Transformation of tangent and cotangent vectors

Suppose $f \in \Omega^0(M)$ is some fixed function, and $\gamma : \mathbb{R} \to M$ is some fixed path with $\gamma(0) = p$. Then

$$\mathbb{R} \xrightarrow{\gamma} M \xrightarrow{f} \mathbb{R}\,,$$

so

$$f \circ \gamma :\ \mathbb{R} \longrightarrow \mathbb{R}$$

is some fixed function. Its derivative at zero is some fixed number

$$(f \circ \gamma)'(0) \in \mathbb{R}.$$

Geometrically, this number represents the directional derivative of f in the direction of $\dot\gamma(0)$.

In local coordinates $x_1, \dots, x_n$ around p, we can write

$$f = f(x_1, \dots, x_n)$$

and

$$\gamma(t) = (\gamma_1(t), \dots, \gamma_n(t))\,.$$

(Note that for the moment, we are suppressing the coordinate chart map $h : U \to \mathbb{R}^n$. Technically, we should be writing $\gamma_i(t) = h_i(\gamma(t))$, but this would make our expressions too complicated.)

By the chain rule,

$$\begin{aligned}(f \circ \gamma)'(0) &= \left.\frac{\partial f}{\partial x_1}\right|_p \left.\frac{d\gamma_1}{dt}\right|_{t=0} + \left.\frac{\partial f}{\partial x_2}\right|_p \left.\frac{d\gamma_2}{dt}\right|_{t=0} + \cdots + \left.\frac{\partial f}{\partial x_n}\right|_p \left.\frac{d\gamma_n}{dt}\right|_{t=0}. \\ &= \underbrace{\left(\left.\frac{\partial f}{\partial x_1}\right|_p \varepsilon_1 + \left.\frac{\partial f}{\partial x_2}\right|_p \varepsilon_2 + \cdots + \left.\frac{\partial f}{\partial x_n}\right|_p \varepsilon_n \right)}_{df|_p} \cdot \underbrace{\left(\left.\frac{d\gamma_1}{dt}\right|_{t=0} e_1 + \left.\frac{d\gamma_2}{dt}\right|_{t=0} e_2 + \cdots + \left.\frac{d\gamma_n}{dt}\right|_{t=0} e_n \right)}_{\dot\gamma(0)},\end{aligned}$$

where $\cdot$ denotes the dual pairing

$$\varepsilon_i \cdot e_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

We abbreviate this formula by

$$(f \circ \gamma)'(0) = df|_p \cdot \dot\gamma(0).$$

This final answer on the right

- is independent of coordinates (because the left side doesn't involve coordinates), and
- depends only on the *first* derivatives of both f at p, and γ at $t = 0$ (from our coordinate expression).

We want to give coordinate-free definitions of tangent vectors $\dot\gamma(0)$ and cotangent vectors $df|_p$.

Definition. A *tangent vector* at $p \in M$ is an equivalence class of smooth paths $\gamma : \mathbb{R} \to M$ which satisfy $\gamma(0) = p$, where equivalence is determined by

$$[\gamma_1] \sim [\gamma_2] \iff \dot{\gamma}_1(0) = \dot{\gamma}_2(0) \text{ in some/any coordinate chart.}$$

(As a consequence of the transformation law for tangent vectors, we will see that if two tangent vectors agree in one chart, then they agree in every other chart.)

Definition. The set of such equivalence classes is called the *tangent space* at p, denoted by T_pM. Thus, if $v \in T_pM$, then we can write $v = \dot{\gamma}(0) = [\gamma]$ for some smooth path γ with $\gamma(0) = p$.

It's not obvious from this definition, but T_pM is a vector space. Given two paths $\gamma_1, \gamma_2 : \mathbb{R} \to M$ with $\gamma_i(0) = p$, we need to define addition $[\gamma_1] + [\gamma_2]$ as some vector $[\gamma_3]$. However, it's nonsensical to define $\gamma_3 = \gamma_1 + \gamma_2$ since we can't add points in a manifold. However, if we choose a coordinate chart for which $h(p) = \vec{0}$, then it makes sense to add the local coordinates of γ_1 and γ_2 to obtain local coordinates for γ_3. Scalar multiplication is similarly defined via a choice of local coordinates.

