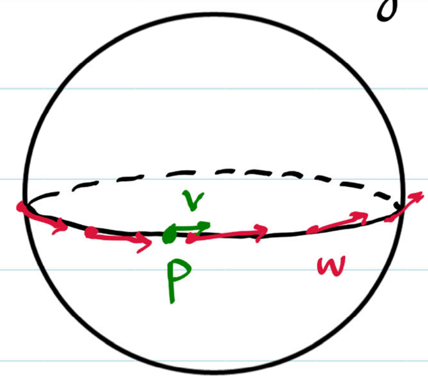
Covariant derivatives (of vector fields)

Earlier we defined what it means for a vector field on a regular surface to be differentiable. We now describe how we differentiate such a vector field.

Let S be a regular surface on $\mathbb{R}^3$. Recall if $f : S \to \mathbb{R}$ is a differentiable function on S, $p \in S$, and $v \in T_p(S)$, then one can differentiate f at p in the direction v, by fitting a C^∞ curve $\gamma(t)$ on S with $\gamma(0) = p$, $\gamma'(0) = v$ and computing $\frac{d}{dt}\Big|_{t=0} f(\gamma(t))$. What if now in place of a differentiable function $f : S \to \mathbb{R}$ we have a differentiable vector field w on S? If $p \in S$, $v \in T_p(S)$, and $\gamma(t)$ is a C^∞ curve on S with $\gamma(0) = p$, $\gamma'(0) = v$, should we differentiate w in the direction v by computing $\frac{d}{dt}\Big|_{t=0} w(\gamma(t))$? No!

Recall that by a vector field w on S we mean a tangent vector field, i.e. w should be tangent to S at every point of S. e.g. S may be the surface of Earth, $w(p)$ may be the velocity of wind at $p \in S$ (which is tangent to S). Suppose wind blows at a constant speed of 10 km/h along the equator, from west to east. Then w is as shown in red in the picture. If p is on the equator and $v \in T_p(S)$ is tangent to the equator, we might want to say that the "derivative" of w at p in the direction v is the zero vector. But if we compute $\frac{d}{dt} w(\gamma(t))$ where $\gamma(t)$ parametrizes the equator, $\frac{d}{dt}\Big|_{t=0} w(\gamma(t))$ is certainly not $(0,0,0)$, because $w(\gamma(t))$ changes direction in $\mathbb{R}^3$ as t varies! Indeed, $\frac{d}{dt} w(\gamma(t))$ is normal to $T_{\gamma(t)}(S)$ $\forall t$, and not tangent to S.

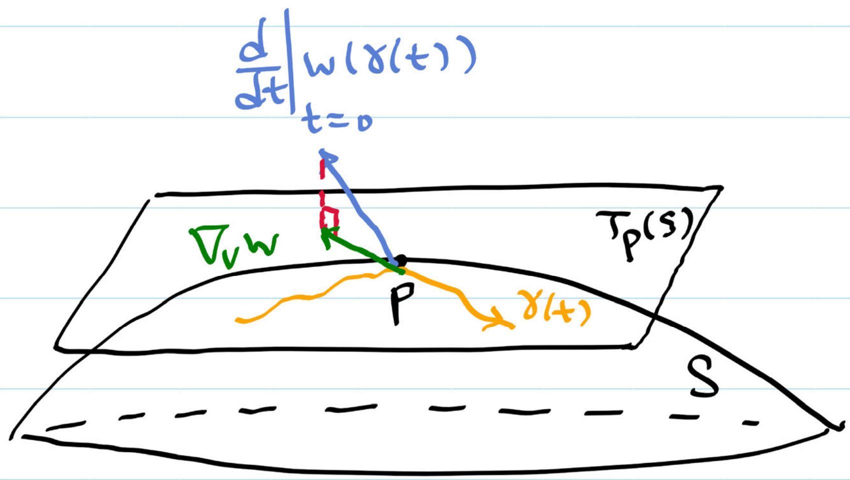


We want a "derivative" of w that gives a (tangent) vector field on S.

Definition. Let w be a differentiable vector field on a regular surface S, $p \in S$, $v \in T_p(S)$. Let $\gamma(t)$ be any C^∞ curve on S with $\gamma(0) = p$ and $\gamma'(0) = v$. The Covariant derivative of w in the direction v is

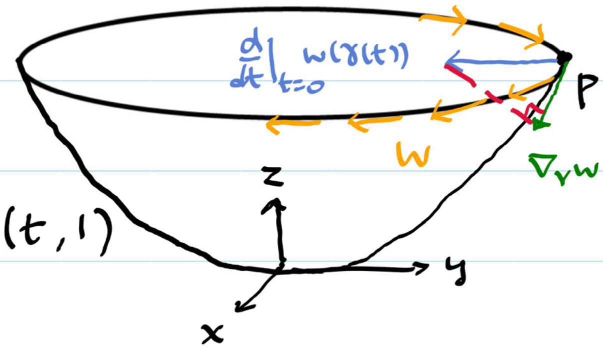
$\nabla_v w :=$ the orthogonal projection of $\left.\frac{d}{dt}\right|_{t=0} w(\gamma(t))$ to $T_p(S)$.

Remark: $\nabla_v w$ is well-defined independent of the choice of $\gamma(t)$ as long as $\gamma(0) = p$ and $\gamma'(0) = v$, because $\left.\frac{d}{dt}\right|_{t=0} w(\gamma(t))$ depends only on v but not on which γ we picked (just as the case of functions).



By definition $\nabla_v w \in T_p(S)$

Example. Let $S = \{(x,y,z) \in \mathbb{R}^3 : z = x^2+y^2\}$ paraboloid in $\mathbb{R}^3$, parametrized by $\underline{x}(u,v) = (u, v, u^2+v^2)$ $\forall (u,v) \in \mathbb{R}^2$. Let w be the vector field on S defined by $w(\underline{x}(u,v)) = v\,\underline{x}_u(u,v) - u\,\underline{x}_v(u,v)$ $\forall (u,v) \in \mathbb{R}^2$. Let $p = \underline{x}(0,1) = (0,1,1) \in S$, and $v = \underline{x}_u(0,1) = (1,0,0) \in T_p(S)$. Find $\nabla_v w$.



Solution. Note $\underline{x}_u = (1, 0, 2u)$ and $\underline{x}_v = (0, 1, 2v)$ so
$w(\underline{x}(u,v)) = (v, -u, 0)$ $\forall (u,v) \in \mathbb{R}^2$. Let $\gamma(t) = \underline{x}(t, 1)$ so that $\gamma(t) \in S$ $\forall t \in \mathbb{R}$, $\gamma(0) = p$, $\gamma'(0) = v$.

