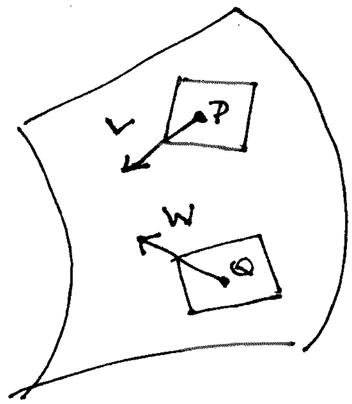
The Connection and Covariant Derivative

On a curved manifold, the tangent spaces at different points are distinct vector spaces. This is intuitively clear if we think of the manifold as imbedded in a higher dimensional space, since the tangent planes at different points are distinct and in general are not parallel to each other.



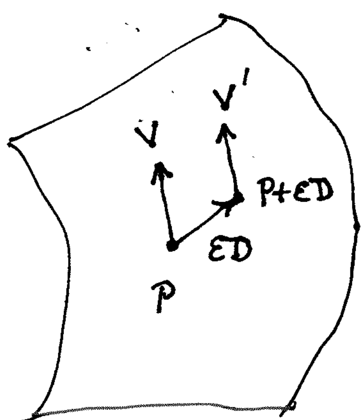
tangent space at P $\neq$ tangent space at Q

On a curved manifold, we must think of a different tangent space at each point. Each tangent space can intuitively be thought of as consisting of small ~~displacements~~ vectors based at the given point; more rigorously (but less pictorially) the vectors at a point P can be thought of as first-order, partial differential operators acting at P.

In general there is no way to take a vector V at P and another W at Q and say whether they are "the same" vector. If we have a coordinate system x^μ we can look at the components V^μ and W^μ, but if $V^\mu = W^\mu$ in one coordinate system then it won't be true in another, so comparing components of vectors is not useful.

To compare vectors at two different points of a curved manifold, we must introduce some geometrical structure that defines how this is done. This geometrical structure is called a <u>connection</u>. We do not attempt to identify vectors in distant tangent spaces; a connection only tells us how to identify vectors in nearby (infinitesimally close) tangent spaces.

Let P be a point of a manifold, and let εD be a displacement to a nearby point we will call "$P+\varepsilon D$".



The ε serves a psychological purpose, to remind us that the vector εD is small. This is using the intuitive idea that vectors are infinitesimal displacements at a point. The symbol "D" is a mnemonic for "displacement". Note that εD is a vector based at P.

Now let V be another vector based at P. We wish to define a vector V' based at $P+\varepsilon D$ that is to be considered "the same" as V, based at P. We require that V' be a linear function of V, and that it depend linearly on the displacement εD.

Let x^μ be a coordinate system in a neighborhood of P, and let V (at P) and V' (at $P+\varepsilon D$) have components V^μ and V'^μ. Since V'^μ is linear function of V^μ, there must be some matrix $M^{\mu}_{\cdot\nu}$ such that $V'^{\mu} = M^{\mu}_{\cdot\nu} V^{\nu}$; the matrix must be close to the identity, $M^{\mu}_{\cdot\nu} = \delta^{\mu}_{\nu} + C^{\mu}_{\cdot\nu}$ where the small correction $C^{\mu}_{\cdot\nu}$ is linear in εD and therefore proportional to ε. That is there must be some coefficients $A^{\mu}_{\alpha\nu}$ such that

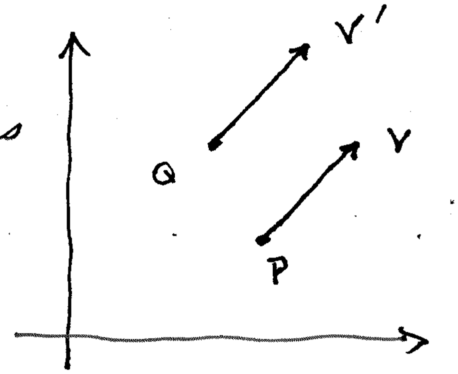
conventional minus sign

$$V'^{\mu} = V^{\mu} - \varepsilon A^{\mu}_{\alpha\nu} D^{\alpha} V^{\nu}. \qquad (1)$$

The coefficients $A^{\mu}_{\alpha\nu}$, which depend on x, specify the connection. Different choices of $A^{\mu}_{\alpha\nu}$ give different rules for identifying vectors at neighboring points.

