§4 Tangent vectors and related objects

Last updated: 23 August (5 September) 2018.

4.1 Tangent space

4.1.1 Definitions of a tangent vector and tangent space

Consider a manifold $M = M^n$. Consider a point $x \in M$ of a manifold $M$.

**Definition 4.1.** A tangent vector $v$ at a point $x \in M$ is defined as a rule assigning to each coordinate system $x^1, \ldots, x^n$ near $x$ an array of numbers $(v^1, \ldots, v^n)$ called the components or coordinates of $v$, so that for any two coordinate systems, say, $x^1, \ldots, x^n$ and $x^1', \ldots, x^n'$ the respective components $v^i$ and $v'^i$ are related by the transformation law

$$v^i = \frac{\partial x^i}{\partial x'^j}(x^1', \ldots, x^n') v'^j.$$  

(1)

Here the partial derivatives are taken at $(x^1', \ldots, x^n') \in \mathbb{R}^n$ corresponding to the point $x \in M$ (the “vector law”).

The transformation law for components of vectors defined by (1) is called the **vector law**. Note that it depends on a point of $M$, since the Jacobi matrix is not, in general, a constant matrix. Therefore tangent vectors are attached to points. It makes no sense to speak of a ‘vector’ on a manifold without referring to a particular point. This is a big difference with $\mathbb{R}^n$.

The set of all tangent vectors at $x \in M$ is called the **tangent space** at $x$ and denoted $T_x M$.

4.1.2 Motivation: velocity vectors

Consider a smooth manifold $M = M^n$. A (smooth) curve in $M$ is a smooth map $\gamma: (a, b) \to M$. In other words, we have $t \mapsto x(t) \in M$. How is it possible to define the velocity of $\gamma$?

**Definition 4.2.** For a curve $\gamma: (a, b) \to M$, which we write as $x = x(t)$, we define the velocity vector

$$\dot{x} = \frac{dx}{dt}$$

at $t \in (a, b)$ to be the array of numbers

$$(\dot{x}^1, \ldots, \dot{x}^n) = \left(\frac{dx^1}{dt}, \ldots, \frac{dx^n}{dt}\right)$$

given for each coordinate system $x^1, \ldots, x^n$, where $x^i = x^i(t)$, $i = 1, \ldots, n$ is the coordinate expression of the curve $x = x(t)$. We may also use the notation such as $\dot{\gamma}$ or $d\gamma/dt$ for the velocity of a curve $\gamma$.

Let us see how the components of velocity in different coordinate systems are related with each other. Suppose we have coordinates that we denote by $x^1, \ldots, x^n$ and another coordinates that we denote by $x^1', \ldots, x^n'$. (We shall refer to them as ‘old’ and ‘new’ coordinates, though these names do not carry any absolute meaning.) Then in the old coordinates the components of
velocity are $\dot{x}^i$, where $i = 1, \ldots, n$, and in the new coordinates they are $\dot{x}'^i$, where $i' = 1', \ldots, n'$. By the chain rule we have

$$\dot{x}^i = \sum_{i'} \frac{\partial x^i}{\partial x'^{i'}} \dot{x}'^{i'}$$

(2)

where the Jacobi matrix is taken at $x(t)$. (More precisely, if the old coordinates are written as functions of the new coordinates as $x^i = x^i(x'^1, \ldots, x'^n)$, the partial derivatives are taken at $x'^1, \ldots, x'^n$ corresponding to the point $x(t)$.) We see that the transformation law (2) depends on a point of $M$ (it varies from point to point).

Therefore, the velocity vector is indeed a tangent vector. Conversely, any tangent vector can be interpreted as the velocity vector for some curve.

The velocity vector as described above corresponds to the usual understanding of velocity in familiar cases.

Let us first examine the case of $M = \mathbb{R}^n$. Then in a straightforward way we can consider

$$\frac{x(t + \Delta t) - x(t)}{\Delta t}.$$ 

The top of the fraction is a vector in $\mathbb{R}^n$ (as the difference of two points), the bottom is a number; hence the whole thing is a vector in $\mathbb{R}^n$. It makes sense to pass to the limit when $\Delta t \to 0$ and, assuming the limit exists, we obtain the vector

$$\dot{x} = \frac{dx}{dt} := \lim_{\Delta \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t},$$

which, by definition, is the velocity of our curve. Unfortunately, we cannot do the same for an arbitrary manifold $M$, because taking difference of two points will not make sense. The only tool we have for manifolds are coordinates. Therefore we have to examine the case of $\mathbb{R}^n$ further and see how velocity can be expressed using coordinates.