Then $w(\gamma(t)) = (1, -t, 0)$ $\forall t \in \mathbb{R}$, so $\left.\frac{d}{dt}\right|_{t=0} w(\gamma(t)) = (0,-1,0)$.

It is not tangent to S.

We orthogonally project it down to $T_p(S)$: let N be a unit normal to $T_p(S)$ and $(0,-1,0) = A\underline{x}_u + B\underline{x}_v + CN$ at p.

Since $\underline{x}_u(0,1) = (1,0,0)$ and $\underline{x}_v(0,1) = (0,1,2)$, taking dot products with them yields $0 = A$ and $-1 = 5B$, and $\nabla_v w$ = the orthogonal projection of $(0,-1,0)$ to $T_p(S) = A\underline{x}_u + B\underline{x}_v = -\frac{1}{5}(0,1,2)$.

↑ drop CN!

At p,
$(0,-1,0)\cdot\underline{x}_u = A\underline{x}_u\cdot\underline{x}_u + B\cancel{\underline{x}_u\cdot\underline{x}_v}^{\,0} + C\cancel{\underline{x}_u\cdot N}^{\,0}$;

$(0,-1,0)\cdot\underline{x}_v = A\cancel{\underline{x}_u\cdot\underline{x}_v}^{\,0} + B\underline{x}_v\cdot\underline{x}_v + C\cancel{\underline{x}_v\cdot N}^{\,0}$

Often $\nabla_v w$ is calculated in the presence of a local parametrization. Let S be a regular surface in $\mathbb{R}^3$, $\underline{x}: \mathcal{U} \to \underline{x}(\mathcal{U}) \subseteq S$ be a local parametrization defined on some open set $\mathcal{U}$ in $\mathbb{R}^2$. The key is to figure out $\nabla_{\underline{x}_u} \underline{x}_u$, $\nabla_{\underline{x}_v} \underline{x}_u$, $\nabla_{\underline{x}_u} \underline{x}_v$ and $\nabla_{\underline{x}_v} \underline{x}_v$ on $\underline{x}(\mathcal{U})$. They are tangent to $\underline{x}(\mathcal{U})$, hence linear combinations of $\underline{x}_u$ and $\underline{x}_v$.

Write
$$\nabla_{\underline{x}_u} \underline{x}_u = \Gamma_{11}^1 \underline{x}_u + \Gamma_{11}^2 \underline{x}_v$$
$$\nabla_{\underline{x}_v} \underline{x}_u = \Gamma_{21}^1 \underline{x}_u + \Gamma_{21}^2 \underline{x}_v$$
$$\nabla_{\underline{x}_u} \underline{x}_v = \Gamma_{12}^1 \underline{x}_u + \Gamma_{12}^2 \underline{x}_v$$
$$\nabla_{\underline{x}_v} \underline{x}_v = \Gamma_{22}^1 \underline{x}_u + \Gamma_{22}^2 \underline{x}_v$$

For $i, j, k \in \{1,2\}$, Γ_{ij}^k are C^∞ functions of (u,v) called the Christoffel symbols on $\underline{x}(\mathcal{U})$.

The Christoffel symbols can be computed by taking dot products of these equations with $\underline{x}_u$ and $\underline{x}_v$. Indeed, if $\nabla_v w$ = orthogonal projection of $\frac{d}{dt}\big|_{t=0} w(\gamma(t))$ to $T_p(S)$, then $\nabla_v w - \frac{d}{dt}\big|_{t=0} w(\gamma(t))$ is normal to $T_p(S)$, so its dot product with any $X \in T_p(S)$ is 0. Hence for X tangent to S, we have $X \cdot \nabla_{\underline{x}_u} \underline{x}_u = X \cdot \underline{x}_{uu}$, $X \cdot \nabla_{\underline{x}_u} \underline{x}_v = X \cdot \underline{x}_{uv} = X \cdot \nabla_{\underline{x}_v} \underline{x}_u$, and $X \cdot \nabla_{\underline{x}_v} \underline{x}_v = X \cdot \underline{x}_{vv}$, which we apply to $X = \underline{x}_u$ and $X = \underline{x}_v$.

Recall the first fundamental forms $E = \mathbf{x}_u \cdot \mathbf{x}_u$, $F = \mathbf{x}_u \cdot \mathbf{x}_v$, $G = \mathbf{x}_v \cdot \mathbf{x}_v$.
Hence Γ_{ij}^k, for $i, j, k \in \{1, 2\}$, can be found by solving

$$\begin{cases} \mathbf{x}_{uu} \cdot \mathbf{x}_u = \Gamma_{11}^1 E + \Gamma_{11}^2 F \\ \mathbf{x}_{uu} \cdot \mathbf{x}_v = \Gamma_{11}^1 F + \Gamma_{11}^2 G \end{cases} \qquad \begin{cases} \mathbf{x}_{vu} \cdot \mathbf{x}_u = \Gamma_{21}^1 E + \Gamma_{21}^2 F \\ \mathbf{x}_{vu} \cdot \mathbf{x}_v = \Gamma_{21}^1 F + \Gamma_{21}^2 G \end{cases}$$

$$\begin{cases} \mathbf{x}_{uv} \cdot \mathbf{x}_u = \Gamma_{12}^1 E + \Gamma_{12}^2 F \\ \mathbf{x}_{uv} \cdot \mathbf{x}_v = \Gamma_{12}^1 F + \Gamma_{12}^2 G \end{cases} \qquad \begin{cases} \mathbf{x}_{vv} \cdot \mathbf{x}_u = \Gamma_{22}^1 E + \Gamma_{22}^2 F \\ \mathbf{x}_{vv} \cdot \mathbf{x}_v = \Gamma_{22}^1 F + \Gamma_{22}^2 G \end{cases}$$

In short, to compute say Γ_{12}^1 and Γ_{12}^2, we first compute $\mathbf{x}_{uv}$, then compute $\mathbf{x}_{uv} \cdot \mathbf{x}_u$ and $\mathbf{x}_{uv} \cdot \mathbf{x}_v$, then compute E, F, G, and solve a system of 2 equations involving Γ_{12}^1 and Γ_{12}^2.
Note $\Gamma_{12}^1 = \Gamma_{21}^1$ and $\Gamma_{12}^2 = \Gamma_{21}^2$, since they obey the same system of equations. Later on we will introduce a trick that allows us to compute dot products such as $\mathbf{x}_{uv} \cdot \mathbf{x}_u$ and $\mathbf{x}_{uv} \cdot \mathbf{x}_v$ more easily. But for now let's look at one example.