Before going on we should mention that the situation described so far simplifies if the manifold is a vector space, such as the $\mathbb{R}^4$ space-time used in special relativity, with coordinates (t, x, y, z). In a vector space we can identify vectors in two distinct tangent spaces, just by parallel-transporting a vector from one tangent space to another, using the vector space structure.

In fact, in a vector space we usually identify all tangent spaces at all points, and consider them to be a single vector space, just by moving the base points of vectors; and we usually identify this global tangent space with the original manifold (itself a vector space). With this understanding, the connection



$V = V'$ in this example.

coefficients $A^{\mu}_{\alpha\nu}$ in (1) are zero, if we use linear coordinates on our vector space (for example, the usual (t, x, y, z) coordinates in special relativity). However, if curvilinear coordinates are used on a vector space, then the coefficients $A^{\mu}_{\alpha\nu}$ are generally nonzero. The case in which the space is the plane $\mathbb{R}^2$ but polar coordinates (r, ϕ) are used is discussed in the book.

In general relativity where the manifolds are not vector spaces there is usually one connection that is used. As above, let P and P+εD be nearby points, and let V (at P) and V′ (at P+εD) be two vectors. We introduce Riemann normal coordinates at P, call these $\bar{x}^{\mu}$. The overbar means R.N.C.; coordinates x^{μ} (without the overbar) are an arbitrary set of coordinates. Let the components of V and V′ in RNC. be $\bar{V}^{\mu}$ and $\bar{V}'^{\mu}$.

Now R.N.C. are coordinates that come as close as we can come to Minkowski coordinates (a Lorentz frame) in special relativity, in a small neighborhood of P. In particular, at P we have

$$\left.\begin{aligned} \bar{g}_{\mu\nu} &= \eta_{\mu\nu} \\ \frac{\partial \bar{g}_{\mu\nu}}{\partial \bar{x}^{\alpha}} &= 0 \end{aligned}\right\} \text{ at } P. \qquad (2)$$

In flat space (special relativity) two vectors at different points are considered equal if they have the same components in a Lorentz frame, with its rectilinear coordinates. Therefore we declare that V (at P) and V' (at P+εD) are "equal" if

$$\bar{V}^{\mu} = \bar{V}'^{\mu}, \qquad (3)$$

that is, if they have the same components in R.N.C. Thus, the connection coefficients at P vanish, $\bar{A}^{\mu}_{\alpha\nu}(P) = 0$, in R.N.C.

This does not mean that the connection coefficients vanish in all coordinate systems, however. Let x^{μ} be an arbitrary set of coordinates in a neighborhood of P, and let V^{μ} and V'^{μ} be the components of V and V' w.r.t. this coordinate system. Then

$$V^{\mu} = \frac{\partial x^{\mu}}{\partial \bar{x}^{\nu}}(P)\bar{V}^{\nu} \qquad (4)$$

and

$$V'^{\mu} = \frac{\partial x^{\mu}}{\partial \bar{x}^{\nu}}(P+\varepsilon D)\,\bar{V}'^{\nu} \qquad (5)$$

But by (3) $\bar{V}'^{\nu} = \bar{V}^{\nu}$, and (5) can be written

$$V'^{\mu} = \left(\frac{\partial x^{\mu}}{\partial \bar{x}^{\nu}}(P) + \varepsilon \frac{\partial^2 x^{\mu}}{\partial \bar{x}^{\nu}\partial \bar{x}^{\sigma}}\bar{D}^{\sigma}\right)\bar{V}^{\nu}$$

$$= V^{\mu} + \varepsilon \frac{\partial^2 x^{\mu}}{\partial \bar{x}^{\nu}\partial \bar{x}^{\sigma}}\bar{D}^{\sigma}\bar{V}^{\nu}$$

