3 The tangent bundle

3.1 Introduction

This is a good time to reflect why we want a notion of tangent spaces and tangent maps in the first place. What are such objects supposed to deliver? What properties should tangent vectors and tangent spaces have? What are the tangent spaces to the line \( \mathbb{R} \) and the plane \( \mathbb{R}^2 \) – two of the most familiar manifolds?

We want to use our experience with tangent lines to curves and tangent planes to surfaces in two- and three dimensional Euclidean spaces as guidance. However, in general we do not want our notion of tangent objects to depend on, or be constrained by imbeddings of the manifold into some Euclidean space. Thus without any surrounding space available, the pictorial arrows become untenable. Before reading on, you should close the notes and brainstorm some ideas . . .

Some ideas which come to mind are:

- Tangent objects should implement linear approximations of objects on manifolds and of maps between manifolds, where infinitesimally linear is synonymous with differentiable.
- The dimension of the linear tangent space(s) should equal the dimension of the manifold.
- Vector fields are intimately connected to differential equations / dynamical systems. Thus tangent vectors should provide a means to describe dynamical systems on manifolds, as well as, more generally, partial differential equations on manifolds. This includes the special case of gradient vector fields, which are the derivatives of some potential function.
- We defined arc-length as an integral of the speed, rather then as the supremum of the length of polygonal approximations. In general tangent vectors may provide a means on which to base a generalized notion of distance.
- We defined curvature for curves in \( \mathbb{R}^n \) in terms of the rate of change of the tangent vectors. Thus we expect that a general notion of (comparing) tangent spaces (at different points) should provide for a notion of curvature.

Should our definition allow tangent spaces at different points of a manifold to have nonempty intersection? E.g. consider the unit-circle \( S^1 \) imbedded in the plane \( \mathbb{R}^2 \). The tangent lines to \( S^1 \) at \( p = (1, 0) \) and at \( q = (0, 1) \) intersect nontrivially at \( (1, 1) \). On the other hand, if we think of the tangent vectors \( v_p = (0, 1) \) and \( v_q = (1, 0) \) as arrows based at \( p \) and \( q \) respectively, then we certainly think of them as different.

This brings up a larger issue of distinguishing vectors (arrows) that may be moved around and vectors that are rooted at a fixed point. There are many applications where it is advantageous to consider equivalence classes of directed line-segments, equivalence meaning that they may be transformed into each other by parallel translation. Compare the first definition as objects characterized by their direction and magnitude, but which have no fixed root. On the other
hand, there are many places where it is appropriate to consider vectors that are rooted, or fixed at their base points, velocity vectors to a curve, and more generally vector fields.

The next section shall be the first effort to bring clarity to these issues and make very precise definitions (which may always be relaxed where this causes no trouble).

### 3.2 Tangent spaces

There are many different ways in which one may motivate an eventual construction of tangent spaces to a general manifold. One typically starts from surfaces in Euclidean spaces, then considers more abstractly immersed manifolds in higher dimensional Euclidean spaces, and eventually tries to develop a notion that works in abstract settings, yet reduces to the familiar ones in Euclidean settings. For a lengthy such discussion see Spivak Vol. I ch. 3.

An intuitive (and very useful) way to define tangent vectors to a manifold $M$ at a point $p$ is as equivalence classes of curves. Roughly, two curves are equivalent if they have the same velocity vector at $p$ — but this would be circular as we don’t have notions of velocity vectors for general curves on manifolds. So the next best thing is to declare any two smooth curves $\sigma, \gamma : (-\varepsilon, \varepsilon) \mapsto M$ (with $\sigma(0) = \gamma(0) = p$) equivalent if for every smooth function $f \in C^\infty(p)$ (that is, smooth function defined on some neighborhood of $p$),

$$\frac{d}{dt}|_{t=0}(f \circ \sigma) = \frac{d}{dt}|_{t=0}(f \circ \gamma).$$

**Project 3.1** Further explore how this leads to a notion of tangent spaces that is basically the same as the one we define below. In particular, equip the collection of equivalence classes with an addition and scalar multiplication (make sure that these are well-defined). Check that the space of tangent vectors at a point is indeed an $m$-dimensional vector space. In a coordinate chart find a basis for the tangent space (e.g. provide representatives (curves) for $m$ equivalence classes that form a basis). Show how to write any tangent vector, that is any equivalence class of curves, as a linear combination of this basis. Analyze how the coordinates of a tangent vector transform under local coordinate changes on the manifold. Match this notion of tangent spaces to the one provided below.

The exercise already hinted at a useful connection between tangent vectors and derivatives. Indeed, going back to Euclidean spaces, say e.g. $M = \mathbb{R}^2$ consider the relation between a vector $\vec{v} = (v^1, v^2) = v^1 \vec{i} + v^2 \vec{j}$ and the directional derivative (operator) $D_{\vec{v}}|_p$, defined by $D_{\vec{v}}f = v^1 \cdot (D_1 f)(p) + v^2 \cdot (D_2 f)(p)$, commonly also written as $D_{\vec{v}}f(p) = \langle \vec{v}, \nabla f(p) \rangle$.

Clearly, any vector $\vec{v}$ uniquely determines a (directional) derivative operator $D_{\vec{v}}|_p$. Conversely, one can easily recover the vector $\vec{v}$ from the directional derivative $D_\vec{v}$ by simply evaluating the latter on suitable functions: For example, evaluating $D_{\vec{v}}|_p$ on the coordinate functions $\pi^1, \pi^2 : \mathbb{R}^2 \hookrightarrow \mathbb{R}$ defined by $\pi^1(x^1, x^2) = x^1$ and $\pi^2(x^1, x^2) = x^2$ immediately recovers the coordinates of $\vec{v}$ as $v^1 = (D_{\vec{v}}\pi^1)(p)$ and $v^2 = (D_{\vec{v}}\pi^2)(p)$. Thus there appears to be no harm in identifying the vector $\vec{v}$ with the directional derivative operator $D_{\vec{v}}(\cdot)(p)$. This idea will prove most beneficial since operators on spaces of functions are automatically endowed with a rich algebraic structure that is ready to use! Following [Chevalley, 1946] define

**Definition 3.1** A tangent vector to a manifold $M$ at a point $p \in M$ is a function $X_p : C^\infty(p) \mapsto \mathbb{R}$ which is linear over $\mathbb{R}$ and is a derivation on $C^\infty(p)$, i.e. which satisfies for $\lambda \in \mathbb{R}$ and for all $f, g \in C^\infty(p)$ on their common domain,

- $X_p(\lambda f + g) = \lambda X_p(f) + X_p(g)$.
- $X_p(f \cdot g) = X_p(f) \cdot g(p) + f(p) \cdot X_p(g)$

The set of all tangent vectors to $M$ at $p$ is called the tangent space to $M$ at $p$, denoted $T_p M$. 
This idea of identifying objects such as tangent vectors as operators on the algebra $C^\infty(M)$ of smooth functions on $M$ is part of a much larger theme. For example, analogous to the above one might consider the multiplicative linear functionals $\hat{p}: C^\infty(M)$, i.e. which satisfy for all $f, g \in C^\infty(M)$ and $c \in \mathbb{R}$ $\hat{p}(f + cg) = \hat{p}(f) + c\hat{p}(g)$ and $\hat{p}(fg) = \hat{p}(f)\hat{p}(g)$. It is easily seen, that each point $p \in M$ gives rise to such a functional via $\hat{p}(f) = f(p)$. More generally, one may start with an algebra $\mathcal{A}$ (with suitable topology or norm), and as whether there exists a manifold $M$ such that $\mathcal{A}$ is isomorphic to $C^\infty(M)$. A critical observation is that functionals of the above form that correspond to evaluations at a point are associated to maximal ideals – think of the ideal of all functions $f \in C^\infty(M)$ which are zero at a fixed point. Indeed, under suitable technical hypotheses it is indeed possible to recover the manifold from the algebra suitably topologizing the space of maximal ideals. But this is just the beginning of a long story. See e.g. the introduction of the textbook *Control Theory from a Geometric Point of View* by Agrachev and Sachkov, for a very quick introduction to these tools. Historically, the starting point for this approach is usually taken to be I. Gelfand’s dissertation in 1935, as well as related work by Marshall Stone and others that started in the late 1920s. The differential geometry part of this course will only use this point of view as related work by Marshall Stone and others that started in the late 1920s.

Recall, $f \in C^\infty(p)$ means that there exists an open neighborhood $U$ of $p$ (depending on $f$) such that $f \in C^\infty(U, \mathbb{R})$. (This is a special case of the earlier definition of $C^\infty(A)$ for subsets $A \subseteq \mathbb{R}^n$ that are not necessarily open.) The following lemma only uses the abstract notion of a derivation, no coordinates are required:

**Lemma 3.1** If $c: M \rightarrow \mathbb{R}$ is a constant function and $X_p \in T_pM$ then $(X_pc) = 0$

**Proof.** Use the linearity and Leibniz rule for differentiating products to establish

$$c \cdot X_p(1) = X_p(c \cdot 1) = (X_p c) \cdot 1 + c \cdot (X_p 1) \quad \text{and hence} \quad (X_pc) = 0.$$

**Observations:**

- $T_pM$ is a vector space: In particular, if $X_p, Y_p \in T_pM$ and $\lambda \in \mathbb{R}$ then $(X_p + \lambda Y_p) \in T_pM$. The addition and scalar multiplication are inherited from pointwise evaluations, i.e. $(X_p + \lambda Y_p)f = (X_p f) + \lambda (Y_p f)$.

- Tangent vectors are local operators: If $f, g \in C^\infty(M)$ agree on some neighborhood $U$ of $p$, i.e. $f|_U \equiv g|_U$ then for all $X_p \in T_pM$, $X_pf = X_pg$.

Technically, the “natural domain” of tangent vectors $X_p \in T_pM$ are germs of functions at $p$: The germ of a function $f \in C^\infty(p)$ is defined as the set of all $g \in C^\infty(p)$ for which there exists an open neighborhood $U$ of $p$ so that $f|_U \equiv g|_U$. 

**Exercise (Spivak I 3.13):** Use $\mathcal{M}_p = \{f \in C^\infty(M): f(p) = 0\}$ and show that $\left(\mathcal{M}_p/\mathcal{M}_p^2\right)^*$ is a model for $T_pM$ (where $\mathcal{M}_p^2 = \{fg: f, g \in \mathcal{M}_p\}$).
• If $M = \mathbb{R}^m$ then the tangent vectors to $M$ at a point $p$ are precisely the directional derivatives evaluated at $p$. (One direction is obvious. For the other see the calculation below for a general manifold.)