Example. Let $S = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2=1\}$, parametrized by

$$\mathbf{x}(u,v) = (\cos u \cos v,\ \cos u \sin v,\ \sin u), \quad (u,v) \in (-\tfrac{\pi}{2}, \tfrac{\pi}{2}) \times (-\pi, \pi)$$

Compute Γ_{12}^1 and Γ_{12}^2.

Solution. (Method 1) Note $\mathbf{x}_u = (-\sin u \cos v,\ -\sin u \sin v,\ \cos u)$

$$\mathbf{x}_v = (-\cos u \sin v,\ \cos u \cos v,\ 0)$$

So $E = 1$, $F = 0$, $G = \cos^2 u$.

Also $\mathbf{x}_{uv} = (\sin u \sin v, -\sin u \cos v, 0)$

So $\mathbf{x}_{uv} \cdot \mathbf{x}_u = 0$, $\mathbf{x}_{uv} \cdot \mathbf{x}_v = -\sin u \cos u$, which gives

$$\begin{cases} \Gamma_{12}^1 \cdot 1 + \Gamma_{12}^2 \cdot 0 = 0 \\ \Gamma_{12}^1 \cdot 0 + \Gamma_{12}^2 \cos^2 u = -\sin u \cos u \end{cases} \Rightarrow \Gamma_{12}^1 = 0,\ \Gamma_{12}^2 = -\tan u.$$

Exercise. Check in this example that $\Gamma_{11}^1 = \Gamma_{11}^2 = \Gamma_{22}^2 = 0$, $\Gamma_{22}^1 = \sin u \cos u$,

(Hint: First show $\mathbf{x}_{uu} \cdot \mathbf{x}_u = \mathbf{x}_{uu} \cdot \mathbf{x}_v = \mathbf{x}_{vv} \cdot \mathbf{x}_v = 0$ and $\mathbf{x}_{vv} \cdot \mathbf{x}_u = \sin u \cos u$).

What is Γ_{21}^1 and Γ_{21}^2 in this example? Answer: $\Gamma_{21}^1 = 0$, $\Gamma_{21}^2 = -\tan u$.

To compute Γ_{ij}^{k} we need to compute $x_{uu}\cdot x_u$, $x_{uu}\cdot x_v$, $x_{uv}\cdot x_u$, $x_{uv}\cdot x_v$, $x_{vv}\cdot x_u$ and $x_{vv}\cdot x_v$. The following trick is often useful.

Fact. $\begin{cases} x_{uu}\cdot x_u = \frac{1}{2}E_u \\ x_{uu}\cdot x_v = F_u - \frac{1}{2}E_v \end{cases}$, $\begin{cases} x_{uv}\cdot x_u = \frac{1}{2}E_v \\ x_{uv}\cdot x_v = \frac{1}{2}G_u \end{cases}$, $\begin{cases} x_{vv}\cdot x_u = F_v - \frac{1}{2}G_u \\ x_{vv}\cdot x_v = \frac{1}{2}G_v \end{cases}$

Here $E_u = \frac{\partial E}{\partial u}$, $E_v = \frac{\partial E}{\partial v}$, $F_u = \frac{\partial F}{\partial u}$, $F_v = \frac{\partial F}{\partial v}$, $G_u = \frac{\partial G}{\partial u}$, $G_v = \frac{\partial G}{\partial v}$.

The proofs are simple consequences of the product rule.

e.g. $\frac{\partial}{\partial u}(x_u\cdot x_u) = 2\, x_{uu}\cdot x_u$ so $x_{uu}\cdot x_u = \frac{1}{2}\frac{\partial}{\partial u}(x_u\cdot x_u) = \frac{1}{2}E_u$.

$\frac{\partial}{\partial u}(x_u\cdot x_v) = x_{uu}\cdot x_v + x_{uv}\cdot x_u = x_{uv}\cdot x_v + \frac{1}{2}\frac{\partial}{\partial v}(x_u\cdot x_u)$

so $x_{uu}\cdot x_v = F_u - \frac{1}{2}E_v$.

All other four identities can be proved similarly (try it!).

So now the Christoffel symbols can be found by solving

$\begin{cases} \Gamma_{11}^{1}E + \Gamma_{11}^{2}F = \frac{1}{2}E_u \\ \Gamma_{11}^{1}F + \Gamma_{11}^{2}G = F_u - \frac{1}{2}E_v \end{cases}$ $\begin{cases} \Gamma_{21}^{1}E + \Gamma_{21}^{1}F = \frac{1}{2}E_v \\ \Gamma_{21}^{1}F + \Gamma_{21}^{2}G = \frac{1}{2}G_u \end{cases}$

$\begin{cases} \Gamma_{12}^{1}E + \Gamma_{12}^{2}F = \frac{1}{2}E_v \\ \Gamma_{12}^{1}F + \Gamma_{12}^{2}G = \frac{1}{2}G_u \end{cases}$ $\begin{cases} \Gamma_{22}^{1}E + \Gamma_{22}^{2}F = F_v - \frac{1}{2}G_u \\ \Gamma_{22}^{1}F + \Gamma_{22}^{2}G = \frac{1}{2}G_v. \end{cases}$

Example. Again let $S = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2=1\}$, parametrized by $\underset{\sim}{x}(u,v) = (\cos u \cos v,\ \cos u \sin v,\ \sin u)$, $(u,v) \in (-\frac{\pi}{2}, \frac{\pi}{2}) \times (-\pi, \pi)$. Find Γ_{ij}^k for $i,j,k \in \{1,2\}$.

Solution (Method 2). Again $\underset{\sim}{x}_u = (-\sin u \cos v,\ -\sin u \sin v,\ \cos u)$

$\underset{\sim}{x}_v = (-\cos u \sin v,\ \cos u \cos v,\ 0)$

So $E = 1$, $F = 0$, $G = \cos^2 u$.

We have $E_u = E_v = 0$, $F_u = F_v = 0$, $G_u = -2\sin u \cos u$, $G_v = 0$.