Writing $x = (x^1, \ldots, x^n)$ and working out the difference $x(t + \Delta t) - x(t)$, we arrive at

$$(x^1(t + \Delta t) - x^1(t), \ldots, x^n(t + \Delta t) - x^n(t)).$$

Therefore calculating the velocity amounts to taking the time derivative of each coordinate, and the velocity vector of a curve in $\mathbb{R}^n$ has the coordinate expression

$$\dot{x} = \frac{dx}{dt} = (\dot{x}^1, \ldots, \dot{x}^n)$$

in standard coordinates. Introducing the vectors $e_i=(0, \ldots, 1, \ldots, 0)$ of the standard basis, with 1 in the $i$-th position and 0 everywhere else, we can re-write this also as

$$\dot{x} = \dot{x}^1 e_1 + \ldots \dot{x}^n e_n.$$ 

**Example 4.1.** For a curve in $\mathbb{R}^2$ we have

$$\dot{x} = (\dot{x}, \dot{y}) = \dot{x} e_x + \dot{y} e_y$$

using the traditional notation for coordinates, where $e_x = (1, 0)$, $e_y = (0, 1)$ is the standard basis (which is otherwise denoted $e_1, e_2$).
Example 4.2. Let us work out the expression for velocity of a curve in $\mathbb{R}^2$ in polar coordinates $r, \theta$ rather than Cartesian coordinates $x, y$. Suppose a curve $x = x(t)$ is defined using polar coordinates as $r = r(t), \theta = \theta(t)$. To calculate the velocity $\dot{x}$ we use the chain rule: a point $x \in \mathbb{R}^2$ is considered as a composite function of the time, first $x = x(r, \theta)$ and then $r = r(t), \theta = \theta(t)$. We obtain

$$\dot{x} = \dot{r} \frac{\partial x}{\partial r} + \dot{\theta} \frac{\partial x}{\partial \theta},$$

where partial derivatives w.r.t. $r$ and $\theta$ are vectors, and $\dot{r}, \dot{\theta}$ are scalar coefficients. If we introduce the vectors

$$e_r := \frac{\partial x}{\partial r} \quad \text{and} \quad e_\theta := \frac{\partial x}{\partial \theta},$$

we may write

$$\dot{x} = \dot{r} e_r + \dot{\theta} e_\theta,$$

very similarly to the expression in Cartesian coordinates.

Note that the vectors $e_r, e_\theta$ associated with polar coordinates on $\mathbb{R}^2$ depend on a point $x \in \mathbb{R}^2$ and make a basis for each $x$ except for the origin, where $e_r$ is not defined and $e_\theta$ vanishes. (Indeed, one can find $e_r = (\cos \theta, \sin \theta), e_\theta = (-r \sin \theta, r \cos \theta)$ in standard coordinates.) Hence the expression $\dot{x} = \dot{r} e_r + \dot{\theta} e_\theta$ is nothing but the expansion over this basis (at a given point $x$). On the other hand, the vectors of the standard basis in $\mathbb{R}^2$ can be also written as partial derivatives w.r.t. the coordinates:

$$e_1 = (1, 0) = \frac{\partial}{\partial x}(x, y) = \frac{\partial x}{\partial x}, \quad \text{and} \quad e_2 = (0, 1) = \frac{\partial}{\partial y}(x, y) = \frac{\partial x}{\partial y}.$$

Generalizing to $\mathbb{R}^n$, we conclude that for arbitrary curvilinear coordinates $y^1, \ldots, y^n$ (we use a different letter to distinguish from standard coordinates denoted above as $x^1, \ldots, x^n$) there is an associated basis of vectors

$$e_1 := \frac{\partial x}{\partial y^1}, \ldots, e_n := \frac{\partial x}{\partial y^n},$$

(although we use the same letters $e_i$, these vectors are not to be confused with the standard basis vectors above), depending on a point with respect to which the velocity of any curve in $\mathbb{R}^n$ has the expansion

$$\dot{x} = \dot{y}^1 e_1 + \ldots + \dot{y}^n e_n.$$

In other words, we may say that in an arbitrary coordinate system on $\mathbb{R}^n$, the array of the time derivatives of the coordinates

$$(\dot{y}^1, \ldots, \dot{y}^n)$$

gives the components of the velocity $\dot{x}$ in this system. This justifies Definition 4.2 for manifolds.