Another technicality is that a path γ may not remain in a single coordinate chart for all t. Since are concerned with the behavior around $\gamma(0) = p$, we need only consider paths $\gamma : [-\varepsilon, \varepsilon] \to M$.

Cotangent vectors behave more nicely. We give a similar definition:

Definition. A *cotangent vector* at $p \in M$ is an equivalence class of functions $f \in \Omega^0(M)$ with

$$[f_1] = [f_2] \iff df_1|_p = df_2|_p$$

in some/any coordinate chart.

Definition. The set of such equivalence classes is called the *cotangent space* at p, denoted by T_p^*M.

There is no difficulty in defining the vector space structure of T_p^*M:

$$\alpha [f_1] + \beta [f_2] = [\alpha f_1 + \beta f_2].$$

At this point, we should understand how vectors and covectors transform under change of coordinates, and verify that our definitions don't depend on the choice of coordinates.

Using the coordinates $h_1 = (x_1, \ldots, x_n)$, we have

$$df = \frac{\partial f}{\partial x_1}\varepsilon_1 + \frac{\partial f}{\partial x_2}\varepsilon_2 + \cdots + \frac{\partial f}{\partial x_n}\varepsilon_n. \tag{1}$$

If we instead use coordinates $h_2 = (y_1, \ldots, y_n)$, then

$$df = \frac{\partial f}{\partial y_1}\tilde{\varepsilon}_1 + \frac{\partial f}{\partial y_2}\tilde{\varepsilon}_2 + \cdots + \frac{\partial f}{\partial x_n}\tilde{\varepsilon}_n,$$

where $\tilde{\varepsilon}_i$ are the covectors under h_2. From the chain rule

$$\frac{\partial f}{\partial y_i} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial y_i} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial y_i},$$

we have

$$
\begin{aligned}
df &= \left(\frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial y_1} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial y_1}\right)\tilde{\varepsilon}_1 + \\
&\quad \vdots \\
&\quad + \left(\frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial y_n} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial y_n}\right)\tilde{\varepsilon}_n \\
&= \frac{\partial f}{\partial x_1}\left(\frac{\partial x_1}{\partial y_1}\tilde{\varepsilon}_1 + \cdots + \frac{\partial x_1}{\partial y_n}\tilde{\varepsilon}_n\right) + \\
&\quad \vdots \\
&\quad + \frac{\partial f}{\partial x_n}\left(\frac{\partial x_n}{\partial y_1}\tilde{\varepsilon}_1 + \cdots + \frac{\partial x_n}{\partial y_n}\tilde{\varepsilon}_n\right).
\end{aligned} \tag{2}
$$

By comparing (1) and (2), we see that our expressions for different coordinate charts are consistent if we identify

$$\varepsilon_i \longleftrightarrow \frac{\partial x_i}{\partial y_1}\tilde{\varepsilon}_1 + \cdots + \frac{\partial x_i}{\partial y_n}\tilde{\varepsilon}_n.$$

This justifies the notation $\varepsilon_i = dx_i$ and $\tilde{\varepsilon}_i = dy_i$, for then

$$dx_i = \frac{\partial x_i}{\partial y_1}dy_1 + \cdots + \frac{\partial x_i}{\partial y_n}dy_n.$$

A similar transformation law for $x_i = \gamma_i(t)$ suggests that

$$\frac{d\gamma}{dt} = \frac{d\gamma_1}{dt}e_1 + \cdots + \frac{d\gamma_n}{dt}e_n$$

be written using

$$e_i = \frac{\partial}{\partial x_i},$$

because under a change of variables,

$$\frac{\partial}{\partial x_i} = \frac{\partial y_1}{\partial x_i}\frac{\partial}{\partial y_1} + \cdots + \frac{\partial y_n}{\partial x_i}\frac{\partial}{\partial y_n}.$$

This notation identifies a tangent vector with its corresponding directional derivative operator! Be warned that order matters, since $\frac{\partial}{\partial y_j}\frac{\partial y_j}{\partial x_i}$ denotes a second partial derivative, while $\frac{\partial y_j}{\partial x_i}\frac{\partial}{\partial y_j}$ is a differential operator.