So $\begin{cases} \Gamma_{11}^1 E + \Gamma_{11}^2 F = \frac{1}{2} E_u \\ \Gamma_{11}^1 F + \Gamma_{11}^2 G = F_u - \frac{1}{2} E_v \end{cases} \Rightarrow \begin{cases} \Gamma_{11}^1 \cdot 1 + \Gamma_{11}^2 \cdot 0 = 0 \\ \Gamma_{11}^1 \cdot 0 + \Gamma_{11}^2 \cdot \cos^2 u = 0 \end{cases} \Rightarrow \begin{cases} \Gamma_{11}^1 = 0 \\ \Gamma_{11}^2 = 0 \end{cases}$,

$\begin{cases} \Gamma_{12}^1 E + \Gamma_{12}^1 F = \frac{1}{2} E_v \\ \Gamma_{12}^1 F + \Gamma_{12}^2 G = \frac{1}{2} G_u \end{cases} \Rightarrow \begin{cases} \Gamma_{12}^1 \cdot 1 + \Gamma_{12}^2 \cdot 0 = 0 \\ \Gamma_{12}^1 \cdot 0 + \Gamma_{12}^2 \cdot \cos^2 u = -\sin u \cos u \end{cases} \Rightarrow \begin{cases} \Gamma_{12}^1 = 0 \\ \Gamma_{12}^2 = -\tan u \end{cases}$

which also implies $\Gamma_{21}^1 = 0$, $\Gamma_{21}^2 = -\tan u$; finally,

$\begin{cases} \Gamma_{22}^1 E + \Gamma_{22}^2 F = F_v - \frac{1}{2} G_u \\ \Gamma_{22}^1 F + \Gamma_{22}^2 G = \frac{1}{2} G_v \end{cases} \Rightarrow \begin{cases} \Gamma_{22}^1 \cdot 1 + \Gamma_{22}^2 \cdot 0 = \sin u \cos u \\ \Gamma_{22}^1 \cdot 0 + \Gamma_{22}^2 \cdot \cos^2 u = 0 \end{cases} \Rightarrow \begin{cases} \Gamma_{22}^1 = \sin u \cos u \\ \Gamma_{22}^2 = 0 \end{cases}$.

Now that we know how to compute the Christoffel symbols Γ_{ij}^k, we know how to compute $\nabla_{x_u} x_u$, $\nabla_{x_v} x_u$, $\nabla_{x_u} x_v$ and $\nabla_{x_v} x_v$. This will be useful in computing $\nabla_v w$ in a local parametrization, once we observe the following key properties for covariant derivatives:

In more advanced course, such a ∇ is called a connection on S.

① If $v_1, v_2 \in T_p(S)$, $c_1, c_2 \in \mathbb{R}$, and w is a differentiable vector field on S, then $\nabla_{c_1 v_1 + c_2 v_2} w = c_1 \nabla_{v_1} w + c_2 \nabla_{v_2} w$.

② If $v \in T_p(S)$ and w_1, w_2 are differentiable vector fields on S, then $\nabla_v (w_1 + w_2) = \nabla_v w_1 + \nabla_v w_2$.

③ If $v \in T_p(S)$, w is a differentiable vector field on S and f is a differentiable function on S, then $\nabla_v (fw) = df_p(v)\, w + f(p) \nabla_v w$.

e.g. for ①, note that if $\underline{x}$ is a local parametrization of S with $p = \underline{x}(0,0)$ and $\gamma_i(t) = \underline{x}(\alpha_i(t))$, $i=1,2$, are C^∞ curves on S such that $\gamma_i(0) = p$ and $\gamma_i'(0) = v_i$, then $\gamma(t) := \underline{x}(c_1\alpha_1(t) + c_2\alpha_2(t))$ is a C^∞ curve on S with $\gamma(0) = p$ and $\gamma'(0) = c_1 v_1 + c_2 v_2$, and chain rule implies $\frac{d}{dt}\Big|_{t=0} w(\gamma(t))$ $= c_1 \frac{d}{dt}\Big|_{t=0} w(\gamma_1(t)) + c_2 \frac{d}{dt}\Big|_{t=0} w(\gamma_2(t))$. So projection onto $T_p(S)$ gives property ①.

The proof of ② uses that $\frac{d}{dt}(w_1+w_2)(\gamma(t)) = \frac{d}{dt} w_1(\gamma(t)) + \frac{d}{dt} w_2(\gamma(t))$, and that of ③ uses that $\frac{d}{dt}\left[f(\gamma(t))\, w(\gamma(t))\right] = \left(\frac{d}{dt} f(\gamma(t))\right) w(\gamma(t)) + f(\gamma(t)) \frac{d}{dt} w(\gamma(t))$. Here $\gamma(t)$ is a C^∞ curve on S with $\gamma(0)=p$, $\gamma'(0)=v$. Details omitted.

To recap, let S be a regular surface, $\underline{x}: U \to \underline{x}(U) \subseteq S$ be a local parametrization, $p \in \underline{x}(U)$, $v \in T_p(S)$, and w be a differentiable vector field on $\underline{x}(U)$. We knew how to compute $\nabla_{\underline{x}_u} \underline{x}_u$, $\nabla_{\underline{x}_u} \underline{x}_v$, $\nabla_{\underline{x}_v} \underline{x}_u$ and $\nabla_{\underline{x}_v} \underline{x}_v$ because we knew how to compute Γ_{ij}^k, $i,j,k \in \{1,2\}$. To compute $\nabla_v w$, we first write $v = c\,\underline{x}_u + d\,\underline{x}_v$ at p and use $\nabla_v w = c \nabla_{\underline{x}_u} w + d \nabla_{\underline{x}_v} w$. Then we write $w(\underline{x}(u,v)) = a(u,v)\, \underline{x}_u + b(u,v)\, \underline{x}_v$ for some differentiable functions $a(u,v)$ and $b(u,v)$, defined for $(u,v) \in U$; then $\nabla_{\underline{x}_u} w = \nabla_{\underline{x}_u}(a\, \underline{x}_u) + \nabla_{\underline{x}_u}(b\, \underline{x}_v) = \frac{\partial a}{\partial u} \underline{x}_u + a \nabla_{\underline{x}_u} \underline{x}_u + \frac{\partial b}{\partial u} \underline{x}_v + b \nabla_{\underline{x}_u} \underline{x}_v$, and similarly $\nabla_{\underline{x}_v} w = \frac{\partial a}{\partial v} \underline{x}_u + a \nabla_{\underline{x}_v} \underline{x}_u + \frac{\partial b}{\partial v} \underline{x}_v + b \nabla_{\underline{x}_v} \underline{x}_v$.