• If $(u, U)$ is a chart about $p \in M$, then for $j = 1, \ldots, m$, $i \frac{\partial}{\partial u^i} |_p \in T_p M$. Thus also $
abla a^j \frac{\partial}{\partial u^j} |_p \in T_p M$ for all $a^j \in \mathbb{R}$. To verify this assertion, recall that for $f \in C^\infty(p)$

\[
\frac{\partial f}{\partial u^i} |_p = D_i(f \circ u^{-1})|_{u(p)}
\]

and use the familiar properties of partial derivatives in $\mathbb{R}^n$ to manipulate e.g.

\[
\frac{\partial (f \circ g)}{\partial u^i} |_p = D_i((f \circ g) \circ u^{-1})|_{u(p)} = D_i((f \circ u^{-1}) \cdot (g \circ u^{-1}))|_{u(p)} = D_i(f \circ u^{-1})|_{u(p)} \cdot (g \circ u^{-1})(u(p)) + (f \circ u^{-1})(u(p)) \cdot D_i(g \circ u^{-1})|_{u(p)} = \frac{\partial f}{\partial u^i} |_p \cdot g(p) + f(p) \cdot \frac{\partial g}{\partial u^i} |_p
\]

More interesting is the converse, i.e. that in any chart $(u, U)$ about $p \in M$ every tangent vector $X_p \in T_p M$ may be expressed as a linear combination of the partial derivatives $\frac{\partial}{\partial u^i}$, $i = 1, \ldots, m$:

**Theorem 3.2** If $(u, U)$ is a chart about $p \in M^m$ then \{ $\frac{\partial}{\partial u^1} |_p, \ldots, \frac{\partial}{\partial u^m} |_p$ \} is a basis for $T_p M$.

**Corollary 3.3** If $(u, U)$ is a chart about $p \in M^m$ and $X_p \in T_p M$ then $X_p = \sum_{i=1}^m (X_p u^i) \cdot \frac{\partial}{\partial u^i} |_p$.

**Corollary 3.4** If $(u, U)$ and $(v, V)$ are charts about $p \in M^m$ then $\frac{\partial}{\partial v^i} |_p = \sum_{i=j}^m \frac{\partial v^i}{\partial u^j} |_p \cdot \frac{\partial}{\partial u^i} |_p$.

Consider this last corollary as a statement about (linear) bases changes in the tangent space associated to (nonlinear) changes of local coordinates on the manifold. Before proving this theorem establish a lemma using an elegant construction

**Lemma 3.5** Let $(u, U)$ be a chart about $p \in M^m$ with $u(p) = x_0$ and $f \in C^\infty(p)$. Then there exist $f_i \in C^\infty(p)$ such that $f_i(p) = \frac{\partial f}{\partial u^i} |_p$ for $i = 1, \ldots, m$, and $f(q) = f(p) + \sum_{i=1}^m (u^i(q) - u^i(p)) f_i(q)$.

This is basically a first-order Taylor expansion with remainder term. Here we have equality (as opposed to an approximation) – but the functions $f_i$ are evaluated at variable points $q$ as opposed to fixed derivatives evaluated at the fixed point $p$.

**Proof** (of the lemma). Using the local coordinates we reduce the proof to Euclidean spaces: Write $x = u(q)$ and $x_0 = u(p)$, rewrite the second statement of the lemma as

\[
f(u^{-1}(x)) = f(u^{-1}(x_0)) + \sum_{j=1}^m (u^j(u^{-1}(x)) - u^j(u^{-1}(x_0))) f_j(u^{-1}(x))
\]
Write \( g \) for \( f \circ u^{-1} : u(U) \subseteq \mathbb{R}^m \rightarrow \mathbb{R} \). The desired functions \( g_i : u(U) \subseteq \mathbb{R}^m \rightarrow \mathbb{R} \) are such that

\[
g(x) = g(x_0) + \sum_{j=1}^{m} (x^j - x_0^j)g_j(x).
\]

After shrinking the neighborhood \( U \), if necessary, we may assume that \( u(U) \subseteq \mathbb{R}^m \) is star-shaped with respect to \( x_0 \), i.e. for every \( x \in u(U) \) the line segment \( \{ x_0 + t \cdot (x - x_0) : t \in [0,1] \} \) is contained in \( u(U) \). For any fixed \( x \in u(U) \) consider the curve \( \sigma_x : [0,1] \rightarrow u(U) \) defined by \( \sigma(t) = x_0 + t \cdot (x - x_0) \). Via the fundamental theorem of calculus and the chain rule

\[
g(x) = g(\sigma_x(1)) = g(\sigma_x(0)) + \int_0^1 \frac{\partial}{\partial t} g(\sigma_x(t)) \, dt
\]

\[
= g(\sigma_x(0)) + \int_0^1 \sum_{j=1}^{m} (D_j g)(\sigma_x(t)) \cdot \frac{\partial x^j}{\partial t}(t) \, dt
\]

\[
= g(\sigma_x(0)) + \sum_{j=1}^{m} (x^j - x_0^j) \int_0^1 (D_j g)(\sigma_x(t)) \, dt
\]

Note the constant derivative \( \sigma_x' = (x - x_0) \) for the curve in \( \mathbb{R}^m \) – this makes no sense on a general manifold, but is coordinate dependent. One immediately verifies that with this definition \( g_j(x_0) = (D_j g)(x_0) \) (in this case \( \sigma_x_0 \) is a constant curve). Consequently \( f_j(p) = g_j(u(p)) = (D_j g)(u(p)) = D_j(f \circ u^{-1})|_{u(p)} = \frac{\partial f}{\partial u^j}
\]

\[
\text{p}. \quad \text{Since } g \in C^\infty, \text{ also } g_j \in C^\infty \text{ and } f_j \in C^\infty. \]

**Proof** (of the theorem). Suppose \((u, U)\) is a chart about \( p \in M^m \), \( X_p \in T_p M \) and \( f \in C^\infty(M) \). Using lemma 3.5 there exist (on some open neighborhood of \( p \)) suitable functions \( f_i \) such that we may rewrite \( X_p f \) as \( X_p f = X_p \left( f(p) + \sum_{j=1}^{m} (u^j - u^j(p)) f_j \right) \). Using the linearity of \( X_p \), the Leibniz rule, and that \( X_p c = 0 \) for any constant function this yields:

\[
X_p f = 0 + \sum_{j=1}^{m} \left( (X_p u^j) - 0 \right) f_j(p) + \left( (u^j(p) - u^j(p)) \cdot (X_p f) \right)
\]

\[
i.e. \quad (X_p f) = \sum_{j=1}^{m} (X_p u^j) \cdot f_j(p).
\]

By lemma 3.5 this is equal to \( (X_p f) = \sum_{j=1}^{m} (X_p u^j) \frac{\partial f}{\partial u^j}
\]

\[
\text{p}. \quad \text{Since this holds for all } f \in C^\infty(p), \text{ we conclude } X_p = \sum_{j=1}^{m} (X_p u^j) \frac{\partial}{\partial u^j} \bigg|_p. \quad \Box
\]

### 3.3 Tangent maps, part I

For every smooth map \( F : \mathbb{R}^n \rightarrow \mathbb{R}^m \) between Euclidean spaces and any point \( x \in \mathbb{R}^n \) the derivative of \( F \) at \( x \) is a linear map \( (DF)(x)(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^m \). Now that we have tangent spaces to manifolds, we are ready to associate analogous (linear) tangent maps (between tangent spaces) to smooth maps (between manifolds).

**Definition 3.2** Suppose \( \Phi : M \rightarrow N \) is a smooth map between manifolds and \( p \in M \). The tangent map \( \Phi_{*p} : T_p M \rightarrow T_{\Phi(p)} N \) (of \( \Phi \) at \( p \)) is defined for \( X_p \in T_p M \) and \( f \in C^\infty(\Phi(p)) \) by

\[
(\Phi_{*p} X_p) f = X_p(f \circ \Phi).
\]

(39)
How else could $\Phi_{sp}$ be defined? It is immediate that if $\Phi = \text{id}_M$ then $\Phi_{sp} = \text{id}_{T_p M}$. Also, it follows immediately from the definition that

**Proposition 3.6** Suppose $\Phi : M^m \to N^n$ and $\Psi : N \to P$ are smooth maps between manifolds, and $p \in M$. Then (note the preservation of the order of $\Phi$ and $\Psi$)

$$(\Psi \circ \Phi)_{sp} = \Psi_{*\Phi(p)} \circ \Phi_{sp}.$$ (40)

**Exercise 3.2** Prove proposition 3.6.

In local coordinates the tangent map is given by matrix multiplication. More specifically, suppose $(u, U)$ and $(v, V)$ are charts about $p \in M^m$ and $\Phi(p) \in N^n$, respectively, and $f \in C^\infty(\Phi(p))$. From the definitions calculate (using the chain-rule)

$$
\left( \Phi_{sp} \frac{\partial}{\partial \nu^i} \right)_p f = \frac{\partial}{\partial \nu^i} f \circ \Phi = D_i (f \circ \Phi \circ u^{-1})|_{u(p)} = D_i ((f \circ v^{-1}) \circ (v \circ \Phi \circ u^{-1}))|_{u(p)} = \sum_{i=1}^n D_i (f \circ v^{-1})|_{v(p)} \cdot D_j (v \circ \Phi \circ u^{-1})|_{u(p)} = \left( \sum_{i=1}^n \frac{\partial (v^i \circ \Phi)}{\partial \nu^i} \right)_p \cdot \frac{\partial}{\partial \nu^j} \Phi.$$

In concrete examples, using local coordinates, it is convenient to express tangent vectors as column vectors. E.g. suppose, as before, $\Phi \in C^\infty(M^m, N^n)$, and $(u, U)$ and $(v, V)$ are charts about $p \in M^m$ and $\Phi(p) \in N^n$. If $X_p \in T_p M$ is a tangent vector at $p$, let $a = (a^1, \ldots, a^m)^T$ be the column vector with components $a_i = X_p \nu^i$, representing $X_p = \sum_{i=1}^m a^i \frac{\partial}{\partial \nu^i} |_p$. Similarly, let $b = (b^1, \ldots, b^n)^T$ be the column vector with components $b^j = (\Phi_{sp} X_p) v^j$, representing the image $\Phi_{sp} X_p \in T_{\Phi(p)} N$. These column vectors $a$ and $b$ are related by matrix multiplication $b = Ca$ where $C$ is the $(n \times m)$ matrix with components $C_{ij} = \frac{\partial (v^i \circ \Phi)}{\partial \nu^j} |_p$.

