\alpha)/\partial x^\beta = (\partial V^\alpha/\partial x^\beta)\mathbf{e}_\alpha + V^\alpha(\partial \mathbf{e}_\alpha/\partial x^\beta)$$

But as we have seen, $\partial \mathbf{e}_\alpha/\partial x^\beta = \Gamma^\mu{}_{\alpha\beta}\mathbf{e}_\mu$

So we have :

$$\partial \mathbf{V}/\partial x^\beta = (\partial V^\alpha/\partial x^\beta)\mathbf{e}_\alpha + V^\alpha \Gamma^\mu{}_{\alpha\beta}\mathbf{e}_\mu$$

Or (with a change to the dummy indexes in the second term, α ↔ μ) :

$$\partial \mathbf{V}/\partial x^\beta = (\partial V^\alpha/\partial x^\beta + V^\mu \Gamma^\alpha{}_{\mu\beta})\mathbf{e}_\alpha$$

Which in Cartesian coordinates (for example) becomes :

$$\begin{aligned}\partial \mathbf{V}/\partial x^\beta = &(\partial V^x/\partial x^\beta + V^x\Gamma^x{}_{x\beta} + V^y\Gamma^x{}_{y\beta} + V^z\Gamma^x{}_{z\beta})\mathbf{e}_x \\ &+ (\partial V^y/\partial x^\beta + V^x\Gamma^y{}_{x\beta} + V^y\Gamma^y{}_{y\beta} + V^z\Gamma^y{}_{z\beta})\mathbf{e}_y \\ &+ (\partial V^z/\partial x^\beta + V^x\Gamma^z{}_{x\beta} + V^y\Gamma^z{}_{y\beta} + V^z\Gamma^z{}_{z\beta})\mathbf{e}_z\end{aligned}$$

and is really three equations, one each for x^β = x, y, z (which means it is a tensor!). This example is just to show how the notation works, since in Euclidean space with Cartesian coordinates every $\Gamma^i{}_{jk}$ is zero.

The symbol ∇_β is often used to represent the covariant derivative "in the x^β direction" or "along the basis vector $\mathbf{e}_\beta$". And since index notation uses the *components* of the vector **V**, the most common notation is :

$$\nabla_\beta V^\alpha = \partial V^\alpha/\partial x^\beta + \Gamma^\alpha{}_{\mu\beta} V^\mu$$

which returns only the <u>α-component of the entire answer</u> $(\nabla_\beta V^1, \nabla_\beta V^2, \nabla_\beta V^3)$, so some authors use $(\nabla_\beta V)^\alpha$. And there is yet another notation, using just a semicolon : $V^\alpha{}_{;\beta}$

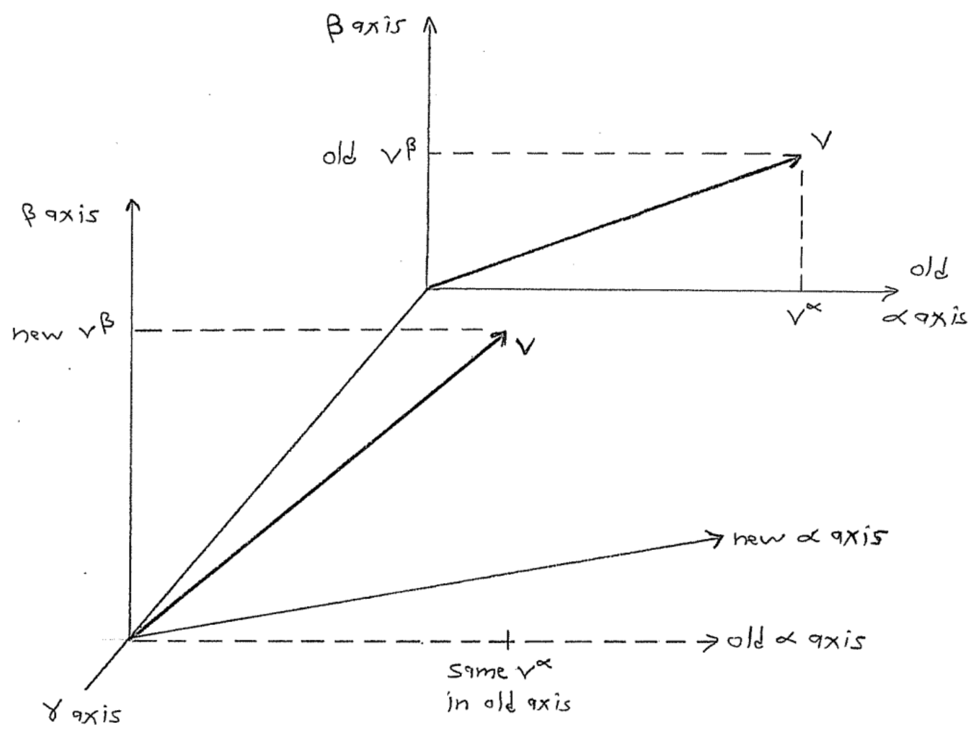
Summing up, the covariant derivative can be expressed in any of the following ways :