$$= V^{\mu} + \varepsilon \frac{\partial^2 x^{\mu}}{\partial \bar{x}^{\nu}\partial \bar{x}^{\sigma}}\frac{\partial \bar{x}^{\sigma}}{\partial x^{\alpha}}\frac{\partial \bar{x}^{\nu}}{\partial x^{\beta}}D^{\alpha}V^{\beta}, \qquad (6)$$

where in the last step we convert $\bar{D}^\alpha$ and $\bar{V}^\beta$ components from the $\bar{x}^\mu$ coordinates to the x^μ coordinates. Comparing (5) to (1), we see that the connection coefficients in an arbitrary coordinate system can be written

$$A^\mu_{\alpha\beta} = - \frac{\partial^2 x^\mu}{\partial \bar{x}^\nu \partial \bar{x}^\sigma} \frac{\partial \bar{x}^\sigma}{\partial x^\alpha} \frac{\partial \bar{x}^\nu}{\partial x^\beta}. \tag{7}$$

From this it is evident that $A^\mu_{\alpha\beta} = + A^\mu_{\beta\alpha}$.

This is not a convenient formula for computing $A^\mu_{\alpha\beta}$, however. Let us consider the Christoffel symbols. See Eq. (8) in the notes on transforming the Christoffel symbols. Identifying coordinates x'^μ in that formula with R.N.C. $\bar{x}^\mu$ here, we have

$$\Gamma^\mu_{\alpha\beta} = \left(\frac{\partial x^\mu}{\partial \bar{x}^\sigma} \bar{\Gamma}^\sigma_{\lambda\gamma} - \frac{\partial^2 x^\mu}{\partial \bar{x}^\lambda \partial \bar{x}^\gamma} \right) \frac{\partial \bar{x}^\lambda}{\partial x^\alpha} \frac{\partial \bar{x}^\gamma}{\partial x^\beta}. \tag{8}$$

But evaluating this at P, where $\bar{\Gamma}^\sigma_{\lambda\mu} = 0$, we have

$$\Gamma^\mu_{\alpha\beta} = - \frac{\partial^2 x^\mu}{\partial \bar{x}^\lambda \partial \bar{x}^\gamma} \frac{\partial \bar{x}^\lambda}{\partial x^\alpha} \frac{\partial \bar{x}^\gamma}{\partial x^\beta}. \tag{9}$$

Comparing this to (7), we find

$$A^\mu_{\alpha\beta} = \Gamma^\mu_{\alpha\beta}. \tag{10}$$

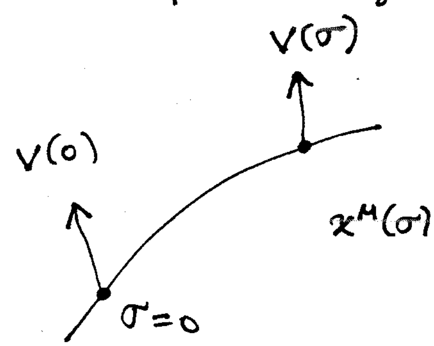
The connection coefficients in general relativity are the same as the Christoffel symbols.

Thus, in relativity theory, we say that vectors V (at P) and V' (at $P+\varepsilon D$) are the "same" if

$$\boxed{V'^{\mu} = V^{\mu} - \varepsilon\, \Gamma^{\mu}_{\alpha\beta} D^{\alpha} V^{\beta}} \qquad (11)$$

The language we use for this is the following: We say that V has been parallel transported from P to $P+\varepsilon D$ along small displacement εD, giving vector V' at $P+\varepsilon D$.

Now, given that we can parallel transport a vector a small step, we can accumulate small steps along a curve to parallel transport a vector between distant tangent spaces. Let $x^{\mu} = x^{\mu}(\sigma)$ be a curve, where σ is a parameter. Let $V^{\mu}(\sigma)$ be the components of a vector that is parallel transported along this curve. Then in the small step $\sigma \to \sigma + d\sigma$, we can write (11) in the form



$$V^{\mu}(\sigma+d\sigma) = V^{\mu}(\sigma) + dV^{\mu}$$

$$= V^{\mu}(\sigma) - d\sigma\, \Gamma^{\mu}_{\alpha\beta}(x(\sigma)) \frac{dx^{\alpha}}{d\sigma} V^{\beta}, \qquad (12)$$

where we have replaced εD^{α} by $d\sigma \frac{dx^{\alpha}}{d\sigma}$.

parallel transport along a curve $x^\mu(\sigma)$.