---

1 The linear independence of the vectors $e_i$ for each $x$ is, in fact, a part of the definition of what is a ‘system of curvilinear coordinates’ on a domain of $\mathbb{R}^n$, the other principal condition being that the correspondence between points and coordinates is one-to-one.
4.1.3 Properties of tangent space

**Theorem 4.1.** The tangent space $T_xM$ is non-empty. With respect to the evident operations (defined component-wise), it is a vector space of dimension $n = \dim M$, for each $x \in M$.

*Proof.* To prove that $T_xM$ is non-empty, one can simply bring up an example of the velocity vector for some curve through $x$. More precisely, if $\gamma: t \mapsto x(t)$ is a curve such that $x(t_0) = x$ (for some $t = t_0$), then $\dot{x}(t_0)$ is an element of $T_xM$. An alternative way of showing the existence of tangent vectors could be as follows. Fix a coordinate system near $x$; suppose the coordinates are denoted $x_1, \ldots, x^n$. To define a tangent vector at $x$, take an arbitrary array $(v^1, \ldots, v^n) \in \mathbb{R}^n$. We want to view $v^i$, $i = 1, \ldots, n$, as the components of a tangent vector w.r.t. the coordinate system $x^i$. We need to define the components of $v$ in all other coordinate systems; we do so by setting

$$v^i = \sum \frac{\partial x^i'}{\partial x^i} v^i$$

(i.e., by using the vector law). For consistency, we need to check that the vector law holds for any two arbitrary coordinate systems, say, $x^i$ and $x^{i'}$. We have

$$v^{i''} = \sum \frac{\partial x^{i''}}{\partial x^{i'}} v^i$$

and can express $v^i$ in terms of $v^{i'}$:

$$v^i = \sum \frac{\partial x^i}{\partial x^{i'}} v^{i'}.$$

Combining these formulas, we arrive at

$$v^{i''} = \sum \sum \frac{\partial x^{i''}}{\partial x^{i'}} \frac{\partial x^i}{\partial x^{i'}} v^{i'} = \sum \frac{\partial x^{i''}}{\partial x^{i'}} v^i,$$

(where we used the chain rule).

Now we need to prove that $T_xM$ has the structure of a vector space. To define it, fix again some coordinate system and define the sum of two vectors and the multiplication of a vector by a number componentwise:

$$(u + v)^i := u^i + v^i, \quad (ku)^i := ku^i$$

where $u, v \in T_xM$ and $k \in \mathbb{R}$. We have to check that this definition does not depend on a choice of coordinates. Indeed, transform $u + v$ and $ku$ into another coordinate system using the vector law:

$$(u + v)^i = \sum \frac{\partial x^i'}{\partial x^i} (u + v)^i = \sum \frac{\partial x^i'}{\partial x^i} (u^i + v^i) = \sum \frac{\partial x^i'}{\partial x^i} u^i + \sum \frac{\partial x^i'}{\partial x^i} v^i = u^{i'} + v^{i'},$$

which shows that the expression for the sum of two vectors will have the same form in all coordinate systems. This holds for the multiplication of a vector by a number as well. □
A conclusion from the proof above is that to define a tangent vector it is sufficient to define its components in a particular coordinate system. Such components can be chosen absolutely arbitrarily — no restrictions. Hence, the tangent space $T_x M$ can be identified with $\mathbb{R}^n$. (Such an identification is not unique and depends on a choice of coordinates near $x$.)