Formally, one may go further, and write $\alpha$ for the row-vector with components $\alpha_i = \Phi_{sp} \frac{\partial}{\partial \nu^i} |_p$ and $\beta$ for the row-vector with components $\beta_j = \frac{\partial}{\partial \nu^j} \Phi(p)$. Then formally, the images of the basis vectors $\frac{\partial}{\partial \nu^i} |_p$ are obtained by right matrix multiplication, i.e. $\alpha = \beta \cdot C$. This matches with the observation that formally $\Phi_{sp} (X_p) = a a = (\beta C) a = \beta(Ca) = \beta b$. While this is all simple (formal) matrix algebra, it is worthwhile to remember that when transforming formal vectors of basis elements these are multiplied by the transformation matrix in a way opposite to the multiplication familiar for transforming specific vectors.

**Exercise 3.3** Suppose $\Phi \in C^\infty(M^m, N^n)$ and $\Psi \in C^\infty(N^n, P^r)$ are smooth maps between manifolds, $p \in M$ and $f \in C^\infty(P)$.

Furthermore, suppose $(u, U)$, $(v, V)$ and $(w, W)$ are local coordinate charts about $p \in M$, $\Phi(p) \in N$ and $(\Psi \circ \Phi)(p) \in P$, respectively. Verify that the matrix representing $(\Psi \circ \Phi)_{sp}$ with respect to $(u, U)$ and $(w, W)$ is the (matrix-)product of the matrices representing $\Phi_{sp}$ (with respect to $(u, U)$ and $(v, V)$) and $\Psi_{*\Phi(p)}$ (with respect to $(v, V)$ and $(w, W)$).

We digress a little to consider tangent spaces of immersed manifolds in $\mathbb{R}^n$ which justify the familiar pictures of tangent planes. Suppose that $\Phi \in C^\infty(M^m, \mathbb{R}^n)$ is an immersion at $p \in M$,
i.e. \( \text{rank}_p \Phi = m \). Using local coordinates \((u,U)\) about \( p \in M^m \) and the standard coordinates \((x,\mathbb{R}^n)\) in the codomain, the rank condition says that the \((n \times m)\)-matrix with components \( \frac{\partial (x^i \circ \Phi)}{\partial u^j} \bigg|_p \) has rank \( m \), and \( \Phi_{*p} \) is a monomorphism (a linear one-to-one map) from \( T_p M \) to \( T_{\Phi(p)} \mathbb{R}^n \).

The tangent vectors \( \Phi_{*p} \left( \frac{\partial}{\partial u} \bigg|_p \right) \in \Phi_{*p} (T_p M) \subseteq T_{\Phi(p)} \mathbb{R}^n \) are linearly independent, and span an \( m \)-dimensional subspace of \( T_{\Phi(p)} \mathbb{R}^n \) which is usually pictured as a tangent line/plane.

The image of any tangent vector \( X_p \in T_p M \) in the standard coordinates may be written in the form \( \Phi_{*p} X_p = \sum_{i=1}^n b_i D_i|_{\Phi(p)} \). Now, if \( f \in C^\infty(\mathbb{R}^n) \) is a function such that \( f \circ \Phi \equiv 0 \) in a neighborhood of \( p \in M \), then

\[
0 \equiv X_p (f \circ \Phi) = (\Phi_{*p} X_p) f = \sum_{i=1}^n b_i (D_i f)|_{\Phi(p)}
\]

which in calculus notation might be written as \( \Phi_{*p} X_p \perp (\nabla f)(\Phi(p)) \), or \( 0 = (b, (\nabla f)(\Phi(p))) \). Recall if \((\nabla f)(\Phi(p)) \neq 0\), then \( f^{-1}(0) \) is (locally) a smooth hypersurface in \( \mathbb{R}^n \)(near \( \Phi(p) \)). Thus \( (\Phi_{*p} X_p) \) may be pictured as lying in the tangent hyperplane to the hypersurface \( f^{-1}(0) \) at \( \Phi(p) \).

**Exercise 3.4** Generalize this discussion to the case when there are functions \( f^1, \ldots, f^{n-m} \in C^\infty(\Phi(p)) \) with \( f^i \cdot \Phi \equiv 0 \) and linearly independent gradients \( (\nabla f^i)(\Phi(p)) \).

**Example 3.1** As a hands-on example consider an immersion of the M-strip into \( \mathbb{R}^3 \). One way to represent the M-strip is as the quotient \( M = \mathbb{R} / \sim \) of the rectangle \( R = [0, 2\pi] \times (-1, 1) \) two of whose edges have been identified by \((0, t) \sim (2\pi, -t)\). Define a map \( \Phi: M \mapsto \mathbb{R}^3 \) by \( \Phi(\theta, t) = \left((2 + t\cos(\frac{1}{2}\theta)) \cos \theta, (2 + t\cos(\frac{1}{2}\theta)) \sin \theta, t \sin(\frac{1}{2}\theta)\right) \).

**Exercise 3.5** Explicitly calculate the images \( \Phi_{*p} \left( \frac{\partial}{\partial \theta} \bigg|_p \right) \) and \( \Phi_{*p} \left( \frac{\partial}{\partial t} \bigg|_p \right) \) at any point \( p = (\theta, t) \in U = (0, 2\pi) \times (-1, 1) \subseteq M \). Verify that \( \Phi \) is indeed an immersion.

The image \( \Phi(M) \subseteq \mathbb{R}^3 \) is a surface in the usual sense. The images of the tangent vectors to \( M \) calculated in the exercise may be visualized as the usual arrows that are tangent to the surface. It is easily seen that the map \( \Phi \) is indeed well-defined on \( M \) (as opposed to only on the rectangle \( R \)) because \( \Phi(0, t) = \Phi(2\pi, -t) \). For \( p \in U \) the map \( \Phi_{*p} \) is well-defined, but problems arise when trying to extend \( \Phi \) and \( \Phi_{*p} \) continuously to all \( q \in M \). Indeed, \(((\theta, t), U)\) is a local coordinate chart of \( M \), but it does not cover all of \( M \). In the language of the next sections \( \Phi_{*} \) maps the coordinate vector fields \( \frac{\partial}{\partial \theta} = \frac{\partial}{\partial \theta} \) and \( \frac{\partial}{\partial t} = \frac{\partial}{\partial t} \) (which are only defined on \( U \)) to vector fields on \( \Phi(U) \), but the vector field \( \frac{\partial}{\partial \theta} \) cannot be continuously extended to a vector field on all of \( M \).

We conclude this first section on tangent maps with a generalization of the earlier local submersion theorem.

**Theorem 3.7** Suppose that \( \Phi \in C^\infty(M^m, N^n) \) is a smooth map between manifolds and \( P^r \subseteq N \) is a smooth submanifold. If for every \( p \in \Phi^{-1}(P) \)

\[
\Phi_{*p}(T_p M) + T_{\Phi(p)} P = T_{\Phi(p)} N
\]

then \( \Phi^{-1}(P) \subseteq M \) is a submanifold of \( M \) of dimension \((m - (n - r))\).
In general one calls a smooth map $\Phi \in C^\infty(M, N)$ between manifolds transverse to (a submanifold) $P \subseteq N$ along (a submanifold) $L \subseteq M$ if $\Phi_p(T_pM) + T_{\Phi(p)}P = T_{\Phi(p)}N$ for all $p \in L \cap \Phi^{-1}(P)$.

The theorem motivates the notion of codimensions as opposed to dimensions of submanifolds. More specifically, for an $r$-dimensional submanifold $P^r \subseteq N^n$ of an $n$-dimensional manifold $N$ the codimension of $P$ (in $N$) is defined as $(n-r)$. The theorem then simply states that if $P \subseteq N$ is a submanifold of codimension $k$ and $\Phi$ is transversal to $P$ (along $M$) then $\Phi^{-1}(P) \subseteq M$ is a submanifold of the same codimension $k$.

**Proof.** Suppose that $\Phi \in C^\infty(M^m, N^n)$ and $p \in \Phi^{-1}(P) \subseteq M$ is in the preimage of a smooth submanifold $P^r \subseteq N$. Using theorem 2.13 choose an adapted chart $(v, V)$ about $\Phi(p) \in N$ such that $v(\Phi(p)) = 0$ and such that the restriction of $w = (v^1, \ldots v^r)$ to the set $W = \{q \in N : v^{r+1}(q) = \ldots = v^n(q) = 0\}$ is a chart of $P$ about $\Phi(p)$.

Define $\Psi: V \mapsto \mathbb{R}^{n-r}$ by $\Psi = (v^{r+1}, \ldots v^n)$. Let $U = \Phi^{-1}(V) \subseteq M$.

Then $p \in \Phi^{-1} \circ \Psi^{-1}(0) = (\Psi \circ \Phi)^{-1}(0)$ and $(\Psi \circ \Phi)_p = \Psi_{*\Phi(p)} \circ \Phi_p: T_pM \mapsto T_0\mathbb{R}^{n-r}$.

Use that $D(\Psi \circ v^{-1}) = (0, I_{n-r})$ and that the kernel of $\Psi_{*\Phi(p)}$ is precisely $T_{\Phi(p)}$. Together with $\Phi_p(T_pM) + T_{\Phi(p)}P = T_{\Phi(p)}N$ this establishes that the restriction of $\Psi_{*\Phi(p)}$ to the image of $\Phi_p$ (i.e. to $\Phi_p(T_pM)$) has full rank, and hence rank$(\psi \circ \Phi)_p = n-p$ (using corollary 2.14).

**Exercise 3.6** Revisit the Hopf map $\Phi: S^3 \mapsto S^2$ (of exercise 2.31), i.e., the restriction (to $S^3$) of $\tilde{\Phi}: \mathbb{R}^4 \mapsto \mathbb{R}^3$ defined by $\tilde{\Phi}(a) = (2a_1a_3 + 2a_2a_4, 2a_2a_3 - 2a_1a_4, a_1^2 + a_2^2 - a_3^2 - a_4^2)$.