This implies

$$\boxed{\frac{dV^\mu}{d\sigma} = - \Gamma^\mu_{\alpha\beta}(x(\sigma))\frac{dx^\alpha}{d\sigma} V^\beta .} \tag{13}$$

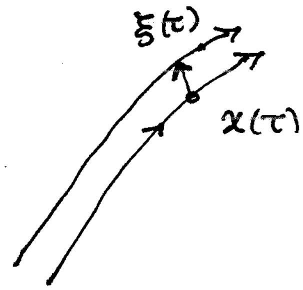
This is a differential equation that the components of the parallel transported vector satisfy. It is understood that the curve $x^\mu(\sigma)$ is given, in this equation, which is therefore a linear, 1st order differential equation for $V^\mu(\sigma)$. A complete set of initial conditions is $V^\mu(0)$ at some point $\sigma = 0$ ~~along~~ on the curve.

The parallel transport equation (13) allows us to identify vectors in distant tangent spaces, but the identification depends on the path. That is, parallel transporting a vector around a closed loop generally does not return the vector to its original value. This is a manifestation of <u>curvature</u>.

To change the game somewhat, suppose now that the curve $x^\mu(\sigma)$ is given, and V^μ is given as a function σ (V is defined along the curve). $V^\mu(\sigma)$ does not need to satisfy the parallel transport equation (13). Here are two examples in which a vector is given as a function of a parameter along a curve, but it is not defined elsewhere in the space. Let the curve $x^\mu(\tau)$ be the world line of a particle (not necessarily a geodesic), with parameter $\sigma = \tau =$ proper time. Then the unit tangent vector u with components

$$u^\mu = \frac{dx^\mu}{d\tau} = u^\mu(\tau) \tag{14}$$

is defined along the curve. Or, let $\xi^\mu(\tau)$ be a small displacement vector between two nearby geodesics; in this case the world line $x^\mu(\tau)$ is a geodesic.



Now let us parallel transport $V(\sigma)$ at $x(\sigma)$ to V' at $x(\sigma + d\sigma)$. According to (11), we have

$$V'^\mu = V^\mu - \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\sigma} V^\beta \, d\sigma \,, \tag{15}$$

where we have replaced εD^α by $\frac{dx^\alpha}{d\sigma} d\sigma$. Now vectors V' and $V(\sigma + d\sigma)$ are defined in the same tangent space, so it makes sense to subtract them. Note that $V(\sigma + d\sigma)$ is not necessarily equal to V', because we are not assuming that V has been parallel transported along the curve. This difference is proportional to $d\sigma$, so we define

$$\frac{DV^\mu}{D\sigma} = \lim_{\Delta\sigma \to 0} \frac{V^\mu(\sigma + \Delta\sigma) - V'^\mu}{\Delta\sigma}$$

$$\boxed{\frac{DV^\mu}{D\sigma} = \frac{dV^\mu}{d\sigma} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\sigma} V^\beta .} \tag{16}$$

We call $\frac{DV}{D\sigma}$ the <u>covariant</u> <u>derivative</u> of V along $x(\sigma)$; its μ-th component $\frac{DV^\mu}{D\sigma}$ is given by (16).

You can think of the covariant derivative $\frac{DV^\mu}{D\sigma}$ as the ordinary derivative w.r.t. σ (the first term of (16)) plus a Γ-correction. Neither of these two terms transforms as a vector, that is, neither obeys the vector transformation law,

$$X'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} X^\nu \qquad (17)$$

where X^μ and X'^μ are the components of a vector X w.r.t. two coordinate systems x^μ and x'^μ. But the sum of the two terms does obey the vector transformation law,

$$\left(\frac{DV^\mu}{D\sigma}\right)' \equiv \frac{DV'^\mu}{D\sigma} = \frac{\partial x'^\mu}{\partial x^\nu}\left(\frac{DV^\nu}{D\sigma}\right). \qquad (18)$$

The covariant derivative maps vectors into vectors.