Altogether, we know how to compute $\nabla_v w$ for a general differentiable vector field w in a general direction v.

Example. Let $S^2 = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2 = 1\}$, $S = (0,0,-1)$, and $S^2 \setminus \{S\}$ be parametrized by $\underline{x}(u,v) = \left(\frac{2u}{1+u^2+v^2}, \frac{2v}{1+u^2+v^2}, \frac{1-u^2-v^2}{1+u^2+v^2}\right)$, $(u,v) \in \mathbb{R}^2$.

Then $w(p) = \begin{cases} -v\,\underline{x}_u(u,v) + u\,\underline{x}_v(u,v) & \text{if } p = \underline{x}(u,v) \\ (0,0,0) & \text{if } p = S \end{cases}$ is a C^∞ vector field on S^2.

Let $p = \left(\frac{2\sqrt{2}}{5}, \frac{2\sqrt{2}}{5}, -\frac{3}{5}\right)$, $v = (-1,1,0)$. Find $\nabla_v w$.

Solution. Note $\underline{x}_u = \left(\frac{2(1-u^2+v^2)}{(1+u^2+v^2)^2}, -\frac{4uv}{(1+u^2+v^2)^2}, -\frac{4u}{(1+u^2+v^2)^2}\right)$

$\underline{x}_v = \left(-\frac{4uv}{(1+u^2+v^2)^2}, \frac{2(1+u^2-v^2)}{(1+u^2+v^2)^2}, -\frac{4v}{(1+u^2+v^2)^2}\right)$

So $E = G = \frac{4}{(1+u^2+v^2)^2}$, $F = 0$. It follows that

$F_u = F_v = 0$, $E_u = G_u = -\frac{16u}{(1+u^2+v^2)^3}$, $E_v = G_v = -\frac{16v}{(1+u^2+v^2)^3}$

$\underline{x}(u,v) = \left(\frac{2\sqrt{2}}{5}, \frac{2\sqrt{2}}{5}, -\frac{3}{5}\right)$

$\Rightarrow v = u$,

$\frac{2u}{1+2u^2} = \frac{2\sqrt{2}}{5}$,

$\frac{1-2u^2}{1+2u^2} = -\frac{3}{5}$

$\Rightarrow u = v = \sqrt{2}$

Hence $\Gamma_{11}^1 = \Gamma_{12}^2 = -\Gamma_{22}^1 = -\frac{2u}{1+u^2+v^2}$, $-\Gamma_{11}^2 = \Gamma_{12}^1 = \Gamma_{22}^2 = -\frac{2v}{1+u^2+v^2}$.

At $p = \left(\frac{2\sqrt{2}}{5}, \frac{2\sqrt{2}}{5}, -\frac{3}{5}\right) = \underline{x}(\sqrt{2},\sqrt{2})$, we thus have $\underline{x}_u = \left(\frac{2}{25}, -\frac{8}{25}, -\frac{4\sqrt{2}}{25}\right)$, $\underline{x}_v = \left(-\frac{8}{25}, \frac{2}{25}, -\frac{4\sqrt{2}}{25}\right)$,

$\Gamma_{11}^1 = \Gamma_{12}^2 = -\Gamma_{22}^1 = -\frac{2\sqrt{2}}{5} = -\Gamma_{11}^2 = \Gamma_{12}^1 = \Gamma_{22}^2$, $\nabla_{\underline{x}_u}\underline{x}_u = -\frac{2\sqrt{2}}{5}\underline{x}_u + \frac{2\sqrt{2}}{5}\underline{x}_v = -\nabla_{\underline{x}_v}\underline{x}_v$,

$\nabla_{\underline{x}_v}\underline{x}_u = \nabla_{\underline{x}_u}\underline{x}_v = -\frac{2\sqrt{2}}{5}\underline{x}_u - \frac{2\sqrt{2}}{5}\underline{x}_v$, all evaluated at p. Hence at p, we have

$$\begin{cases} \nabla_{\underline{x}_u} w = \nabla_{\underline{x}_u}(-v\underline{x}_u + u\underline{x}_v)\Big|_{(u,v)=(\sqrt{2},\sqrt{2})} = -\cancelto{0}{\frac{\partial v}{\partial u}}\underline{x}_u + \frac{\partial u}{\partial u}\underline{x}_v - \sqrt{2}\nabla_{\underline{x}_u}\underline{x}_u + \sqrt{2}\nabla_{\underline{x}_u}\underline{x}_v = -\frac{3}{5}\underline{x}_v = \frac{(24,-6,12\sqrt{2})}{125} \\ \nabla_{\underline{x}_v} w = \nabla_{\underline{x}_v}(-v\underline{x}_u + u\underline{x}_v)\Big|_{(u,v)=(\sqrt{2},\sqrt{2})} = -\frac{\partial v}{\partial v}\underline{x}_u + \cancelto{0}{\frac{\partial u}{\partial v}}\underline{x}_v - \sqrt{2}\nabla_{\underline{x}_v}\underline{x}_u + \sqrt{2}\nabla_{\underline{x}_v}\underline{x}_v = \frac{3}{5}\underline{x}_u = \frac{(6,-24,-12\sqrt{2})}{125} \end{cases}$$

Common error: forgot to differentiate $\underline{x}_u$ and $\underline{x}_v$.

But $v = (-1,1,0) = -\frac{5}{2}\underline{x}_u + \frac{5}{2}\underline{x}_v$ at p, so $\nabla_v w = -\frac{5}{2}\nabla_{\underline{x}_u} w + \frac{5}{2}\nabla_{\underline{x}_v} w = \left(-\frac{9}{25}, -\frac{9}{25}, -\frac{12\sqrt{2}}{25}\right)$.