Remark 4.1. We used velocities of parametrized curves as a model for defining tangent vectors. In fact, for a given point $x \in M$, every vector $v \in T_x M$ is the velocity vector for some curve through $x$. Indeed, in an arbitrary (but fixed) coordinate system near $x$ set

$$x^i(t) = x^i + tv^i,$$

where $x^i$ are the coordinates of the point $x$ and $v^i$, the components of the vector $v$. It is a curve passing through $x$ at $t = 0$ and we have $\dot{x}(0) = v$. (The curve looks as a ‘straight line’ in the chosen coordinate system, but in another chart it would have a different appearance.)

Each coordinate system near $x \in M$ defines a basis in $T_x M$. The basis vectors are $e_i = \frac{\partial x}{\partial x_i}$. Recall that partial derivatives are defined as follows: all independent variables except one are fixed and only one is allowed to vary; then the partial derivative is the ordinary derivative w.r.t. this variable. Hence the vectors $e_i = \frac{\partial x}{\partial x_i}$ are precisely the velocity vectors of the coordinate lines, i.e., the curves obtained by fixing all coordinates but one, which is the parameter on the curve. By applying the definition of the velocity vector we see that in the coordinate system $x^1, \ldots, x^n$ the basis vectors $e_1, \ldots, e_n$ are represented by the standard basis vectors of $\mathbb{R}^n$:

$$e_1 = \frac{\partial x}{\partial x^1} \quad \text{has components} \quad \frac{\partial}{\partial x^1}(x^1, \ldots, x^n) = (1, 0, \ldots, 0)$$

$$e_2 = \frac{\partial x}{\partial x^2} \quad \text{has components} \quad \frac{\partial}{\partial x^2}(x^1, \ldots, x^n) = (0, 1, \ldots, 0)$$

$$\vdots$$

$$e_n = \frac{\partial x}{\partial x^n} \quad \text{has components} \quad \frac{\partial}{\partial x^n}(x^1, \ldots, x^n) = (0, 0, \ldots, 1)$$

Hence $e_1, \ldots, e_n$ is indeed a basis in $T_x M$ (for all $x$ in the region where this coordinate system is defined). It is called the coordinate basis (for a given coordinate system), or the basis associated with it.

4.1.4 Practical description of tangent vectors

Consider practical ways of describing tangent vectors.

Firstly, if we are given a chart, the tangent space at a point $x$ is the linear span of the coordinate basis vectors $e_i = \frac{\partial x}{\partial x_i}$. This is helpful, in particular, if our manifold is defined as a parametrized surface in some $\mathbb{R}^N$.

Example 4.3. For $S^2 \subset \mathbb{R}^3$ consider a parametrization by angles $\theta, \varphi$:

$$x = \cos \varphi \sin \theta, \quad y = \sin \varphi \sin \theta, \quad z = \cos \theta.$$

We obtain the vectors $e_\theta, e_\varphi$,

$$e_\theta = (\cos \varphi \cos \theta, \sin \varphi \cos \theta, -\sin \theta)$$

$$e_\varphi = (- \sin \varphi \sin \theta, \cos \varphi \sin \theta, 0)$$

(as elements of $\mathbb{R}^3$). Hence the tangent space $T_x S^2$ at $x \in S^2$ can be described as the subspace of $\mathbb{R}^3$ spanned by $e_\theta, e_\varphi$. 

5
Secondly, if a manifold is specified by equations, then the tangent space at each point appears as the space of solutions of the corresponding “linearized” system. Suppose $M$ is defined by a system of $k$ equations in $\mathbb{R}^N$

$$f^1(x^1, \ldots, x^N) = 0$$

$$\ldots$$

$$f^k(x^1, \ldots, x^N) = 0.$$  

Then to obtain the equations for the tangent space we assume that a vector $v \in T_{\mathbf{x}}M$ is the velocity of a curve $\mathbf{x}(t)$ on $M$. By differentiating the above system w.r.t. $t$, we arrive at the linear system

$$\sum_{a=1}^{N} \frac{\partial f^\mu}{\partial x^a} \dot{x}^a = 0, \quad (\mu = 1, \ldots, k)$$

or

$$\sum_{a=1}^{N} \frac{\partial f^\mu}{\partial x^a} v^a = 0, \quad (\mu = 1, \ldots, k)$$

specifying a subspace of $\mathbb{R}^N$. It is exactly $T_{\mathbf{x}}M$. (Note that the coefficients of the matrix of this system are functions of a point $\mathbf{x} \in M$.) The dimension of the space of solutions is $N - r$, where $r$ is the rank of the matrix $\frac{\partial f^\mu}{\partial x^a}$.