$$\begin{aligned} V^\alpha{}_{;\beta} = \nabla_\beta V^\alpha = (\nabla_\beta V)^\alpha &= (\partial V^\alpha/\partial x^\beta)\mathbf{e}_\alpha + V^\alpha \Gamma^\mu{}_{\alpha\beta}\mathbf{e}_\mu \\ &= (\partial V^\alpha/\partial x^\beta + V^\mu \Gamma^\alpha{}_{\mu\beta})\mathbf{e}_\alpha \\ &= \partial V^\alpha/\partial x^\beta + \Gamma^\alpha{}_{\mu\beta} V^\mu \\ &= \partial_\beta V^\alpha + V^\mu \Gamma^\alpha{}_{\mu\beta} \\ &= V^\alpha{}_{,\beta} + \Gamma^\alpha{}_{\mu\beta} V^\mu \end{aligned}$$

All of which stand for "the amount of change in the α component of the covariant derivative of vector V when we move in the x^β direction".

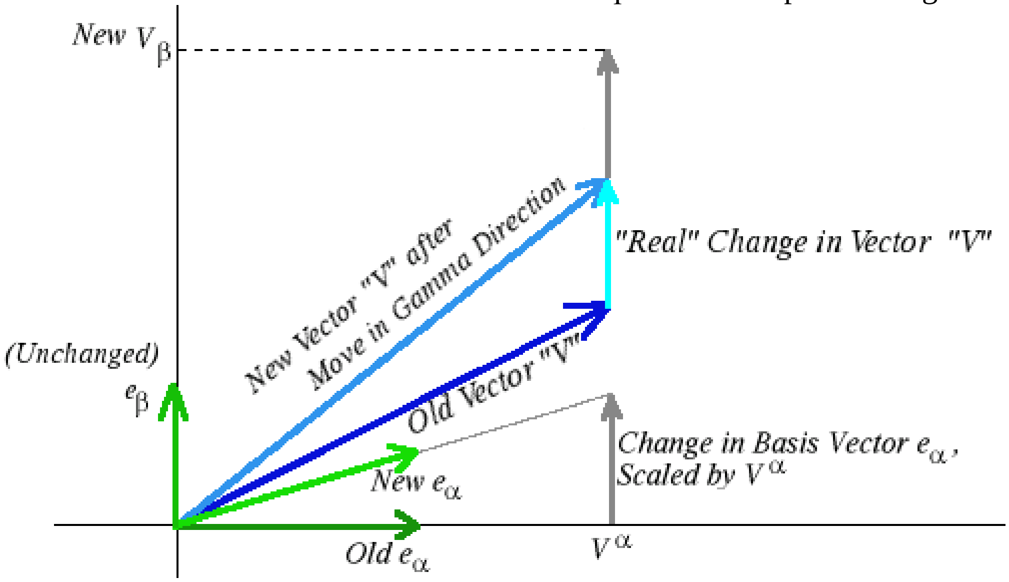
And just to make things more confusing, "D" is also sometimes used for ∇, as in D(V) or $D_\beta V$. This actually comes in handy when you want to take the covariant derivative with respect to a path parameter, such as D/Dλ. In addition, some authors use D to indicate only the 3-D <u>spatial</u> covariant derivatives.

What does all this mean? Let's say we want to find the amount of change in the β component of the covariant derivative of a vector **v** when we move in the x^γ direction, or $(\nabla_\gamma v)^\beta$:



Since **v** is a function of its position in spacetime, the v^β component will change just because we're moving in the x^γ direction (in the figure above, the new v^β is larger than the old v^β). The amount of this change is $\partial v^\beta/\partial x^\gamma$ (this is just the ordinary partial derivative), and is the "real" change in the picture below.