Common error: In calculating say $\nabla_{\underline{x}_u}(a\underline{x}_u + b\underline{x}_v)$, one most common mistake is to differentiate only the coefficients, and think that it is equal to $\frac{\partial a}{\partial u}\underline{x}_u + \frac{\partial b}{\partial u}\underline{x}_v$. This would not work. Indeed, if $\underline{x}, \underline{y}$ are two different local parametrizations near some $p \in S$, and w is a differentiable vector field on S with

$$\begin{cases} w(\underline{x}(u,v)) = a(u,v)\,\underline{x}_u(u,v) + b(u,v)\,\underline{x}_v(u,v) \\ w(\underline{y}(\xi,\eta)) = \tilde{a}(\xi,\eta)\,\underline{y}_\xi(\xi,\eta) + \tilde{b}(\xi,\eta)\,\underline{y}_\eta(\xi,\eta) \end{cases}$$

then in general, $\frac{\partial a}{\partial u}\underline{x}_u + \frac{\partial b}{\partial u}\underline{x}_v \neq \frac{\partial \tilde{a}}{\partial u}\underline{y}_\xi + \frac{\partial \tilde{b}}{\partial u}\underline{y}_\eta$ (more precisely, the LHS at $(u,v) \neq$ the right hand side at $(\xi,\eta) = \underline{y}^{-1}\circ\underline{x}(u,v)$; indeed,

$$\begin{cases} \underline{x}_u = \frac{\partial \xi}{\partial u}\underline{y}_\xi + \frac{\partial \eta}{\partial u}\underline{y}_\eta \\ \underline{x}_v = \frac{\partial \xi}{\partial v}\underline{y}_\xi + \frac{\partial \eta}{\partial v}\underline{y}_\eta \end{cases} \Rightarrow \begin{cases} \tilde{a}(\xi,\eta) = a(u,v)\frac{\partial \xi}{\partial u} + b(u,v)\frac{\partial \xi}{\partial v} \\ \tilde{b}(\xi,\eta) = a(u,v)\frac{\partial \eta}{\partial u} + b(u,v)\frac{\partial \eta}{\partial v} \end{cases}, \text{ so } \begin{cases} \frac{\partial \tilde{a}}{\partial u} = \frac{\partial a}{\partial u}\frac{\partial \xi}{\partial u} + \frac{\partial b}{\partial u}\frac{\partial \xi}{\partial v} + a\frac{\partial^2 \xi}{\partial u^2} + b\frac{\partial^2 \xi}{\partial u \partial v} \\ \frac{\partial \tilde{b}}{\partial u} = \frac{\partial a}{\partial u}\frac{\partial \eta}{\partial u} + \frac{\partial b}{\partial u}\frac{\partial \eta}{\partial v} + a\frac{\partial^2 \eta}{\partial u^2} + b\frac{\partial^2 \eta}{\partial u \partial v} \end{cases}$$

which gives

$$\frac{\partial \tilde{a}}{\partial u}\underline{y}_\xi + \frac{\partial \tilde{b}}{\partial u}\underline{y}_\eta = \frac{\partial a}{\partial u}\underline{x}_u + \frac{\partial b}{\partial u}\underline{x}_v + \underbrace{\left(a\frac{\partial^2 \xi}{\partial u^2} + b\frac{\partial^2 \xi}{\partial u \partial v}\right)\underline{y}_\xi + \left(a\frac{\partial^2 \eta}{\partial u^2} + b\frac{\partial^2 \eta}{\partial u \partial v}\right)\underline{y}_\eta}_{\text{not necessarily zero!}}\Big)$$

So we do need to take into account $\nabla_{\underline{x}_u}\underline{x}_u$ and $\nabla_{\underline{x}_u}\underline{x}_v$ when we compute $\nabla_{\underline{x}_u}(a\underline{x}_u + b\underline{x}_v)$, which makes the derivative ∇ "covariant".

Let's understand this covariance in our previous example. We had a parametrization $\underline{x}(u,v) = \left(\frac{2u}{1+u^2+v^2}, \frac{2v}{1+u^2+v^2}, \frac{1-u^2-v^2}{1+u^2+v^2}\right)$ of S^2, $w(\underline{x}(u,v)) = -v\underline{x}_u(u,v) + u\underline{x}_v(u,v)$, and we calculated $\nabla_v w$ at $p = \left(\frac{2\sqrt{2}}{5}, \frac{2\sqrt{2}}{5}, -\frac{3}{5}\right)$, $v = (-1,1,0) \in T_p(S^2)$. We could also have parametrized S^2 by $\underline{y}(\theta,\phi) = (\cos\phi\cos\theta, \cos\phi\sin\theta, \sin\phi)$. Then $\underline{x}(u,v) = \underline{y}(\theta,\phi)$ if and only if $u = \tan(\frac{\pi}{4}+\frac{\phi}{2})\cos\theta$ and $v = \tan(\frac{\pi}{4}+\frac{\phi}{2})\sin\theta$, hence $w(\underline{y}(\theta,\phi)) = \underline{y}_\theta(\theta,\phi)$. We also have $\underline{y}_\theta = (-\cos\phi\sin\theta, \cos\phi\cos\theta, 0)$ so at $p = \left(\frac{2\sqrt{2}}{5}, \frac{2\sqrt{2}}{5}, -\frac{3}{5}\right) = \underline{y}(\sin^{-1}(-\frac{3}{5}), \frac{\pi}{4})$, we have $v = (-1,1,0) = \frac{5}{2\sqrt{2}}\underline{y}_\theta$. So $\nabla_v w$ at p can also be written $\frac{5}{2\sqrt{2}}\nabla_{\underline{y}_\theta}\underline{y}_\theta$ at $(\phi,\theta) = (\sin^{-1}(-\frac{3}{5}), \frac{\pi}{4})$. If we thought $\underline{y}_\theta$ is "constant" with respect to θ and declared $\nabla_{\underline{y}_\theta}\underline{y}_\theta = 0$, then we would have got a different and incorrect answer $(0,0,0)$ for $\nabla_v w$ at p. But we really should arrive at the same answer even if we calculated in the local parametrization $\underline{y}$: indeed then $\underline{y}_\theta$ is as above, $\underline{y}_\phi = (-\sin\phi\cos\theta, -\sin\phi\sin\theta, \cos\phi)$, so $E = \cos^2\phi$, $F = 0$, $G = 1$, and $\Gamma_{11}^1\cos^2\phi = \frac{1}{2}E_\theta = 0 \Rightarrow \Gamma_{11}^1 = 0$, $\Gamma_{11}^2 \cdot 1 = F_\theta - \frac{1}{2}E_\phi = \cos\phi\sin\phi \Rightarrow \Gamma_{11}^2 = \cos\phi\sin\phi$, so at $p = \underline{y}(\sin^{-1}(-\frac{3}{5}), \frac{\pi}{4})$, we have $\nabla_v w = \frac{5}{2\sqrt{2}}\nabla_{\underline{y}_\theta}\underline{y}_\theta = \frac{5}{2\sqrt{2}}\Gamma_{11}^2\underline{y}_\phi = \frac{5}{2\sqrt{2}}\left(\frac{4}{5}\cdot(-\frac{3}{5})\right)\underline{y}_\phi$ $= -\frac{6\sqrt{2}}{5}\left(\frac{3}{5\sqrt{2}}, \frac{3}{5\sqrt{2}}, \frac{4}{5}\right) = \left(-\frac{18}{25}, -\frac{18}{25}, \frac{24\sqrt{2}}{25}\right)$ as before, as long as we did calculate $\nabla_{\underline{y}_\theta}\underline{y}_\theta$.