Notice that the equation of parallel transport (13) can be written,

$$\frac{DV^\mu}{D\sigma} = 0. \qquad (19).$$

Notice also that if a geodesic $x^\mu(\tau)$ is given, then the unit tangent vector

$$u^\mu = \frac{dx^\mu}{d\tau} \qquad (20)$$

is its own parallel transport along the geodesic, that is,

$$\frac{Du^\mu}{D\tau} = \frac{du^\mu}{d\tau} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\tau}u^\beta = \frac{du^\mu}{d\tau} + \Gamma^\mu_{\alpha\beta}u^\alpha u^\beta = 0. \qquad (21)$$

The meaning of this equation is that a geodesic is not only a curve of extremal distance or time, but it is also the straightest curve. To make a straight curve, we start at a point P and choose an initial unit time-like vector $u(0)$ (for a time-like geodesic). We then choose some small increment $\Delta\tau$ in proper time, and make a step $\Delta x^\mu = u^\mu(0)\,\Delta\tau$ away from point P. Now the problem is, how do we know what direction to make the next step? We parallel transport $u^\mu(0)$ from P (with coordinates x_P^μ) to the point $x_P^\mu + \Delta x^\mu = x_P^\mu + u^\mu(0)\,\Delta\tau$. This gives the vector $u'^\mu(0)$:

$$\begin{aligned} u'^\mu(0) &= u^\mu(0) - \Gamma^\mu_{\alpha\beta}\,\Delta x^\alpha\, u^\beta(0) \\ &= u^\mu(0) - \Gamma^\mu_{\alpha\beta}\, u^\alpha(0)\, u^\beta(0)\,\Delta\tau. \end{aligned} \qquad (22)$$

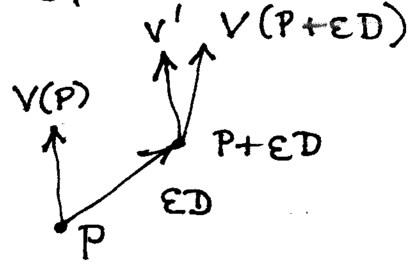
Call $u'^\mu(0) \equiv u^\mu(1)$. Then use $u(1)$ to take the next small step. This procedure amounts to integrating the equations,

$$\left.\begin{aligned} \frac{dx^\mu}{d\tau} &= u^\mu \\ \frac{du^\mu}{d\tau} &= -\Gamma^\mu_{\alpha\beta}(x)\, u^\alpha u^\beta \end{aligned}\right\} \qquad (23)$$

which are the geodesic equations.

Now let us change the game again, and let V be a <u>vector field</u>, defined everywhere in a neighborhood of a point P (instead of just along a curve, as previously).

Let D be a vector based at P. We wish to compute the <u>directional derivative</u> of V along D.



This means we want to compare the values of the vector field V at the base and tip of the vector εD, where ε is small, that is, the vectors V(P) and V(P+εD). But these vectors are in different tangent spaces, so we cannot compare them directly. Instead, we must compare V(P+εD) with V', the parallel transport of V(P) along εD. Then we define the directional derivative of V along D by

$$(\nabla_D V)^\mu \equiv \nabla_D V^\mu = \lim_{\varepsilon \to 0} \frac{V^\mu(P+\varepsilon D) - V'}{\varepsilon}$$

$$= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[\left(V^\mu(P) + \varepsilon \frac{\partial V^\mu}{\partial x^\alpha} D^\alpha\right) - \left(V^\mu(P) - \Gamma^\mu_{\alpha\beta}\, \varepsilon D^\alpha V^\beta\right)\right] \tag{24}$$

or,

$$\boxed{\nabla_D V^\mu = \left(\frac{\partial V^\mu}{\partial x^\alpha} + \Gamma^\mu_{\alpha\beta} V^\beta\right) D^\alpha .} \tag{25}$$