Assembling the above results, we have

Transformation law The cotangent vector dx_i transforms linearly into the $\{dy_j\}$ basis according to the matrix $\left[\frac{\partial x_i}{\partial y_j}\right]$. The tangent vector $\frac{\partial}{\partial x_i}$ transforms linearly into the $\left\{\frac{\partial}{\partial y_j}\right\}$ basis according to the matrix $\left[\frac{\partial y_j}{\partial x_i}\right]$.

These are Jacobian matrices for the coordinate transformations, going in opposite directions. Therefore, the matrices are inverses of each other.

Our definitions of tangent and cotangent vectors are well-defined, because equality in one coordinate chart differs from equality in another by an invertible linear map.

Theorem. *If M^n is a smooth manifold, then T_pM and T_p^*M are dual spaces of dimension n. The duality pairing of $v \in T_pM$ with $\alpha \in T_p^*M$ is given by directional differentiation*

$$\alpha \cdot v := (f \circ \gamma)'(0) \text{ where } \alpha = [f] \text{ and } v = [\gamma].$$

Proof. Since these computations involve only first-derivatives, which are uniquely determined by α and v, it suffices to check this formula for any particular representatives for α and v.

Consider a local coordinate chart $U \subset V = \mathbb{R}^n$. For simplicity, suppose that the coordinates for p are $\vec{0}$. As usual, let $\left\{e_i = \frac{\partial}{\partial x_i}\right\}_{i=1}^n$ denote the standard basis of V, and $\{\varepsilon_i = dx_i\}_{i=1}^n$ the dual basis of V^*. Under this coordinate chart, any velocity vector $v \in T_pM$ can be uniquely expressed as $v = \sum v_i e_i$, and any covector can be uniquely expressed as $\alpha = \sum \alpha_i \varepsilon_i$. The dual pairing is

$$\alpha \cdot v = \left(\sum \alpha_i \varepsilon_i\right) \cdot \left(\sum v_j e_j\right) = \sum \alpha_i v_i.$$

We need to pick representatives $[f]$ for α and $[\gamma]$ for v, and then verify the identity

$$\alpha \cdot v \stackrel{?}{=} (f \circ \gamma)'(0).$$

Set $\gamma = (v_1 t, v_2 t, \dots, v_n t)$, and $f = \alpha_1 x_1 + \cdots + \alpha_n x_n$. By construction, it's clear that $\dot{\gamma}(0) = v$, and $df|_p = \alpha$. We compute that $f \circ \gamma(t) = \alpha_1 v_1 t + \cdots + \alpha_n v_n t$, so that $(f \circ \gamma)'(0) = \sum \alpha_i v_i = \alpha \cdot v$, as desired. □

Restoring the coordinate charts

Suppose we have a coordinate chart (U_1, h_1) around $p \in M$ so that $h_1(p) = \vec{0}$. Let $x_1, \dots, x_n$ denote the local coordinates of $h_1(U_1) \subset \mathbb{R}^n$. Given a function $f \in C^\infty(M)$, the corresponding function on $C^\infty(h_2(U))$ is $f \circ h_1^{-1}(x_1, \dots, x_n)$, which is denoted by $f_1 = (h_1^{-1})^*(f)$. Similarly, if (U_2, h_2) is another coordinate chart around p, the local representative for f in $C^\infty(h_2(U_2))$ is $f_2 = (h_2^{-1})^*(f)$.

The transition function from $h_1(U_1 \cap U_2)$ to $h_2(U_1 \cap U_2)$ is

$$h_1(U_1 \cap U_2) \xrightarrow{h_{21}} h_2(U_1 \cap U_2)$$

given by

$$h_{21} = h_2 \circ h_1^{-1}.$$

It relates f_1 and f_2 via

$$f_2 = (h_{21}^{-1})^*(f_1).$$