Summary. If S is a regular surface and $\underline{x}(u,v)$ is a local parametrization the Christoffel symbols are determined by

$$\begin{cases} \underline{x}_{uu}\cdot\underline{x}_u = \Gamma_{11}^1 E + \Gamma_{11}^2 F \\ \underline{x}_{uu}\cdot\underline{x}_v = \Gamma_{11}^1 F + \Gamma_{11}^2 G \end{cases} \qquad \begin{cases} \underline{x}_{vu}\cdot\underline{x}_u = \Gamma_{21}^1 E + \Gamma_{21}^2 F \\ \underline{x}_{vu}\cdot\underline{x}_v = \Gamma_{21}^1 F + \Gamma_{21}^2 G \end{cases}$$

$$\begin{cases} \underline{x}_{uv}\cdot\underline{x}_u = \Gamma_{12}^1 E + \Gamma_{12}^2 F \\ \underline{x}_{uv}\cdot\underline{x}_v = \Gamma_{12}^1 F + \Gamma_{12}^2 G \end{cases} \qquad \begin{cases} \underline{x}_{vv}\cdot\underline{x}_u = \Gamma_{22}^1 E + \Gamma_{22}^2 F \\ \underline{x}_{vv}\cdot\underline{x}_v = \Gamma_{22}^1 F + \Gamma_{22}^2 G \end{cases}$$

and
$$\begin{cases} \underline{x}_{uu}\cdot\underline{x}_u = \frac{1}{2}E_u \\ \underline{x}_{uu}\cdot\underline{x}_v = F_u - \frac{1}{2}E_v \end{cases}, \quad \begin{cases} \underline{x}_{uv}\cdot\underline{x}_u = \frac{1}{2}E_v \\ \underline{x}_{uv}\cdot\underline{x}_v = \frac{1}{2}G_u \end{cases}, \quad \begin{cases} \underline{x}_{vv}\cdot\underline{x}_u = F_v - \frac{1}{2}G_u \\ \underline{x}_{vv}\cdot\underline{x}_v = \frac{1}{2}G_v \end{cases}$$

We then have $\nabla_{\underline{x}_u}\underline{x}_u = \Gamma_{11}^1\underline{x}_u + \Gamma_{11}^2\underline{x}_v$, $\nabla_{\underline{x}_u}\underline{x}_v = \Gamma_{12}^1\underline{x}_u + \Gamma_{12}^2\underline{x}_v$,

$\nabla_{\underline{x}_v}\underline{x}_u = \Gamma_{21}^1\underline{x}_u + \Gamma_{21}^2\underline{x}_v$, $\nabla_{\underline{x}_v}\underline{x}_v = \Gamma_{22}^1\underline{x}_u + \Gamma_{22}^2\underline{x}_v$.

In particular, $\nabla_{\underline{x}_u}\underline{x}_v = \nabla_{\underline{x}_v}\underline{x}_u$ (in general, $\nabla_v w \neq \nabla_w v$).

Finally, if $w(\underline{x}(u,v)) = a(u,v)\,\underline{x}_u(u,v) + b(u,v)\,\underline{x}_v(u,v)$, then

$$\begin{cases} \nabla_{\underline{x}_u} w = \frac{\partial a}{\partial u}\underline{x}_u + a\,\nabla_{\underline{x}_u}\underline{x}_u + \frac{\partial b}{\partial u}\underline{x}_v + b\,\nabla_{\underline{x}_u}\underline{x}_v \\ \nabla_{\underline{x}_v} w = \frac{\partial a}{\partial v}\underline{x}_u + a\,\nabla_{\underline{x}_v}\underline{x}_u + \frac{\partial b}{\partial v}\underline{x}_v + b\,\nabla_{\underline{x}_v}\underline{x}_v \end{cases}$$; furthermore,

if $v = c\,\underline{x}_u + d\,\underline{x}_v$, then $\nabla_v w = c\,\nabla_{\underline{x}_u} w + d\,\nabla_{\underline{x}_v} w$.

Further properties of covariant differentiation

Two further properties of our covariant derivative are worth mentioning. To describe these we introduce a convention and a definition. Recall that if S is a regular surface in $\mathbb{R}^3$, $p \in S$, $v \in T_p(S)$, and $f: S \to \mathbb{R}$ is a differentiable function, then the directional derivative of f at p in direction v is $df_p(v)$. This is sometimes simply written as $v(f)(p)$, or $v(f)$ if the base point p is clear from context. In other words, $v(f)(p) := df_p(v)$.

Proposition. If w_1, w_2 are differentiable vector fields on S, then there exist a unique vector field w on S such that for any C^∞ function $f: S \to \mathbb{R}$, we have

$$w(f) = w_1(w_2(f)) - w_2(w_1(f)) \quad \text{everywhere on } S.$$

If w_1, w_2 are C^∞, then so is w.

This vector field w is called the commutator of w_1 and w_2, written $w = [w_1, w_2]$.

Proof. Let $\underline{x}: U \to \underline{x}(U) \subseteq S$ be a local parametrization. If $w_1 = a\,\underline{x}_u + b\,\underline{x}_v$ and $w_2 = c\,\underline{x}_u + d\,\underline{x}_v$ on $\underline{x}(U)$ (sometimes written as $w_1 = a\frac{\partial}{\partial u} + b\frac{\partial}{\partial v}$ and $w_2 = c\frac{\partial}{\partial u} + d\frac{\partial}{\partial v}$), then for any C^∞ $f: S \to \mathbb{R}$, if $F(u,v) := f(\underline{x}(u,v))$, then $w_1(f)(\underline{x}(u,v)) = a\frac{\partial F}{\partial u} + b\frac{\partial F}{\partial v}$, $w_2(f)(\underline{x}(u,v)) = c\frac{\partial F}{\partial u} + d\frac{\partial F}{\partial v}$,

so
$$w_1(w_2(f))(\underline{x}(u,v)) = ac\frac{\partial^2 F}{\partial u^2} + (ad+bc)\frac{\partial^2 F}{\partial u \partial v} + bd\frac{\partial^2 F}{\partial v^2} + \left(a\frac{\partial c}{\partial u} + b\frac{\partial c}{\partial v}\right)\frac{\partial F}{\partial u} + \left(a\frac{\partial d}{\partial u} + b\frac{\partial d}{\partial v}\right)\frac{\partial F}{\partial v}$$

and similarly
$$w_2(w_1(f))(\underline{x}(u,v)) = ac\frac{\partial^2 F}{\partial u^2} + (ad+bc)\frac{\partial^2 F}{\partial u \partial v} + bd\frac{\partial^2 F}{\partial v^2} + \left(c\frac{\partial a}{\partial u} + d\frac{\partial a}{\partial v}\right)\frac{\partial F}{\partial u} + \left(c\frac{\partial b}{\partial u} + d\frac{\partial b}{\partial v}\right)\frac{\partial F}{\partial v}$$