Considering the usual imbeddings of the spheres into $\mathbb{R}^4$ and $\mathbb{R}^3$, respectively, describe the preimages $\Phi^{-1}(P_c)$ of the meridians $P_c = S^2 \cap \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = c\}$ for $-1 \leq c \leq 1$. 
3.4 The tangent bundle

It is natural to assemble all tangent spaces of a manifold together into a new object – and conceivably, this set should again have a natural manifold structure. We will omit some of the more technical details of this construction and refer to e.g. Spivak vol.1 ch.3, especially exercise 1. As discussed earlier, it is desirable to distinguish between tangent vectors at different points, and we define:

**Definition 3.3** The tangent bundle $TM$ of a manifold $M$ is (as a set) the (disjoint) union of all tangent spaces to $M$ at all points $p \in M$.

$$TM = \left\{ (p, X_p) \in M \times \bigcup_{p \in M} T_pM \mid X_p \in T_pM \right\}.$$ (43)

The bundle projection $\pi: TM \to M$ is defined by $\pi(p, X_p) = p$. The fiber over $p \in M$ is the preimage $\pi^{-1}(p) = \{p\} \times T_pM$. A section of $TM$, or tangent vector field, is a map $X: M \to TM$ that satisfies $\pi \circ X = \text{id}_M$.

Basically, vector fields are functions that assign to each $p \in M$ a tangent vector in $T_pM$. Often we conveniently identify the fiber $\pi^{-1}(p)$ with $T_pM$ and the pair $(p, X_p) \in T_M$ with tangent vector $X(p) = X_p \in T_pM$. Technically this involves a tacit projection onto the second factor or a tacit use of the inclusion map $i_p: T_pM \hookrightarrow \{p\} \times T_pM \subseteq TM$. However, in some instances, e.g. when working with $TTM$, more precision is indicated.

To illustrate that tangent bundles of manifolds are candidates to be considered manifolds themselves, consider the example of $M = S^1$. The *naive* collection of all tangent lines to the imbedded circle $S^1 \subseteq \mathbb{R}^2$ is full of intersections. More suitable for our purposes is to imbed the circle in $\mathbb{R}^3$ as $\tilde{S}^1 = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 = 1, \ x_3 = 0\}$ and attach at every point $p \in \tilde{S}^1$ a vertical (!) line, yielding a cylinder. As a set, this cylinder is in bijection with the (disjoint) collection of all tangent lines to the circle imbedded in the plane. Intuitively one can consistently choose an orientation of the lines, and even more a consistent scaling. For example, identify the naive tangent vector $\left((\cos \Theta, \sin \Theta), (-L \sin \Theta, L \cos \Theta)\right)$ with the point $(\cos \Theta, \sin \Theta, L) \in \mathbb{R}^3$.

Within this picture, a vector field on the circle may be visualized as the graph of a function $(\cos \Theta, \sin \Theta) \mapsto L(\theta)$. If the vector field is continuous and nonvanishing, then the graph lies entirely above, or entirely below the plane $x_3 = 0$.

In analogy, we may intuitively think of the tangent bundle $T\mathbb{R}$ of the real line $\mathbb{R}$ as the plane $\mathbb{R}^2$. However, due to dimensional reasons it is clear that these two examples are the only tangent bundles amenable to such immediate visualization. How quickly things get complicated becomes clear if one tries to think of $TS^2$ as a sphere with a plane attached to each of its points. A vector field on the sphere simply selects one point on each plane. However, from algebraic topology it is known that there does not exist any continuous (yet to be defined!) vector field on the sphere that vanishes nowhere. In our picture this means that it is impossible to continuously select one point on each tangent plane avoiding the origin (zero-vector) in each $T_pS^2$. Intuitively $TS^2$ must be a nontrivially twisted, (when compared to e.g. $TS^1$ which is the very tame cylinder), i.e. it must be very different from the trivial Cartesian product $S^2 \times \mathbb{R}^2$. 


We proceed more abstractly to endow the tangent bundle $TM$ of a smooth manifold $M$ with a manifold structure. The key idea is that locally, above a chart $(u,U)$ (which itself is homeomorphic to $\mathbb{R}^m$) the tangent bundle basically looks like $\mathbb{R}^m \times \mathbb{R}^m \cong \mathbb{R}^{2m}$. This observation is captured in the concept of local triviality, i.e., if $(u,U)$ is a chart of an $m$-dimensional manifold $M$, then $TU \cong U \times \mathbb{R}^m$. Thus the topology and geometry of $M$ are captured in the global structure of the tangent bundle, by how the trivial bundles are pieced together in some twisted way.

Starting with an atlas of local coordinate charts for $M$ we shall find it easy to obtain a candidate atlas of charts for $TM$. There is a natural candidate for a topology on $TM$ that makes all candidate charts for $TM$ into homeomorphisms, and automatically guarantees that their transition maps are smooth. However, it takes a little more advanced arguments to show that this topology is indeed sufficiently nice (metrizable or paracompact) so that $TM$ qualifies as a manifold.

Suppose $(u,U)$ is a chart of $M$ about $p$. Consider the subset $\bar{U} = \pi^{-1}(U) \subseteq TM$. Every point $Q \in \bar{U}$ is a pair $Q = (q,X_q)$ with $X_q \in T_qM$. Since $\{\frac{\partial}{\partial u^j} | q : 1 \leq j \leq m\}$ is a basis for $T_qM$ there exists functions $w^j : \bar{U} \mapsto \mathbb{R}$ (indeed, $w^j(Q) = X_q u^j$) such that

$$Q = \left( \pi(Q), \sum_{j=1}^{m} w^j(Q) \frac{\partial}{\partial u^j}|_{\pi(Q)} \right). \quad (44)$$

It is natural to define a map $\bar{u} : \bar{U} \mapsto \mathbb{R}^{2m}$ by

$$\bar{u} = \left( u^1 \circ \pi, \ldots, u^m \circ \pi, w^1, \ldots, w^m \right). \quad (45)$$

It is clear that $\bar{u}$ is a bijection from $\bar{U}$ to $\mathbb{R}^{2m}$. (This assumes that $u$ is a bijection from $U$ to $\mathbb{R}^m$, as originally mandated. Alternatively, $\bar{u}$ is a bijection from $\bar{U}$ onto $u(U) \times \mathbb{R}^m \subseteq \mathbb{R}^{2m}$.)

For each chart $(u,U)$ in the $C^\infty$-differentiable structure of $M$, we want the associated map $\bar{u}$ to be a homeomorphism. This is achieved by endowing $TM$ with the weakest (i.e., coarsest) topology in which all maps $\bar{u}$ are continuous. More constructively, we consider the collection $T$ of all subsets $O \subseteq TM$ which are such that for every point $Q \in TM$ and every chart $(u,U)$ of $M$ for which $\pi(Q) \in U$, there exists an open set $W \in \mathbb{R}^{2m}$ containing $\bar{u}(Q)$ such that $\bar{u}^{-1}(W) \subseteq O$.

Exercise 3.7 Show that the collection $T$ of subsets of $TM$ is a topology on $TM$, i.e., show that $\emptyset, TM \in T$, and that $T$ is closed under finite intersections and arbitrary unions.

Exercise 3.8 Check that when $TM$ is endowed with this topology $T$ then for every chart $(u,U)$ the associated map $\bar{u}$ is a homeomorphism, i.e. is continuous with continuous inverse.

Next calculate the transition maps $(\bar{v} \circ \bar{u}^{-1}) : \bar{u}(\bar{U} \cap \bar{V}) \subseteq \mathbb{R}^{2m} \mapsto \bar{v}(\bar{U} \cap \bar{V}) \subseteq \mathbb{R}^{2m}$ to verify that the charts are indeed $C^\infty$-related.

Thus assume that $Q \in \bar{U} \cap \bar{V} \subseteq TM$ and that $\bar{u}(Q) = (x,y) \in \mathbb{R}^m \times \mathbb{R}^m$. Calculate

$$\bar{v}(Q) = (\bar{v} \circ \bar{u}^{-1})(x,y) = \bar{v}(\bar{u}^{-1}(x), \sum_{j=1}^{m} y^j \frac{\partial}{\partial u^j}|_{\bar{v}^{-1}(x)}). \quad (46)$$
To obtain the second $m$-components of $\bar{v}(Q)$ change the basis in $T_{\pi(Q)}M$ from the \( \{ \frac{\partial}{\partial v^i}|_{\pi(Q)} \}^m_{j=1} \) to the \( \{ \frac{\partial}{\partial w^i}|_{\pi(Q)} \}^m_{j=1} \), yielding
\[
(\bar{v} \circ \bar{u}^{-1})(x, y) = \bar{v} \left( (u^{-1}(x), \sum_{j=1}^{m} y^j \sum_{i=1}^{m} \frac{\partial v^i}{\partial w^j}|_{u^{-1}(x)} \frac{\partial}{\partial v^i}|_{u^{-1}(x)} \right). \tag{47}
\]
Interchange the order of summation and regroup to read off the components of $\bar{v}(Q)$
\[
(\bar{v} \circ \bar{u}^{-1})(x, y) = ((v \circ u^{-1})(x), \sum_{j=1}^{m} y^j \frac{\partial v^1}{\partial w^j}|_{u^{-1}(x)} , \ldots , \sum_{j=1}^{m} y^j \frac{\partial v^m}{\partial w^j}|_{u^{-1}(x)} ). \tag{48}
\]
It is easily seen that the map \( (\bar{v} \circ \bar{u}^{-1}): \bar{u}(\bar{U} \cap \bar{V}) \subseteq \mathbb{R}^{2m} \mapsto \bar{v}(\bar{U} \cap \bar{V}) \subseteq \mathbb{R}^{2m} \) is a smooth map: By hypothesis the first $m$-components are smooth maps. The second $m$ components are linear in $y$ and smooth functions of $x$, and hence the combined map is smooth. If working in the class of $C^\alpha$-manifolds, this calculation shows that when one starts with a $C^\alpha$ atlas for $M$, then one obtains, as might be expected, a $C^{\alpha-1}$ atlas for $TM$.

