General Relativity by Robert M. Wald

Chapter 2: Manifolds and Tensor Fields

2.2. Vectors (Partial)

Note. In this section, we introduce a way to discuss vectors tangent to a manifold intrinsically (that is, without an appeal to a "hyperspace" in which the manifold is embedded—curvature will have to be dealt with similarly).

Note. In Calculus III (MATH 2110), we used a vector (a *unit* vector) to define a directional derivative in $\mathbb{R}^2$ and $\mathbb{R}^3$. See Section 14.5 of my Calculus III notes (`http://faculty.etsu.edu/gardnerr/2110/notes-12e/c14s5.pdf`). For $(v^1, v^2, \ldots, v^n) \in \mathbb{R}^n$ we have the directional derivative operator on $f(x^1, x^2, \ldots, x^n)$ defined as $\displaystyle\sum_{\mu=1}^{n} v^n \frac{\partial}{\partial x^\mu}[f]$ (and conversely, any directional derivative corresponds to a vector). Wald states (page 15) that "Directional derivatives are characterized by their linearity and 'Leibniz's Rule' [a version of the Product Rule] behavior when acting on functions." This motivates the following definition.

Definition. For manifold M, let $\mathcal{F}$ be the collection of all C^∞ function from M into $\mathbb{R}$. A *tangent vector v at a point $p \in M$* is a function $v : \mathcal{F} \to \mathbb{R}$ which satisfies:

(1) Linearity: $v(af + bg) = av(f) + bv(g)$ for all $f, g \in \mathcal{F}$ and $a, b \in \mathbb{R}$, and

(2) Leibniz Rule: $v(fg) = f(p)v(g) + g(p)v(f)$.

Note. Notice that the only place the point p plays a role in the definition of a tangent vector is in "Leibniz's Rule."

Note. If $h \in \mathcal{F}$ is a constant function, say $h(q) = c$ for all $q \in M$, then at point p by Leibniz's Rule

$$v(h^2) = v(h\,h) = h(p)v(h) + h(p)v(h) = 2cv(h)$$

and

$$\begin{aligned} v(h^2) &= v(ch) \text{ since } h(q) = c \\ &= cv(h). \end{aligned}$$

So $v(h^2) = 2cv(h) = cv(h)$, and so $v(h) = 0$.

Definition. Let V_p denote the collection of all tangent vectors at p to manifold M. For $a, b \in \mathbb{R}$, define the linear combination $av_1 + bv_2 \in V_p$ as

$$(av_1 + bv_2)(f) = av_1(f) + bv_2(f)$$

for all $f \in \mathcal{F}$.

Note. V_p is "clearly" a vector space (a vector space of linear operators on $\mathcal{F}$). The following result confirms that if manifold M is of dimension n, then V_p is of dimension n as well.

Theorem 2.2.1. Let M be an n-dimensional manifold. Let $p \in M$ and let V_p denote the tangent space at p. Then $\dim(V_p) = n$.

Proof. We will construct a basis for V_p. Let $\psi : O \to U \subset \mathbb{R}^n$ be a chart with $p \in O$.

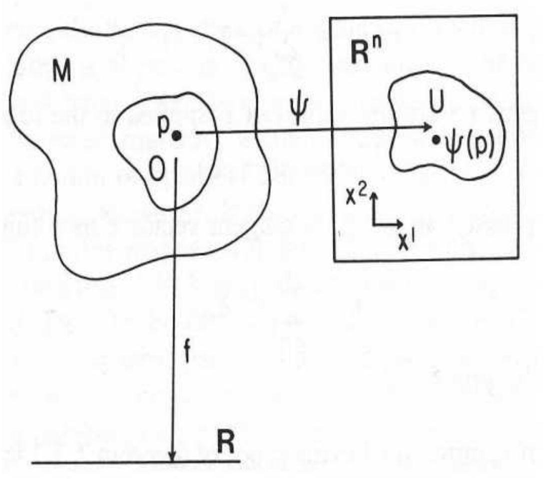

Figure 2.3 from Wald, page 15.

If $f \in \mathcal{F}$ then by the definition of "C^∞ function" we have $f \circ \psi^{-1} : U \to \mathbb{R}$ is C^∞. (We defined $f : M \to M'$ as C^∞ in Section 2.1 and involved $\psi'_\beta : M' \to \mathbb{R}^n$, but here $M' = \mathbb{R}$ so we take ψ'_β as the identity and $\psi'_\beta \circ f \circ \psi^{-1} = f \circ \psi^{-1}$.) For $\mu = 1, 2, \ldots, n$ define $X_\mu : \mathcal{F} \to \mathbb{R}$ by

$$X_\mu(f) = \frac{\partial}{\partial x^\mu} \left[f \circ \psi^{-1} \right] \Big|_{\psi(p)}$$

where $(x^1, x^2, \ldots, x^\mu)$ are the Cartesian coordinates of $\mathbb{R}^n$. Notice that

$f \circ \psi^{-1} : U \to \mathbb{R}$ and $U \subset \mathbb{R}^n$, so in fact $f \circ \psi^{-1}$ is a function of $x^1, x^2, \ldots, x^n$. Then, since $X_1, X_2, \ldots, X_n$ are defined using partial derivatives, then X_1, X_2, $\ldots, X_n$ satisfy linearity and Leibniz's Rule and so are tangent vectors. To see that $X_1, X_2, \ldots, X_n$ are linearly independent, consider $f_\mu(x) = x^\mu$ for $\mu = 1, 2, \ldots, n$. Then

$$X_i(f_i) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

and so $(a_1 X_1 + a_2 X_2 + \cdots + a_n X_n)(f_i) = 0$ if and only if $a_i = 0$. So by applying $a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$ to $f_1, f_2, \ldots, f_n$ and setting each equal to 0 implies that $a_1 = a_2 = \cdots = a_n = 0$. So $X_1, X_2, \ldots, X_n$ are linearly independent.