Now look at the other basis vector $\mathbf{e}_\alpha$ ("α axis") in the figure above. As we move in the x^γ direction, $\mathbf{e}_\alpha$ changes because the space is not flat. We also need to know how much $\mathbf{e}_\alpha$ changes *in the* β *direction*. The following figure combines information from both the "old" and the "new" planes in the previous figure :



So we need to add the change in $\mathbf{e}_\alpha$ to the "real" change in v^β, <u>and</u> we need to scale the change in $\mathbf{e}_\alpha$ by [the part of $\mathbf{v}$ which lies along the direction of $\mathbf{e}_\alpha$ ($= v^\alpha$)]. In other words, v^α contributes a change to v^β which is proportional to v^α times [the component in the β direction of the change in $\mathbf{e}_\alpha$ as we move in the x^γ direction]. This scaling factor in the brackets is just $\mathbf{e}^\beta \cdot \partial\mathbf{e}_\alpha/\partial x^\gamma \equiv \Gamma^\beta{}_{\alpha\gamma}$. So the "extra" change due to $\mathbf{e}_\alpha$ is $\Gamma^\beta{}_{\alpha\gamma}v^\alpha$ (for the *specific values* of α, β, γ – no summing!).

Now, the change to the $\mathbf{e}_\beta$ basis vector will contribute $\Gamma^\beta{}_{\beta\gamma}v^\beta$ and the change to $\mathbf{e}_\gamma$ will contribute $\Gamma^\beta{}_{\gamma\gamma}v^\gamma$, so summing over all of them with a dummy index μ gives us :

$$(\nabla_\gamma V)^\beta = \partial V^\beta/\partial x^\gamma + \Gamma^\beta{}_{\mu\gamma}V^\mu$$

QED (which is Latin for "ta-daa!!!")

The covariant derivative turns a type (r,s) tensor into a type (r,s+1) tensor, increasing the rank by one. So for the above examples, it turns a (1,0) contravariant vector into a (1,1) tensor of rank 2.

In general relativity, the covariant derivative is associated with the Levi-Civita affine connection. By definition, the Levi-Civita connection preserves the metric under parallel transport, therefore, the covariant derivative is zero when acting on the metric tensor (as well as its inverse). This means we can take the metric tensor (or its inverse) in and out of the derivative and use it to raise and lower indices :

$$\nabla_\beta V^\alpha = \nabla_\beta(g^{\alpha\gamma}V_\gamma) = g^{\alpha\gamma}\nabla_\beta V_\gamma$$

And it can be used to take the "contravariant derivative" :

$$\nabla^\alpha V = g^{\alpha\beta}\nabla_\beta$$

The covariant derivative can be extended to tensors of any rank –

For a type (1,0) contravariant vector field V^a : $\qquad \nabla_c V^a = \partial_c V^a + \Gamma^a{}_{ic}V^i$

For a type (0,1) covariant vector field V_a : $\qquad \nabla_c V_a = \partial_c V_a - \Gamma^i{}_{ac}V_i$

For a type (2,0) tensor field T^{ab} : $\nabla_c T^{ab} = \partial_c T^{ab} + \Gamma^a{}_{ic} T^{ib} + \Gamma^b{}_{ic} T^{ai}$

For a type (0,2) tensor field T_{ab} : $\nabla_c T_{ab} = \partial_c T_{ab} - \Gamma^i{}_{ac} T_{ib} - \Gamma^i{}_{bc} T_{ai}$

For a type (1,1) tensor field $T^a{}_b$: $\nabla_c T^a{}_b = \partial_c T^a{}_b + \Gamma^a{}_{ic} T^i{}_b - \Gamma^i{}_{bc} T^a{}_i$

For a type (1,1) tensor field $T_a{}^b$: $\nabla_c T_a{}^b = \partial_c T_a{}^b - \Gamma^i{}_{ac} T_i{}^b + \Gamma^b{}_{ic} T_a{}^i$

The pattern is : take the ordinary derivative, then add a Γ term for each contravariant letter and subtract a Γ term for each covariant letter.

Things to know about covariant derivatives :

Covariant derivatives are "in" the surface. The covariant derivative in Euclidean space with Cartesian coordinates is the same as the ordinary partial derivative, because the basis vectors never change.

The covariant derivative of a scalar field V is the ordinary partial derivative : $\nabla_\alpha V = \partial V/\partial x^\alpha$

The covariant derivative of a vector field $\nabla_\beta V^\alpha = (\nabla_\beta V^1, \nabla_\beta V^2, \nabla_\beta V^3)$ is the "partial derivative" of **V** in the $\mathbf{e}^\beta$ direction, and is often equated to the divergence (except that divergence returns a scalar value!).