Hence $[w_1(w_2(f)) - w_2(w_1(f))](\underline{x}(u,v)) = w(f)(\underline{x}(u,v))$ if
$$w(\underline{x}(u,v)) = \left(a\frac{\partial c}{\partial u} + b\frac{\partial c}{\partial v} - c\frac{\partial a}{\partial u} - d\frac{\partial a}{\partial v}\right)\frac{\partial}{\partial u} + \left(a\frac{\partial d}{\partial u} + b\frac{\partial d}{\partial v} - c\frac{\partial b}{\partial u} - d\frac{\partial b}{\partial v}\right)\frac{\partial}{\partial v}$$
$$= (w_1(c) - w_2(a))\,\underline{x}_u + (w_1(d) - w_2(b))\,\underline{x}_v.$$
This shows
$$[a\underline{x}_u + b\underline{x}_v, c\underline{x}_u + d\underline{x}_v] = \left(\left(a\frac{\partial}{\partial u} + b\frac{\partial}{\partial v}\right)c\right)\underline{x}_u + \left(\left(a\frac{\partial}{\partial u} + b\frac{\partial}{\partial v}\right)d\right)\underline{x}_v - \left(\left(c\frac{\partial}{\partial u} + d\frac{\partial}{\partial v}\right)a\right)\underline{x}_u - \left(\left(c\frac{\partial}{\partial u} + d\frac{\partial}{\partial v}\right)b\right)\underline{x}_v.$$

If w_1, w_2 are C^∞, then so are $w_1(c) - w_2(a)$ and $w_1(d) - w_2(b)$, hence $w \in C^\infty$.

Examples: ① $[\underline{x}_u, \underline{x}_v] = 0$ since then $a = d = 1$, $b = c = 0$. ② $[w, w] = 0$ $\forall$ C^∞ vector field w.

③ $[\underline{x}_u, v\underline{x}_u - u\underline{x}_v] = -\underline{x}_v$.

We may now state two properties of our connection ∇.

1. If $\langle \cdot , \cdot \rangle_p$ is the inner product on $T_p(S)$ that determines the first fundamental form on S, then for any differentiable vector fields $w, \tilde{w}$ on S, and any $p \in S$, $v \in T_p(S)$, we have

$$v\langle w, \tilde{w}\rangle = \langle \nabla_v w, \tilde{w} \rangle + \langle w, \nabla_v \tilde{w} \rangle$$

Indeed, if $\gamma(t)$ is a C^∞ curve on S with $\gamma(0) = p$ and $\gamma'(0) = v$, then $\frac{d}{dt}\Big|_{t=0} \langle w(\gamma(t)), \tilde{w}(\gamma(t))\rangle = \langle \frac{d}{dt}\Big|_{t=0} w(\gamma(t)), \tilde{w}(p)\rangle + \langle w(p), \frac{d}{dt}\Big|_{t=0} \tilde{w}(\gamma(t))\rangle$, but $\nabla_v w - \frac{d}{dt}\Big|_{t=0} w(\gamma(t))$ is orthogonal to $T_p(S)$, hence to $\tilde{w}(p)$. Similarly for $\nabla_v \tilde{w} - \frac{d}{dt}\Big|_{t=0} \tilde{w}(\gamma(t))$.

We say ∇ is compatible with the metric on S.

2. For any C^∞ vector fields v, w on S, we have $\nabla_v w - \nabla_w v - [v, w] = 0$

We say ∇ is torsion free. In particular, $\nabla_v w = \nabla_w v$ if and only if $[v, w] = 0$.

In more advanced courses, a connection satisfying 1 and 2 is called a Levi-Civita connection. It is uniquely determined by the first fundamental form.

A further conceptual point

Let S be a regular surface in $\mathbb{R}^3$ and $p \in S$.

① If v_1, v_2 are vector fields on S, and $v_1(p) = v_2(p)$, then $df_p(v_1) = df_p(v_2)$ for any differentiable function f on S.

On the other hand, if f, g are differentiable functions on S, and we only know $f(p) = g(p)$, then it is not true that "$df_p(v) = dg_p(v)$ for any $v \in T_p(S)$".

What is true is that if $f = g$ in an open set on S containing p, then $df_p(v) = dg_p(v)$ for any $v \in T_p(S)$; also,

if $v \in T_p(S)$ and $f = g$ along a C^∞ curve γ on S with $\gamma(0) = p$, $\gamma'(0) = v$, then $df_p(v) = dg_p(v)$ for this particular v.

② If v_1, v_2 are vector fields on S, and $v_1(p) = v_2(p)$, then $\nabla_{v_1} w = \nabla_{v_2} w$ at p, for any differentiable vector field w on S.

On the other hand, if $w, \tilde{w}$ are differentiable vector fields on S, and we only know $w(p) = \tilde{w}(p)$, then it is <u>not</u> true that "$\nabla_v w = \nabla_v \tilde{w}$ at p for any $v \in T_p(S)$".

What is true is that if $w = \tilde{w}$ in an open set on S containing p, then $\nabla_v w = \nabla_v \tilde{w}$ at p for any $v \in T_p(S)$; also if $v \in T_p(S)$ and $w = \tilde{w}$ along a C^∞ curve γ on S with $\gamma(0) = p$, $\gamma'(0) = v$, then $\nabla_v w = \nabla_v \tilde{w}$ at p for this particular v.

This last point will allow us to covariant differentiate a vector field defined only along a curve, and parallel transport a vector along a C^∞ curve on a regular surface. Before we go into that, we digress to discuss one other consequence of our discussion of Christoffel symbols, with an application towards calculation of Gaussian curvatures.