In order for $TM$ to qualify as a smooth manifold we still need that the topology is reasonably nice (metrizable, or equivalently that $(T M, T)$ is paracompact). For the technical details we refer to Spivak vol.I ch.3, especially exercise 1. Here we sketch only some basic ideas. Since a manifold $M$ is locally homeomorphic to a Euclidean space and it is assumed to be metrizable (or equivalently paracompact), each connected component of $M$ is second countable. [\[This means that there is a countable basis for the topology on each connected component of $M$. A basis for a topology is a collection of open sets that covers the space, and such that whenever a point $x$ is contained in basic open sets $B_1$ and $B_2$, then there exists a basic open set $B_3$ such that $x \in B_3 \subseteq B_1 \cap B_2$.\]] Following Spivak vol.I ch.3 ex. 1, construct a sequence of functions that separates points and closed sets, use these to produce a sequence of bounded metrics $d_i$ and finally piece these together e.g. via $d = \sum_{i=1}^{\infty} 2^{-i} d_i$.

This construction of the tangent bundle shall serve as a model for similar constructions of more general vector bundles in which the tangent spaces $T_p M$ are replaced by other suitable vector spaces. Formally, a vector bundle is a triple \((E, B, \pi)\) (or actually, a five-tuple \((E, B, \pi, \oplus, \otimes)\)) consisting of a total space $E$, a base space $B$ and a bundle projection $\pi: E \mapsto B$ which is a continuous surjective map. The linear operations $\oplus$ and $\otimes$ are defined on the fibres $\pi^{-1}(p)$ for $p \in B$, making each fibre a vector space. A distinguishing condition is that a vector bundle is locally trivial, i.e. every $p \in B$ has an open neighborhood $U$ together with a homeomorphism $\beta: \pi^{-1}(U) \mapsto U \times \mathbb{R}^m$ such that for each $q \in U$ the restriction $\beta|_{\pi^{-1}(q)}$ is a vector space isomorphism from the fibre $\pi^{-1}(q)$ to $\{q\} \times \mathbb{R}^m$.

The Möbius strip is an example of a nontrivial line-bundle over the circle $S^1$. An upcoming section will introduce the cotangent bundle in which the fibre are the spaces of linear functionals on the corresponding tangent spaces. In some places it is convenient to work with functions that assign to each point $p \in M$ a pair, triple, or $m$-tuple of (co)tangent vectors. These may be thought of as sections of bundles in which each fibre is a product of two, three, or $n$ copies of the (co)tangent space.

Beyond vector bundles are fibre bundles in which the fibres need not necessarily be vector spaces. Arguably one of the most important ones is the principal bundle in which each fibre is a copy of the general linear group $GL(m, \mathbb{R})$ (the space of all invertible linear maps from $\mathbb{R}^m$ to...
$\mathbb{R}^m$). Its distinguishing feature is that each section $L: M \mapsto P$ (with $\pi \circ L = \text{id}_M$) acts on the manifold, e.g., if $(u, U)$ is a local coordinate chart, then $(L \circ u, U)$ is another chart (to be read as $(L \circ u: q \mapsto L_q \circ u(q), U)$ where $L_q: \mathbb{R}^m \mapsto \mathbb{R}^m$ is a linear map.)

**To be added:** Use tangent bundles for a geometric definition of orientability for a manifold $M$, or of vector bundle - as opposed to the purely algebraic condition in terms of charts, whether there exists atlas $\mathcal{A}$ such that $\text{det}(D(v \circ u^{-1})) > 0$ for all $(u, U), (v, V) \in \mathcal{A}$.

We digress with a brief discussion of the (lack of) triviality of the tangent bundles of spheres and its consequences. Consider the usual imbeddings of the spheres $S^m \hookrightarrow \mathbb{R}^{m+1}$ and use the standard coordinates in $\mathbb{R}^{m+1}$. Note that the tangent bundles $TS^m$ are diffeomorphic to the subsets $\{(a, b) \in S^m \times \mathbb{R}^{m+1} : \langle a, b \rangle = 0 \} \subseteq \mathbb{R}^{m+1} \times \mathbb{R}^{m+1}$, using the standard inner product in $\mathbb{R}^{m+1}$. [[This is completely different from asserting that $TS^m$ were trivial, or diffeomorphic to $S^m \times \mathbb{R}^m$]]

When $m = (2k - 1)$ is odd, then $X = x_2 D_1 - x_1 D_2 + x_1 D_3 - x_3 D_4 + \ldots + x_{2k} D_{2k-1} - x_{2k-1} D_{2k}$ (all $D_j$ evaluated at $x$) is a global nonvanishing (tangent) vector field on $S^m \subseteq \mathbb{R}^{m+1}$. Conversely, using tools from algebraic topology one may show that if $m = 2k$ is even, then not even a single globally defined continuous nonvanishing vector field exists on the sphere $S^{2k}$. Very special is $S^3$ which admits three smooth vector fields that are everywhere linearly independent:

\[
\begin{align*}
X &= -x_2 D_1 + x_1 D_2 - x_4 D_3 + x_3 D_4 \\
Y &= -x_3 D_1 + x_4 D_2 + x_1 D_3 - x_2 D_4 \\
Z &= -x_4 D_1 - x_3 D_2 + x_2 D_3 + x_1 D_4.
\end{align*}
\]

Mimicking (and repeating this construction, similar to the example for $S^{2k-1}$ above) one may construct from these three vector fields on $S^3$ three everywhere linearly independent vector fields on any sphere $S^{4k-1}$. However, it can be shown that on $S^{4k+1}$ any two smooth vector fields are linearly dependent at some point. This example of the frame of the three everywhere linearly independent vector fields on $S^3$ motivates the notion of a parallelizable manifolds [[Abraham-Marsden p.218; Boothby p.219; not in Spivak]]:

**Definition 3.4** A manifold $M^m$ is called parallelizable if it admits a frame of $m$ everywhere linearly independent vector fields.

It is straightforward to see that a [[finite-dimensional, c.f. Abraham-Marsden]] manifold is parallelizable if and only if its tangent bundle is trivial. So far we have seen that Euclidean spaces $\mathbb{R}^m$, hence all coordinate charts $(u, U)$ are parallelizable. Also, every Lie group is parallelizable.

**Should have been done much earlier, in chapter 2:** A Lie group is a differentiable manifold $G$ with a group structure such that both the multiplication : $G \times G \mapsto G$ defined by $(p, q) \mapsto pq$ and the inverse $G \mapsto G$, defined by $p \mapsto p^{-1}$ are $C^\infty$ maps.

We already encountered several examples of Lie groups: the general linear groups $\text{GL}(n, \mathbb{R})$ of invertible linear maps on $\mathbb{R}^n$, the special linear groups $\text{SL}(n, \mathbb{R})$ (linear maps with determinant one), and the orthogonal groups $\text{O}(n)$ and special orthogonal groups $\text{SO}(n)$. Note that as a manifold $S^1$ is diffeomorphic to $\text{SO}(2)$. Similarly, $S^3$ is a double-cover of the projective space $P^3$ which is diffeomorphic to $\text{SO}(3)$ – thus shedding some light on this most versatile example. In quantum mechanics these Lie groups are supplanted by their complex sisters, the unitary $U(n)$ and special unitary groups $\text{SU}(n)$. 

Exercise 3.9 Suppose $f \in C^\infty(\mathbb{R}^m, \mathbb{R})$. Show that the graph $\{(x, f(x)) : x \in \mathbb{R}^m\}$ is a parallelizable submanifold of $\mathbb{R}^{m+1}$. Is the same necessity true for functions $f : \mathbb{R}^m \mapsto \mathbb{R}^n$?

Returning to the tangent bundles of the spheres: It is known that the only parallelizable spheres are $S^1$, $S^3$, and $S^7$ [[Spivak vol.I, ch.3, ex. 19]]. It is no coincidence that these are the only dimensions in which one may endow the Euclidean spaces with some sort of a multiplicative structure: In $\mathbb{R}^2$ this is the field of complex numbers, in $\mathbb{R}^4$ this yields the noncommutative, but still associative quaternions (or Hamilton numbers), and in $\mathbb{R}^8$ these are the Cayley numbers whose multiplication is not even associative.

3.5 Smooth vector fields and Lie products

Recall, a vector field on a manifold is defined as a section of the tangent bundle, that is, a function $X : M \mapsto TM$ such that its composition $\pi \circ X$ with the bundle projection is the identity on $M$. Rather than considering arbitrary such functions, our interest is primarily in those that vary smoothly (in topological considerations continuity may suffice). Since a vector field is defined as a map between manifolds $M$ and $TM$ we already have a notion of smoothness: A vector field $X : M \mapsto TM$ is $C^r$ if for every point $p \in M$ and coordinate charts $(u, U)$ and $(\bar{v}, \bar{V})$ about $p$ and $X(p)$, respectively, the map $\bar{v} \circ X \circ u^{-1} : \mathbb{R}^m \mapsto \mathbb{R}^{2m}$ is a $C^r$-map between Euclidean spaces. We write $\Gamma^\infty(M)$ for the set of all smooth vector fields on $M$.

On the other hand recall that every tangent vector $X_p \in T_pM$ maps $C^\infty(p) \mapsto \mathbb{R}$. Consequently, we may view a vector field $X$ as a mapping of the algebra $C^\infty(M)$ of smooth functions to itself. We expect that if $X$ is a smooth vector field and $f \in C^\infty(M)$ then $(Xf) \in C^\infty(M)$. In particular, any smooth vector field is a derivation on the algebra $C^\infty(M)$ (i.e. it satisfies $X(fg) = (Xf)g + f(Xg)$ for all $f, g \in C^\infty(M)$).

Proposition 3.8 A vector field $X : M \mapsto TM$ is a $C^\infty$ vector field, written $X \in \Gamma^\infty(M)$, if and only if for every open set $U \subseteq M$ and every function $f \in C^\infty(U)$ the function $(Xf) : p \mapsto X(p)f$ is again in $C^\infty(U)$.

This proposition follows easily from the following exercise upon expanding $(Xf)(p)$ on a chart $(u, U)$ about $p$ in terms of local coordinates $(Xf)(q) = \sum_{j=1}^m (Xu^j)(q) \frac{\partial}{\partial u^j} \Big|_q$ and writing out the components of the map $\bar{v} \circ X \circ u^{-1} : \mathbb{R}^m \mapsto \mathbb{R}^{2m}$.

Exercise 3.10 Verify directly that a vector field $X : M \mapsto TM$ is $C^\infty$ if and only if for every coordinate chart $(u, U)$ of $M$ the functions $Xu^j : U \mapsto \mathbb{R}$ are smooth.