The covariant derivative of the metric tensor is zero everywhere on the surface : $\nabla_c\, g_{ab} = 0$

Note that $\nabla_\alpha V^\beta$ is a covariant derivative (because V^β implies a vector), while $\nabla_\alpha V$ is the gradient of the scalar V (no indexes).

Directional Covariant Derivative

We can generalize the covariant derivative by taking the derivative of **V** in the direction of another vector **T** (for example, the path tangent vector to some path $x^\alpha(\lambda)$). As before, let M be a surface, let **V** be a tangent vector to M at a point P, and let $\mathbf{T}(\lambda)$ be a **path** tangent vector at P. Then the directional covariant derivative (which is *independent of the curve used to create* **T**), is

$$\nabla_T V^\alpha = T^\beta \nabla_\beta T^\alpha = T^\beta\,(\partial V^\alpha/\partial x^\beta + \Gamma^\alpha{}_{\beta\gamma} V^\gamma)\,\mathbf{e}_\alpha$$

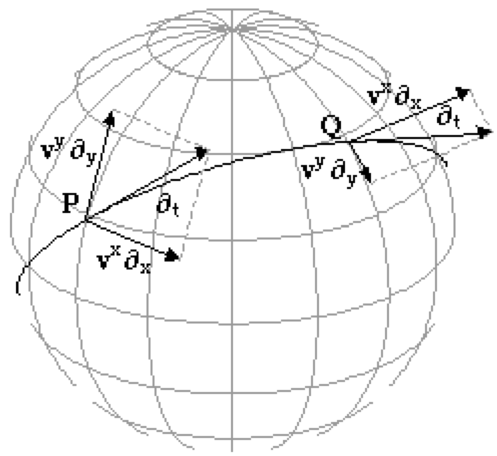
The directional covariant derivative of a type (r,s) tensor is also of type (r,s).

When a vector **V** moves along a path $x^\alpha(\lambda)$ that has path tangent vectors $\mathbf{T}(\lambda)$, and $\nabla_T V^\alpha = 0$ for all α, then V was **parallel transported** along the path. This is equivalent to moving **V** along the path so that it remains parallel to itself and does not change length.

When the vector field **V** is described in terms of a parameter λ, the covariant derivative becomes :

$$\nabla_\lambda V^\alpha = (\nabla_\gamma V^\alpha)(\partial x^\gamma/\partial\lambda)$$
$$= (\partial V^\alpha/\partial x^\gamma + \Gamma^\alpha{}_{\beta\gamma} V^\beta)\,(\partial x^\gamma/\partial\lambda) = \partial V^\alpha/\partial\lambda + \Gamma^\alpha{}_{\beta\gamma} V^\beta(\partial x^\gamma/\partial\lambda)$$

A path that parallel-transports its own path tangent vector $\mathbf{T}(\lambda)$, $\nabla_T T^\alpha = 0$, is a **geodesic**. This means that in addition to the path tangent vector remaining parallel to itself with constant length, the path appears straight within the surface. The figure below shows a couple *path tangent vectors* of a geodesic, with their components along the *tangent basis vectors* $\mathbf{e}_x = \partial_x$ and $\mathbf{e}_y = \partial_y$ (note how they line up with their local lines of latitude and longitude) :



It is now possible to give a mathematical definition of a non-rotating inertial frame :

$$\nabla_{\mathbf{e}_o}\mathbf{e}_i = 0$$

In particular, i=0 implies that there is no acceleration (= inertial), and i=1…3 implies non-rotation.

<u>Lie Derivative and Lie Bracket</u>

While the covariant derivative requires an affine connection for transport, the Lie derivative uses the <u>flow of another vector field</u> to describe the direction of travel.

If a vector field V is defined by a set of differential equations, then its <u>congruence of curves</u> (or just "congruence") is the family of solutions (= the integrals of the equations), and is often parameterized. A single path in this congruence is a <u>trajectory</u>, whose tangent vector at each point is in turn the value of V at that point. The collection of tangent vectors (across all trajectories) is called a <u>flow</u>.