Theorem 1.5 of Section 1 can now interpreted as follows: for a set specified in $\mathbb{R}^N$ by a system of equations as above, without any conditions for the rank, we may still define a tangent space at each point as the space of solutions of the linear system; the condition that the rank is constant is therefore equivalent to the condition that the dimension of the tangent space is the same for all points (as we expect for a manifold).

**Example 4.4.** For the sphere $S^2 \subset \mathbb{R}^3$, if we specify it by the equation

$$x^2 + y^2 + z^2 = 1,$$

we obtain by differentiating

$$x\dot{x} + y\dot{y} + z\dot{z} = 0,$$

(after dividing by 2). For each $\mathbf{x} = (x, y, z) \in S^2$, it is the equation of the tangent space $T_{\mathbf{x}}S^2$ as a subspace of $\mathbb{R}^3$ (of dimension 2).

**Example 4.5.** Consider the set of all real orthogonal matrices $n \times n$. It is specified by the matrix equation

$$AA^T = E$$

(where $E$ denotes the identity matrix) and has the standard notation $O(n)$. Consider a curve $A = A(t)$. By differentiating w.r.t. $t$ we obtain the equation

$$\dot{A}A^T + AA^T = 0$$

or

$$(\dot{A}A^T) + (\dot{A}A^T)^T = 0.$$
In other words, a matrix $\dot{A}$ belongs to $T_A O(n)$ if and only if the matrix $B = \dot{A}A^T = \dot{A}A^{-1}$ is antisymmetric. We have an isomorphism of $T_A O(n)$ with the vector space of all antisymmetric matrices $n \times n$, for all $A \in O(n)$. Therefore the dimension of $T_A O(n)$ is the same for all $A$, the system of equations specifying $O(n)$ has constant rank, and by Theorem 1.5 of Section 1, $O(n)$ is a manifold. (Check that its dimension is $n(n - 1)/2$.)

### 4.2 Tangent bundle and tangent maps

#### 4.2.1 Tangent bundle

**Definition 4.3.** The union of the tangent spaces $T_x M$ for all points $x \in M$, i.e., the collection of all tangent vectors to $M$, is called the tangent bundle of $M$ and denoted $TM$,

$$
TM = \bigcup_{x \in M} T_x M.
$$

(3)

Let us emphasize that tangent spaces at different points are, by the definition, different vector spaces, which cannot have common elements. Hence the union (3) is a disjoint union.

**Remark 4.2.** One should not be confused by the picture of tangent planes to a surface in $\mathbb{R}^3$ where the planes may intersect: at this picture, the planes are not the actual tangent spaces of the corresponding manifold $M = M^2$ (which are disjoint by the definition), but their images in the ambient $\mathbb{R}^3$ under the maps $i_x: v \mapsto x + v$ where both $x \in M$ and $v \in T_x M$ are regarded as elements of $\mathbb{R}^3$.

There is a natural map $p: TM \to M$ that send each vector $v \in T_x M$ to the point $x$ to which it is attached. It is called the projection on $TM$.

The set $TM$ has a natural structure of a manifold of dimension $2n$, induced by the manifold structure of $M^n$. Suppose $\mathcal{A} = \{\varphi_\alpha: U_\alpha \to V \subset M\}$ is at atlas for $M$. We define the corresponding atlas for $TM$ as follows: each chart $\varphi_\alpha: U_\alpha \to V \subset M$ for $M$ gives rise to a chart

$$
\tilde{\varphi}_\alpha: U_\alpha \times \mathbb{R}^n \to TV_\alpha \subset TM,
$$