Since every (smooth) vector field $X \in \Gamma^\infty(M)$ maps $C^\infty(M)$ back into itself, it is natural to consider compositions of two vector fields $X, Y \in \Gamma^\infty(M)$. Clearly $X \circ Y$, also written $XY$, is again a map from $C^\infty(M)$ into itself. However, for two functions $f, g \in C^\infty(M)$ we calculate

$$XY(fg) = X \left( (Yf)g + f(Yg) \right) = (XYf)g + (Yf)(Xg) + (Xf)(Yg) + f(XYg). \quad (50)$$

In general there is no reason for the terms $(Yf)(Xg)$ and $(Xf)(Yg)$ to cancel each other, hence in general $XY$ is not a derivation, and thus is not a vector field! However, the commutator $XY - YX$ clearly will be a derivation. Thus we have a product structure on $\Gamma^\infty(M)$ – which equips this space of all smooth vector field with an important algebraic structure that invites deeper study:
Definition 3.5 The Lie bracket or Lie product of vector fields is the map
\[ [\cdot, \cdot] : \Gamma^\infty(M) \times \Gamma^\infty(M) \rightarrow \Gamma^\infty(M), \text{ defined for } f \in C^\infty(M) \text{ by } [X,Y]f = X(Yf) − Y(Xf). \]

Exercise 3.11 Let \( \xi = (\xi^1, \ldots, \xi^m)^T \) and \( \eta = (\eta^1, \ldots, \eta^m)^T \) be column vector fields representing two vector fields \( X, Y \in \Gamma^\infty(M) \) in a coordinate chart \((u, U)\), i.e. \( \xi^i = (Xu^i) \) and \( \eta^j = (Yu^j) \). Verify that in these coordinates the Lie product \( [X,Y] \) is represented by the column vector \( (D\eta)\xi − (D\xi)\eta \) where \( D \) denotes the Jacobian matrix of partial derivatives.

Definition 3.6 A linear vector space \( L \) equipped with a bilinear mapping \([\cdot, \cdot]: L \times L \rightarrow L\) is a Lie algebra if this map is anti-commutative and satisfies the Jacobi identity:
\[
\begin{align*}
\text{for all } x, y \in L, & \quad 0 = [x, y] + [y, x] \\
\text{for all } x, y, z \in L, & \quad 0 = [x, [y, z]] + [y, [z, x]] + [z, [x, y]].
\end{align*}
\]

Exercise 3.12 Verify that \( \mathbb{R}^3 \) equipped with the standard cross-product is a Lie algebra.

Exercise 3.13 Verify that the space \( \text{so}(3) \) of skew symmetric \( 3 \times 3 \)-matrices with the product \( [A,B] = AB − BA \) (matrix product) is a three dimensional Lie algebra. Find a basis for \( \text{so}(3) \) and establish a Lie algebra isomorphism from \( \text{so}(3) \) to \( \mathbb{R}^3 \) with the cross-product – i.e. explicitly give a bijective linear map (between vector spaces) that is also a Lie algebra homomorphism, meaning in this case \( \Phi([A,B]) = \Phi(A) \times \Phi(B) \) for all \( A, B \in \text{so}(3) \).

Exercise 3.14 Verify by direct calculation that the Lie product of vector fields as defined above equips \( \Gamma^\infty(M) \) with a Lie algebra structure. Note that this means verifying that \([\cdot, \cdot, \cdot]\) is linear over \( \mathbb{R} \), i.e. \([aX + Y, Z] = a[X, Z] + [Y, Z] \), that it is anticommutative (obvious) and that it satisfies the Jacobi identity – simply expand \([X, [Y, Z]]f + [Y, [Z, X]]f + [Z, [X, Y]]f\).

Exercise 3.15 Show that any associative algebra \((A, \cdot)\) is a Lie algebra with the commutator product that is defined by \([x, y] = x \cdot y − y \cdot x\). In particular, the set \( \mathcal{D}(A) \) of derivations on an associative algebra, that is of linear maps \( \ell : A \rightarrow A \) satisfying \( \ell(xy) = (\ell(x))y + x\ell(y) \) for all \( x, y \in A \) is an associative algebra under composition and thus a Lie algebra under the commutator as above.

Exercise 3.16 (This is a preview of an example which geometrically is situated in the cotangent bundle and symplectic geometry). On the set of all smooth functions \( C^\infty(\mathbb{R}^{2m}) \) define a product, the Poisson bracket, using coordinates \((q_1, \ldots, q_m, p_1, \ldots, p_m)\) on \( \mathbb{R}^{2m} \) by
\[
\{f, g\} = \sum_{i=1}^m \frac{\partial f}{\partial p^i} \frac{\partial g}{\partial q^i} − \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p^i}. \tag{52}
\]
Verify that this product equips \( C^\infty(\mathbb{R}^{2m}) \) with the structure of a Lie algebra.

Recall that if \( X, Y \in \Gamma^\infty(M) \) and \( f, g \in C^\infty(M) \) then \( fX \in \Gamma^\infty(M) \), and the usual distributive and mixed associative properties hold, e.g. \((fg)X = f(gX), f(X + Y) = fX + fY, (f + g)X = fX + gX, 1 \cdot X = X\). This means that \( \Gamma^\infty(M) \) is not only a vector space over \( \mathbb{R} \), but also a (left) \( C^\infty(M) \)-module. (It is not a vector space over \( C^\infty(M) \) since the ring of smooth functions
is not a field.) Given this \( C^\infty(M) \)-module structure it is natural to ask how the Lie bracket on \( \Gamma^\infty(M) \) relates to it. For \( X, Y \in \Gamma^\infty(M) \) and \( f, g \in C^\infty(M) \) calculate

\[
[fX,Y]_g = (fX)(Yg) - Y(fX(g)) = f \left( X(Yg) - Y(Xg) \right) - (Yf) \cdot (Xg) = (f[X,Y] - (Yf)X) g
\]

and conclude that the Lie bracket \([ \cdot, \cdot ]\) is not linear over \( C^\infty(M) \).

In a chart \((u,U)\) (compare also exercise 3.20) calculate

\[
\left[ \frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^j} \right] f = \frac{\partial}{\partial u^i} (D_j (f \circ u^{-1}) \circ u) - \frac{\partial}{\partial u^j} (D_i (f \circ u^{-1}) \circ u)
\]

\[
= D_i (D_j (f \circ u^{-1}) \circ u \circ u^{-1}) \circ u - \frac{\partial}{\partial u^j} (D_i (f \circ u^{-1}) \circ u \circ u^{-1}) \circ u
\]

\[
= \left( D_i D_j (f \circ u^{-1}) - D_j D_i (f \circ u^{-1}) \right) \circ u
\]

\[
\equiv 0
\]

since the mixed partial derivatives on \( C^\infty \mathbb{R}^m \) are equal. As an important corollary we obtain:

**Proposition 3.9** If \( X, Y \in \Gamma^\infty(M) \) and \( U \subseteq M \) is an open set with \( [X,Y]|_U \neq 0 \) then there does not exist a map \( u : U \rightarrow \mathbb{R}^m \) such that \((u,U)\) is a chart on \( M \) with \( X|_U = \frac{\partial}{\partial u^1} \) and \( Y|_U = \frac{\partial}{\partial u^2} \).

Indeed, in subsequent sections we will see that in the neighborhood of any point \( p \) at which a smooth vector field \( X \) does not vanish, there are always coordinates \((u,U)\) such that \( X = \frac{\partial}{\partial u^1} \). On the other hand, generalizing the above criterion to sets of vector fields will lead to important Frobenius integrability theorem.

**Exercise 3.17** [[This exercise is somewhat frivolous – but it is a good practice for hands-on calculations, and it hits hard at common misperceptions.]] Consider the upper half plane \( M = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 > 0\} \) with standard rectangular coordinates \((x,y)\), with polar coordinates \((r,\theta)\) and with the mixed coordinates \((\rho,\xi)\) defined by \( \rho = r \) and \( \xi = x \).

- Explicitly express the coordinate vector fields \( \frac{\partial}{\partial r} \) and \( \frac{\partial}{\partial \theta} \) as linear combinations of \( \frac{\partial}{\partial x} \) and \( \frac{\partial}{\partial y} \). (In particular express the coefficients in terms of \( x \) and \( y \)).
- Use these expressions to verify by direct calculation that \([ \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta} ] \equiv 0 \).
- Verify by direct calculation that \([ \frac{\partial}{\partial \rho}, \frac{\partial}{\partial x} ] \neq 0 \).
- Explain why this does not contradict that \((\xi,\rho) = (x,r)\) are admissible local coordinates. Calculate the \((\xi,\rho)\) coordinates of the points \((1,0.1)\), \((1,1)\), and \((0,1)\).
- Calculate \( \frac{\partial}{\partial \xi} \) and \( \frac{\partial}{\partial \rho} \), e.g. write these as linear combinations of \( \frac{\partial}{\partial x} \) and \( \frac{\partial}{\partial y} \), and sketch these coordinate vector fields as arrows in the half-plane. Describe in words in which directions these arrows point.
- Explain where such possible misconceptions might come from. In some sense, for partial derivatives it is less important what varies than what held fixed ... Revisit this discussion later using differential forms \( dx, d\xi, d\rho, d\eta \).
3.6 The tangent map and vector fields

Having assembled the tangent spaces $T_pM$ at all points $p \in M$ into the tangent bundle $TM$ as a manifold, it is natural to combine the tangent maps $\Phi_{sp}$ associated to a map $\Phi \in C^\infty(M,N)$ between manifolds into a map $\Phi_* : TM \mapsto TN$. This is a straightforward definition with no or few ensuing surprises. However, in general, tangent maps need not map vector fields to vector fields.

**Definition 3.7** For any map $\Phi \in C^\infty(M,N)$ define an associated tangent map $\Phi_* : TM \mapsto TN$ for $q \in M, \ (q, X_q) \in \pi^{-1}(q)$ by

$$\Phi_*(q, X_q) = (\Phi(q), \Phi_{sq}(X_q)).$$

(55)

Note that the tangent map $\Phi_*$ has the map $\Phi$ built in. It is straightforward to verify the following:

**Proposition 3.10** If $\Phi \in C^\infty(M,N)$ and $\Psi \in C^\infty(N,P)$ then (note preservation of order)

$$(\Psi \circ \Phi)_* = \Psi_* \circ \Phi_*.$$  (56)

**Exercise 3.18** Show that if $\Phi \in C^\infty(M,N)$ then $\Phi_* \in C^\infty(TM, TN)$. (Use the definition of differentiability of a map between manifolds in terms of charts $(\bar{u}, \bar{U})$ and $(\bar{v}, \bar{V})$ for $M$ and $N$, respectively.)