For example, if V is defined by

$\partial V_x/\partial\theta = -r \sin\theta$ $\quad\quad$ $\partial V_y/\partial\theta = r \cos\theta$

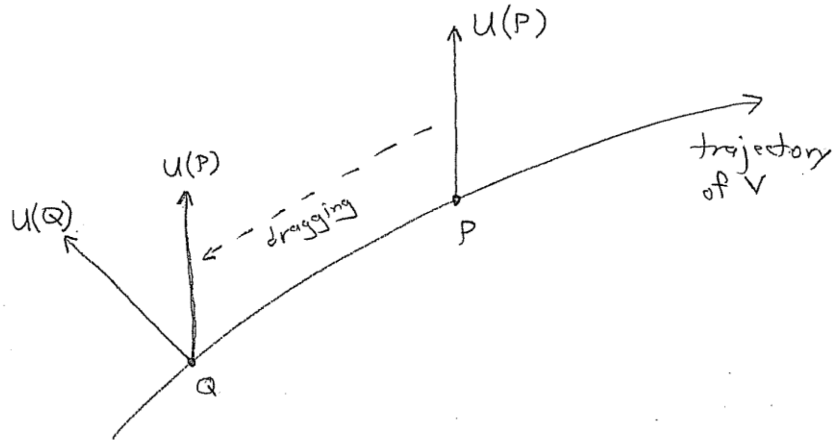
Then the solution to V is

$V_x = r \cos\theta$ $\quad\quad$ $V_y = r \sin\theta$

where r and θ are parameters. The last two equations result in a series of concentric circles as you vary r and θ. → The concentric circles are the trajectories, the tangent vectors (some of which are shown in blue) are defined by (–r sinθ, r cosθ), and the set of all tangent vectors is the flow.

A **congruence** in spacetime is a <u>family of world lines.</u> One example of a *timelike geodesic congruence* is the trajectories of many free-falling test particles. A *null geodesic congruence* could be a family of propagating light rays. But not all congruences <u>have</u> to be geodesics.

The idea of <u>Lie dragging</u> a function along a congruence leads to a definition of the Lie derivative. Consider a small coordinate change in the surface that results from dragging along a trajectory of the vector field V, using its tangent vector at point P to point the way. After moving a small distance ε along the vector, we are now at point *Q*. Let U be some "object" (scalar/vector/tensor field) defined on the surface. There are two values we can consider at Q : U evaluated at Q, and the value that results from dragging the value of U at P to the point Q thru the coordinate system. Since the surface is not flat, we need to subtract out coordinate transformation to get the actual change in U (as was true of the covariant derivative as well). These two different values at Q allow us to form the Lie derivative at Q.



The Lie derivative is usually denoted by L_V where V is the vector field along whose congruence the Lie derivative is taken. The Lie derivative takes a tensor of type (r,s) and creates a tensor of type (r,s).

The Lie derivative **of a <u>scalar</u> field** S=f(x,y,z) is just the directional derivative of *f* along the vector field V :

$$L_V S = V^a S_{,a} = V^a \partial_a S = V^a (\partial S/\partial x^a)$$

The Lie derivative **of a (contravariant) <u>vector</u> field** U is :

$$L_V U^a = V^i U^a{}_{,i} - U^i V^a{}_{,i} = V^i (\partial U^a/\partial x^i) - U^i (\partial V^a/\partial x^i)$$

and if U is covariant, just turn all the superscripts on U into subscripts.

Another notation is the <u>Lie bracket</u> [X,Y] which is the Lie derivative of a pair of vector fields X and Y acting on a function f :

$$[X,Y](f) = [X^b(\partial Y^a/\partial x^b) - Y^b(\partial X^a/\partial x^b)]\, \partial f/\partial x^a$$

Note : a set of *holonomic* basis vectors $\mathbf{e}_i$ satisfies $[\mathbf{e}_i,\mathbf{e}_j]=0$

The Lie derivative **of a tensor <u>T</u>** is given by :

$$L_V T_{ab} = T_{ai} V^i{}_{,b} + T_{ib} V^i{}_{,a} + V^i T_{ab,i}$$
$$L_V T_{ab} = T_{ai}(\partial V^i/\partial x^b) + T_{ib}(\partial V^i/\partial x^a) + V^i(\partial T_{ab}/\partial x^i)$$

and the Lie derivative **of the metric tensor** is :

$$L_V g_{ab} = \nabla_b V_a + \nabla_a V_b$$

If the Lie derivative of an object is everywhere zero, the object possesses a symmetry : $L_V T = 0$ means that T's structure is constant along the flows of V. This leads to the concept of Killing vectors, which are discussed in GR3x.