where $v$ is the tangent vector at $x = \varphi_\alpha(x^1, \ldots, x^n)$ with the components $(v^1, \ldots, v^n)$ in the chart $\varphi_\alpha$. The inverse map $\tilde{\varphi}_\alpha^{-1}: TV_\alpha \to U_\alpha \times \mathbb{R}^n$ sends a tangent vector $v \in T_x M$, where $x \in V_\alpha$, to $(x^1, \ldots, x^n, v^1, \ldots, v^n) \in U_\alpha \times \mathbb{R}^n$ where $(x^1, \ldots, x^n) = \varphi_\alpha^{-1}(x)$ and $(v^1, \ldots, v^n)$ are the components of $v$ in the chart $\varphi_\alpha$. The changes of coordinates $\tilde{\varphi}_\alpha^{-1} \circ \tilde{\varphi}_\beta: U_{\beta\alpha} \times \mathbb{R}^n \to U_{\alpha\beta} \times \mathbb{R}^n$ consist of the changes of coordinates on $M$ and the corresponding transformations of the components of vectors:

$$
x^i = x^i(x'^1, \ldots, x'^n),
$$

$$
v^i = \sum_{i'} \frac{\partial x^i}{\partial x'^{i'}}(x'^1, \ldots, x'^n) v^{i'},
$$

where we denoted by $x^i, v^i$ coordinates in the chart $\tilde{\varphi}_\alpha$ and by $x'^i, v'^i$ coordinates in the chart $\tilde{\varphi}_\beta$.

The projection map $p: TM \to M$ in coordinates has the form

$$
(x^1, \ldots, x^n, v^1, \ldots, v^n) \mapsto (x^1, \ldots, x^n),
$$

7
i.e., the standard projection $U_{\alpha} \times \mathbb{R}^n \to U_{\alpha}$. We see that the tangent bundle locally looks like a direct product (but it is not a product globally.)

**Remark 4.3.** A manifold $E = E^{n+k}$ that can be presented as the union of a family of vector spaces (of the same dimension $k$)

$$E = \bigcup_{x \in M} E_x,$$

parametrized by the points of another manifold $M = M^n$, and endowed with an atlas where the projection $E \to M$ has the form $(x^i, v^\mu) \mapsto (x^i)$ (where $x^i$ are coordinates on $M$ and $v^\mu$ coordinates on $E_x$) and the changes of coordinates have the form

$$x^i = x^i(x^{1'}, \ldots, x^{n'}),$$

$$v^\mu = \sum_{\mu'} T_{\mu\mu'}^{\mu}(x^{1'}, \ldots, x^{n'}) v^{\mu'},$$

is called a *vector bundle* with base $M$. Hence $TM$ is an example of a vector bundle (here $k = n$). The notion of vector bundles and of more general *fibre bundles* is fundamental in differential geometry. Fibre bundles are informally referred to as “twisted products” because a fibre bundle with base $M$ locally looks like the product of an open subset $U \subset M$ with some fixed manifold $F$ (called the “standard fibre”). For vector bundles such as the tangent bundle the standard fibre is the vector space $\mathbb{R}^k$ or $\mathbb{C}^k$.

### 4.2.2 Tangent map and its matrix

Consider a smooth map $F: M \to N$. Fix a point $x \in M$ and denote $y = F(x) \in N$. We shall define a linear map from the vector space $T_x M$ to the vector space $T_y N$ ‘induced’ by the map $F$.

Each tangent vector $v \in T_x M$ can be interpreted as the velocity vector of some curve $\gamma: (a, b) \to M$ through $x$: $v = \dot{x} = \frac{dx}{dt}$.

For concreteness let us assume that $(a, b) = (-\varepsilon, \varepsilon)$ and $x = x(0)$; the time derivative above is therefore also taken at $t = 0$.

Consider the composition $F \circ \gamma$; it is a smooth curve in $N$.

**Definition 4.4.** The *tangent map* for the map $F$ at $x \in M$ maps a vector $v = \dot{x} \in T_x M$ to the velocity vector of the curve $F \circ \gamma$ in $N$ at $t = 0$:

$$\dot{x} \mapsto \frac{d}{dt} F(x(t)) \quad (4)$$

(the derivative is taken at $t = 0$). Notation for the tangent map: $dF(x)$ or $TF(x)$ or $DF(x)$ or $F_\ast(x)$. (The indication to a point $x$ is often dropped if it is clear from the context.)

The tangent map for a smooth map $F$ is also referred to as the *differential* or *derivative* of $F$ (at a point $x \in M$).
Remark 4.4. In elementary differential calculus we are used to distinguish derivatives and differentials (say, of a function of one variable). On a more advanced level such a distinction disappears. We shall see it in examples.