By Problem 2.2, if $F : \mathbb{R}^n \to \mathbb{R}$ is C^∞, then for each $a = (a^1, a^2, \ldots, a^n) \in \mathbb{R}^n$, there exists C^∞ functions H_μ such that for all $x \in \mathbb{R}^n$ we have

$$F(x) = F(a) + \sum_{\mu=1}^{n} (x^\mu - a^\mu) H_\mu(x) \text{ and } H_\mu(a) = \left.\frac{\partial F}{\partial x^\mu}\right|_{x=a}. \qquad (2.2.3/2.2.4)$$

We take $F = f \circ \psi^{-1} : \mathbb{R}^n \to \mathbb{R}$ and $a = \psi(p)$ to get from Problem 2.2 that for all $q \in O$ (where $\psi(q) = x \in \mathbb{R}^n$; think of both x and q as variables)

$$F(x) = (f \circ \psi^{-1})(\psi(q)) = f(q) = F(\psi(p)) + \sum_{\mu=1}^{n} [x^\mu \circ \psi(q) - x^\mu \circ \psi(p)] H_\mu(\psi(q))$$

where $x^\mu \circ \psi(q)$ denotes the μ-th coordinate of $\psi(q) \in \mathbb{R}^n$. Also, $F(\psi(p)) = (f \circ \psi^{-1})(\psi(p)) = f(p)$, so

$$f(q) = f(p) + \sum_{\mu=1}^{n} (x^\mu \circ \psi(q) - x^\mu \circ \psi(p)) H_\mu(\psi(q)). \qquad (2.2.4)$$

Let $v \in V_p$. We now show that v is a linear combination of $X_1, X_2, \ldots, X_n$ (and hence $X_1, X_2, \ldots, X_n$ is a basis for V_p). Let $f \in \mathcal{F}$. We have

$$v(f) \;=\; v(f(q))|_{q=p}$$

$$
\begin{aligned}
&= v\left[f(p) + \sum_{\mu=1}^{n}[x^\mu \circ \psi(q) - x^\mu \circ \psi(p)] H_\mu(\psi(q))\right] \text{ by (2.2.4)} \\
&= v[f(p)] + \sum_{\mu=1}^{n} v\left[[x^\mu \circ \psi(q) - x^\mu \circ \psi(p)] H_\mu(\psi(q))\right] \text{ since } v \text{ is linear} \\
&= v[f(p)] + \sum_{\mu=1}^{n} \Big\{ (x^\mu \circ \psi(q) - x^\mu \circ \psi(p))|_{q=p}\, v[H_\mu(\psi(q))] \\
&\quad + v[x^\mu \circ \psi(q) - x^\mu \circ \psi(p)] H_\mu(\psi(q))|_{q=p} \Big\} \text{ by Leibniz's Rule} \\
&= 0 + \sum_{\mu=1}^{n} \{0 + v[x^\mu \circ \psi(q) - x^\mu \circ \psi(p)] H_\mu(\psi(p))\} \text{ since } f(p) \text{ is constant} \\
&= \sum_{\mu=1}^{n} v[(x^\mu \circ \psi)(q)](H_\mu \circ \psi)(p) \text{ since } (x^\mu \circ \psi)(p) \text{ is constant.}
\end{aligned}
$$

But by equation (2.2.3),

$$H_\mu \circ \psi(p) = H_\mu(a) = \left.\frac{\partial F}{\partial x^\mu}\right|_{x=a} = X_\mu(f)|_{x=a}.$$

So

$$v(f) = \sum_{\mu=1}^{n} v((x^\mu \circ \psi)(q))\, X_\mu(f)|_{x=a}.$$

With v^μ set equal to $v((x^\mu \circ \psi)(q))$ we have

$$v(f) = \left(\sum_{\mu=1}^{n} v^\mu X_\mu\right)(f)$$

and so $v \sum_{\mu=1}^{n} v^\mu X_\mu$ and $X_1, X_2, \ldots, X_n$ is a spanning set for V_p. Therefore, $X_1, X_2, \ldots, X_n$ is a basis for V_p. ▮

Definition. The basis $\{X_\mu\}_{\mu=1}^{n}$ of V_p (the n-dimensional tangent space to M at p) of Theorem 2.2.1 is a *coordinate basis.*

Note. Notice that

$$X_\mu(f) = \frac{\partial}{\partial x^\mu}\left[f \circ \psi^{-1}\right]\bigg|_{\psi(p)}$$

so basis vector X_μ depends on the coordinate system ψ. We could use a different coordinate system ψ' to produce a different coordinate basis $\{X'_\nu\}$ at point p. We then want to relate the coordinate bases using the Chain Rule.

Last updated: June 19, 2016