It is important to understand that in general there is no hope that a tangent map associated to a smooth $\Phi : M \mapsto N$ between manifold will map a vector field $X$ on $M$ to a vector field on $N$. This is immediately clear if we recall that a vector field on $N$ is a function from $N$ to $TN$. Thus if $p_1 \neq p_2 \in M$ but $\Phi(p_1) = \Phi(p_2) \in N$ then problems arise unless $\Phi_{sp_1}X_{p_1} = \Phi_{sp_2}X_{p_2}$. Similarly, if $\Phi$ is not onto, then $\Phi_*$ can at best yield a partially defined vector field on $N$.

If $\Phi : M \mapsto N$ is a diffeomorphism and $X \in \Gamma^\infty(M)$ then, with some abuse of notation, define $\Phi_* X : N \mapsto TN$ by

$$(\Phi_* X)_q = \Phi_{s\Phi^{-1}(q)}X_{\Phi^{-1}(q)} \quad \text{for} \quad q \in N.$$  (57)

Sometimes this is written suggestively as $(\Phi_* X)_q = (\Phi_* X \circ \Phi^{-1})(q)$. It is clear that $\pi \circ (\Phi_* X) = id_N$ and that $\Phi_*$ is a smooth map, and hence $(\Phi_* X) \in \Gamma^\infty(N)$.

An important application is when $\Phi = u : U \mapsto \mathbb{R}^m$ is a coordinate map. Indeed, we have routinely used the map $u_*$ which maps e.g. coordinate vector fields $\partial \overline{u}_{ji}$ to the fields $D_j$ on $\mathbb{R}^m$. Note that it is not required for $\Phi$ to be a diffeomorphism in order for $(\Phi_* X)$ to make sense as a vector field on $N$ – as long as $\Phi$ is a smooth map such that $p_1 = p_2 \in M$ implies $\Phi_{sp_1}X_{p_1} = \Phi_{sp_2}X_{p_2}$ the definition (57) still makes sense. The following example gives a preview on how this may be used in the case that the vector field has some infinitesimal symmetries as they will be defined in the section on Lie derivatives.

**Exercise 3.19** Consider $M = \mathbb{R}^2 \setminus \{0\}$ and $N = \mathbb{P}^1$. Let $X(x) = (ax^1 + bx^2) \ D_1 |_x + (cx^1 + dx^2) \ D_2 |_x$ be a linear vector field on $M$. Define a relation $\sim$ on $M$ by $x \sim y$ if there exists $\lambda \in \mathbb{R}$ such that $x = \lambda y$. Verify that $\sim$ is an equivalence relation on $M$.

Let $\Phi : M \mapsto N = \mathbb{P}^1 = \mathbb{R}_*$ be the canonical projection map which maps each $x \in M$ to its equivalence class $[x] = \{y \in M : y \sim x\}$. Verify that if $x \sim y$ then $\Phi_{sx}(X_x) = \Phi_{sy}(X_y)$ and hence we may $\Phi_* (X)$ does define a smooth vector field on $\mathbb{P}^1$.

Consider the local coordinate chart $(m, U)$ on $\mathbb{P}^1$ where $U = \left\{\{x_1, x_2\} : x_1 \neq 0\right\}$ and $m(\{(x_1, x_2)\})$
is the slope of the line through the points \([(x_1, x_2)]\), i.e. \(m([x_1, x_2]) = \frac{y_2}{x_1}\). Find an explicit expression for \((\Phi, X)_m = f(m) \frac{\partial}{\partial m} |_m\). Interpret \((\Phi, X)\) as defining a dynamical system \(\dot{m} = f(m)\) on the space of lines through the origin. In detail discuss the special cases when \(a = d\) and either \(b = c = 0\) or \(b = -c = -1\). In general relate the stationary points of \(\Phi, X\) (i.e. the zeros of \(f(\theta)\)) to the eigenspaces of the \(2 \times 2\)-matrix with entries \(a, b, c\) and \(d\).

**Exercise 3.20** Extend the previous exercise 3.19 to a higher dimensional case. Let \(X(x) = \sum_{i,j=1}^{m} a_{ij} x^i D_j |_x\) be a linear vector field on \(\mathbb{R}^m\). Use the coordinates \(y = (y^1, \ldots, y^{m-1})\) on a suitable subset \(U \subseteq \mathbb{P}^{m-1} = (\mathbb{R}^m \setminus \{0\})/\sim\) defined by \(y^i([x]) = \frac{x^i}{x^m}\) and verify that in these coordinates \((\Phi, X)\) is a quadratic vector field (representing a Riccati differential equation).

As an illustration explicitly write out the formula for \((\Phi, X)\) in the case of \(m = 1\). For fun explore the case where the matrix \((a_{ij})\) has a triple eigenvalue with a single Jordan block, e.g. \(a_{ii} = \lambda \neq 0, a_{12} = a_{23} = 1\) and \(a_{ij} = 0\) else. In particular, sketch the phase portrait for \((\Phi, X)\) near \(y = 0\) and relate it to the integral curves on \(\mathbb{P}^2\) (or on \(S^2\) which may be easier to visualize).

### 3.7 The cotangent bundle and differential one-forms

Associated to each tangent space \(T_p M\) of a manifold \(M\) at a point \(p\) is a well-defined dual space whose elements are the linear functionals on \(T_p M\). Assembling all these dual spaces one obtains the cotangent bundle. Its sections, the analogues to (tangent) vector fields, are differential forms. While such dual objects appear to be considerably less tangible to the novice, they do have better algebraic properties than tangent vector fields. This makes them the preferred choice in the many settings where one may choose between describing objects and properties using tangent fields or cotangent fields. We begin with a brief linear algebra review.

Let \(V\) be a finite dimensional vector space (over a field, here always taken to be \(\mathbb{R}\)). A linear functional on \(V\) is a linear map \(\lambda : V \mapsto \mathbb{R}\) (i.e. \(\lambda(c v + w) = c\lambda(v) + \lambda(w)\) for all \(v, w \in V\) and all \(c \in \mathbb{R}\)). The set \(V^*\) of all linear functionals on \(V\) inherits a scalar multiplication and addition from the codomain \(\mathbb{R}\), i.e. for linear functionals \(\lambda_1, \lambda_2\) on \(V\), \(c \in \mathbb{R}\), and \(v \in V\) define \((c\lambda_1 + \lambda_2)(v) = c\lambda_1(v) + \lambda_2(v)\). It is a straightforward to check that with these operations the set \(V^*\) is a vector space over \(\mathbb{R}\).

**Exercise 3.21** Suppose \(\beta = \{v_1, \ldots, v_m\}\) is a basis for a vector space \(V\). Consider the maps \(\lambda^j : V \mapsto \mathbb{R}\) defined by

\[
\lambda^j \left( \sum_{j=1}^{m} c^j v_j \right) = c^j \quad \text{where} \quad c^j \in \mathbb{R}.
\]

- Verify that \(\lambda^j \in V^*\).
- Show that \(\gamma = \{\lambda^1, \ldots, \lambda^m\}\) are linearly independent.
- Show that every linear functional \(\lambda \in V^*\) is a linear combination of \(\gamma\).

The exercise establishes, in particular, that \(V^*\) is of the same dimension as \(V\). The basis \(\gamma\) for \(V^*\), described in this exercise, is called the dual basis to \(\beta\).

Novices to linear algebra often seem troubled that unlike the elements of the given vector space \(V\) the elements of \(V^*\) seem to be less tangible, and that they can be represented by an somewhat arbitrary collection of different objects. However, this drawback is easily compensated for by their superior algebraic properties . . . The following exercise may help a little pinning down what the linear functionals are (and what they are not).
Exercise 3.22 [[This is not meant to be deep, but should be fun and provide a hands-on different point of view.]] Consider the vector space $V$ of all quadratic polynomial functions on the real line. In the usual shorthand notation $V = \{a + bx + cx^2: a, b, c \in \mathbb{R}\}$.

- Verify that $\lambda_1: p \mapsto p(1)$, $\lambda_2: p \mapsto p'(23)$, $\lambda_3: p \mapsto \int_0^1 p(t) \, dt$, and $\lambda_4: p \mapsto \int_\infty^{-2} e^{-t^2} p(t) \, dt$, are linear functionals on $V$.
- Show that $\{\lambda_1, \lambda_2, \lambda_3\}$ is a basis for $V^*$.
- Write $\lambda_4$ as a linear combinations of $\lambda_1, \lambda_2$ and $\lambda_3$.
- Find a basis for $V^*$ that is dual to the basis $\{1, x, x^2\}$ for $V$.
- Find a basis for $V^*$ that is dual to the point evaluations $\epsilon_j: p \mapsto p(j)$ for $j = 1, 2, 3$.
- Explain why for every fixed integer $N > 0$ and every fixed interval $[a, b]$ there exist fixed numbers $\alpha_j, \xi_j \in \mathbb{R}$ (not necessarily in $[a, b]$) such that for every polynomial function $p$ of degree at most $(N - 1)$, $\int_a^b p(t) \, dt = \sum_{j=1}^N \alpha_j p(\xi_j)$. (E.g. use that Vandermonde matrices are nonsingular.)

This example will be revisited in the next chapter in the context of inner product spaces.

Returning to differential geometry, define

Definition 3.8 Suppose $M$ is a smooth manifold and $p \in M$. The cotangent space to $M$ at $p$, denoted $T_p^* M$, is the space of all linear functionals on $T_p M$, i.e. $T_p^* M = (T_p M)^*$.

Recall that we defined tangent vectors $X_p \in T_p M$ to be linear mappings from $C^\infty(p)$ to $\mathbb{R}$. Turning this around we define:

Definition 3.9 For $p \in M$ and $f \in C^\infty(p)$ define a map $(df)_p: T_p M \mapsto \mathbb{R}$, called the differential of $f$ at $p$, by

$$(df)_p(X_p) = (X_p f).$$

Exercise 3.23 Verify that for each $p \in M$ and each $f \in C^\infty(p)$ the differential $(df)_p$ is a linear functional on $T_p M$, i.e. $(df)_p \in T_p^* M$.