Theorem 4.2. The tangent map is a linear transformation $T_x M \to T_{F(x)} N$.

Proof. Suppose $v \in T_x M$ is the velocity of curve $\gamma: (-\varepsilon, \varepsilon) \to M$ at $t = 0$. Let us calculate the action of $dF(x)$ on $v$ using coordinates. Fix coordinate systems on $M$ and $N$ so that coordinates on $M$ are denoted as $x^i$, $i = 1, \ldots, n$, and on $N$, as $y^a$, $a = 1, \ldots, m$. Let the curve $\gamma$ be represented in coordinates as $x^i = x^i(t), i = 1, \ldots, n$, and the map $F$, as $y^a = y^a(x^1, \ldots, x^n)$. We have $v^i = \dot{x}^i(0)$. The image of $v$ under $dF(x)$ is the vector $d\frac{dt}{d} (F(x(t)))$, (the derivative at $t = 0$), hence in coordinates it will be

$$ (dF(x)(v))^a = \frac{d}{dt} y^a(x^1(t), \ldots, x^n(t)) = \sum_{i=1}^{n} \frac{\partial y^a}{\partial x^i} \frac{dx^i}{dt} = \sum_{i=1}^{n} \frac{\partial y^a}{\partial x^i} v^i. $$

Hence the tangent map $dF$ at $x \in M$ is a linear map (as claimed), with the matrix

$$ \left( \frac{\partial y^a}{\partial x^i}(x^1, \ldots, x^n) \right) $$

w.r.t. the coordinate bases $e_i$ in $T_x M$ and $e_a$ in $T_{F(x)} N$.

4.2.3 Tangent map for a composition

Suppose we have two smooth maps, $G: M \to N$ and $F: N \to P$. Consider their composition $F \circ G: M \to P$. Let $x \in M$, $y = G(x) \in N$, and $z = F(y) \in P$. How to calculate the tangent map to the composition $F \circ G$ at $x$? It is supposed to be a linear transformation $T_x M \to T_z P$.

Theorem 4.3. The tangent map $d(F \circ G)$ for the composition $F \circ G$ at $x$ is the composition of the linear transformations $dF$ at $y = F(x)$ and $dG$ at $x$:

$$ d(F \circ G)(x) = dF(y) \circ dG(x). $$

Proof. One way of proving this is to write down both sides in coordinates. If $x^1, \ldots, x^n$ are coordinates on $M$, $y^1, \ldots, y^m$ are coordinates on $N$, and $z^1, \ldots, z^r$ are coordinates on $P$, and the maps $F$ and $G$ are specified, respectively, by the equations $z^\mu = z^\mu(y^1, \ldots, y^m)$ and $y^a = y^a(x^1, \ldots, x^n)$, then the LHS of (5) has the matrix

$$ \left( \frac{\partial z^\mu}{\partial x^i} \right) $$

and the matrix of the RHS is the product of matrices

$$ \left( \frac{\partial z^\mu}{\partial y^a} \right) \left( \frac{\partial y^a}{\partial x^i} \right) = \left( \sum_a \frac{\partial z^\mu}{\partial y^a} \frac{\partial y^a}{\partial x^i} \right). $$

9
which coincides with (6) by the chain rule known from multivariate calculus. In other words, the composition formula (5) is but an abstract version of the chain rule.

It is possible to come to (5) more geometrically, working directly from the definition. Suppose \( v \in T_xM \) is the velocity of a curve \( x(t) \) at \( t = 0 \). Then the image of \( v \) under \( d(F \circ G)(x) \) is

\[
\frac{d}{dt} F(G(x(t)))
\]

(at \( t = 0 \)). On the other hand, the image of \( dG(x)(v) \) under \( dF(y) \), i.e., the RHS of (5) applied to \( v \), is

\[
\frac{d}{dt} F(y(t))
\]

for an arbitrary curve \( y(t) \) on \( N \) such that its velocity at \( t = 0 \) is \( dG(x)(v) \). We can take \( y(t) := G(x(t)) \) as such a curve, and plugging it into (9), we arrive at (8). Hence the LHS and the RHS of (5) coincide. \( \square \)