Proposition 3.11 Suppose that $(u, U)$ is a chart about $p \in M^m$. Then the set $\{(du^1)_p, \ldots, (du^m)_p\}$ of differentials at $p$ is a basis for $T_p^* M$, dual to the basis $\{\frac{\partial}{\partial u^1}|_p, \ldots, \frac{\partial}{\partial u^m}|_p\}$ of $T_p M$.

Proof. From the definition it is clear that

$$(du^i)_p(\frac{\partial}{\partial u^j}|_p) = D_j(u^i \circ u^{-1})|_{u(p)} = \delta^i_j$$

which shows the linear independence of $\gamma = \{(du^1)_p, \ldots, (du^m)_p\}$. Since the cardinality of $\gamma$ matches the dimension on $T_p M$, this also establishes that $\gamma$ is a basis for $T_p^* M$. ■

Note that in a chart $(u, U)$ the coordinates $\omega_j$ of any element $\omega = \sum_{j=1}^m \omega_j (du^j)_p \in T_p^* M$ are immediately obtained by evaluating $\omega_j = \omega(\frac{\partial}{\partial u^j}|_p)$. In particular, if $f, g \in C^\infty(p)$ are such that

for all $j = 1, \ldots, m$, $\frac{\partial f}{\partial u^j}|_p = \frac{\partial g}{\partial u^j}|_p$ then $(df)_p = (dg)_p$ as elements of $T_p^* M$.

It is useful to compare the notion of differential forms developed here to the common usage in calculus. For illustration consider the function $z = x^2 + y^2$ (i.e. $z: \mathbb{R}^2 \mapsto \mathbb{R}$), whose differential is $dz = 2x \, dx + 2y \, dy$. Commonly $dz$ is considered as a function of the four variables $x, y, dx$ and
Suppose $C X$. Indeed, with differential forms we now alternatively may express a tangent vector structure and properties of $\Omega_{u,U}$. In our notation this tangent vector is written as $0$. Note that it is consistent with our language to use $dz$ and $dy$ as coordinates in the tangent plane – we merely may regard them as linear functions, here on $T_{(2,3)} \mathbb{R}^2$. On manifolds, we clearly distinguish between $(dx)_p$ and $(dx)_q$ at different points (just as we associate tangent vectors to fixed points). In particular, $dx = 0.2$ is simply a shorthand for $(dx)_{(2,3)}(0.2 \frac{\partial}{\partial x}|_{(2,3)} - 0.1 \frac{\partial}{\partial y}|_{(2,3)}) = 0.2$.

Indeed, with differential forms we now alternatively may express a tangent vector $X_p \in T_p M$ in a chart $(u, U)$ about $p$

$$X_p = \sum_{j=1}^{m} (X_p u^j) \frac{\partial}{\partial u^j} \big|_p \quad \text{or} \quad X_p = \sum_{j=1}^{m} (du^j)_p (X_p) \frac{\partial}{\partial u^j} \big|_p \quad (61)$$

and $(du^j)_p$, $j = 1, \ldots, m$ are legitimate coordinate functions, or simply “coordinates” (?) of tangent vectors.

In complete analogy to the tangent bundle assemble all cotangent spaces $T^*_p M$ into the cotangent bundle, denoted $T^* M$. It is a vector bundle over $M$ with bundle projection again denoted by $\pi$. For any chart $(u, U)$ of $M$ define $\bar{U} = \pi^{-1}(U)$, and, $\bar{u}: \bar{U} \mapsto \mathbb{R}^m$ by

$$\bar{u}(p, \omega_p) = (u^1(p), \ldots, u^m(p), \omega_p(\frac{\partial}{\partial u^1}|_p), \ldots, \omega_p(\frac{\partial}{\partial u^m}|_p)). \quad (62)$$

Using proposition 3.11 it is clear that $\bar{u}$ is a bijection onto its image. As in the case of $TM$, it is possible to equip $T^* M$ with a topology such that the maps $\bar{u}$ are homeomorphisms (onto their respective images). In more technical forms one may show that the topology is metrizable, and via the next exercise, $T^* M$ is a smooth manifold.

**Exercise 3.24** Suppose $(u, U)$ and $(v, V)$ are charts on $M$, and $(\bar{u}, \bar{U}), (\bar{v}, \bar{V})$ are defined as above. Verify that the transition maps $\bar{v} \circ \bar{u}^{-1}: \bar{u}(\bar{U} \cap \bar{V}) \mapsto \bar{v}(\bar{U} \cap \bar{V}) \subseteq \mathbb{R}^m$ are smooth bijections between subsets of Euclidean spaces.

**Definition 3.10** The (smooth) sections of the cotangent bundle, that is, the (smooth) functions $\omega: M \mapsto T^* M$ satisfying $\pi \circ \omega = \text{id}_M$ are called (smooth) differential one-forms. The space of all smooth differential one-forms on $M$ is denoted by $\Omega^1(M)$.

As a map between smooth manifolds, a section $\omega: M \mapsto T^* M$ is smooth if for all coordinate charts $(u, U)$ of $M$ and $(\bar{v}, \bar{V})$ of $T^* M$ the maps $\bar{v} \circ \omega \circ u^{-1}: u(U \cap V) \mapsto \bar{v}(U \cap V) \subseteq \mathbb{R}^m$ are smooth as maps from subsets of $\mathbb{R}^m$ to subsets of $\mathbb{R}^2m$.

In particular, in a chart $(u, U)$ a differential one-form $\omega \in \Omega^1(M)$ may be written as a linear combination $\omega = \sum_{j=1}^{m} \omega_j du^j$ with smooth functions $\omega_j \in C^\infty(M)$ defined by $\omega_j = \omega(\frac{\partial}{\partial u^j})$.

For every function $f \in C^\infty(U)$ defined on an open subset $U \subseteq M$ define a smooth differential one-form $df \in \Omega^1(U)$ by $(df)(p) = (df)_p$ for $p \in U$. (More pedantic writers may prefer $(df)(p) = (p, (df)_p)$.) In a coordinate chart $(u, U)$ this becomes $df = \sum_{j=1}^{m} \frac{\partial f}{\partial u^j} du^j$.

After this brief discussion of smoothness and local representations we take a look at the algebraic structure and properties of $\Omega^1(M)$. We already have routinely combined smooth functions $f, g \in C^\infty(M)$ and differential forms $\omega, \eta \in \Omega^1(M)$ to write e.g. $(f \omega + g \eta)$. A brief reflection shows,
that this is permissible, and indeed yields new differential forms in $\Omega^1(M)$: Indeed any $T^*_p M$ is a $\mathbb{R}$-vector space and for any $X_p \in T_p M$ we interpret $(f\omega + g\eta)_p(X_p)$ as $f(p)\omega_p(X_p) + g(p)\eta_p(X_p)$. The next exercise addresses the smoothness.

**Exercise 3.25** Suppose $f, g \in C^\infty(M)$ and $\omega, \eta \in \Omega^1(M)$. Argue (from the definition of smoothness of maps between manifolds) why $f\omega + g\eta$ is indeed a smooth differential form on $M$.

It is a straightforward to verify that the usual mixed associative and mixed distributive laws hold, and thus $\omega \in \Omega^1(M)$ has the structure of a $C^\infty(M)$ module.

On the other hand, every differential one-form $\omega \in \Omega^1(M)$ is naturally also a functional that maps $\omega: \Gamma^\infty(M) \mapsto C^\infty(M)$, defined pointwise by $\omega(X)(p) = \omega_p(X_p)$. To verify that $\omega(X)$ is indeed a smooth map locally expand $\omega(X)$ in a coordinate chart $(u, U)$

$$\omega(X) = (\sum_{i=1}^{m} \omega_i du^i)(\sum_{j=1}^{m} (Xu^j) \frac{\partial}{\partial u^j}) = \sum_{i=1}^{m} \omega_i \cdot (Xu^i) \quad (63)$$

and use that $\omega_i$ and $(Xu^i)$ are smooth functions since $\omega$ and $X$ are a smooth differential form and a smooth vector field, respectively.

Moreover, one readily observes that if $f \in C^\infty(M)$ (in a chart $(u, U)$

$$\omega(fX) = \sum_{i=1}^{m} \omega_i du^i (\sum_{j=1}^{m} (fXu^j) \frac{\partial}{\partial u^j}) = \sum_{j=1}^{m} f \cdot \omega_i \cdot Xu^i = f \sum_{i=1}^{m} \omega_i du^i (\sum_{j=1}^{m} (Xu^j) \frac{\partial}{\partial u^j}) = f\omega(X) \quad (64)$$

establishing that any $\omega \in \Omega^1(M)$ is not only an $\mathbb{R}$-linear map, but indeed a $C^\infty(M)$-linear map from $\omega \in \Gamma^\infty(M)$ to $C^\infty(M)$, written

$$\Omega^1(M) \subseteq \text{Hom}_{C^\infty(M)}(\Gamma^\infty(M), C^\infty(M)). \quad (65)$$

### 3.8 Cotangent maps and pullbacks of differential forms

The next step is to analyze the analogues of the tangent maps associated to a smooth function between manifolds. Recall from linear algebra that every linear map $\phi: V \mapsto W$ between vector spaces induces a dual map $\phi^*: W^* \mapsto V^*$, defined by $(\phi^*\lambda)(v) = (\lambda \circ \phi)(v)$ for $v \in V$ and $\lambda \in W^*$.

**Definition 3.11** Suppose $\Phi \in C^\infty(M, N)$ is a smooth map and $p \in M$. Define the cotangent map $\Phi^*_p: T^*_p(N) \mapsto T^*_p M$ as the dual of the tangent map $\Phi_p$, i.e. $\Phi^*_p = (\Phi_p)^*$. Note that this means if $\omega_{\Phi(p)} \in T^*_p(N)$ and $X_p \in T_p M$ then

$$(\Phi^*_p \omega_{\Phi(p)})(X_p) = \omega_{\Phi(p)}(\Phi_p X_p). \quad (66)$$

**Exercise 3.26** Let $\Phi \in C^\infty(M, N)$, $\Psi \in C^\infty(N, P)$, and $p \in M$. Verify $(\Psi \circ \Phi)^*_p = \Phi^*_p \circ \Psi^*_p(p)$. Unlike the situation of the tangent bundle it is in general not possible to combine all maps $\Phi^*_p$, $p \in M$ together to get a well-defined map from $T^*N$ to $T^*M$. Indeed, the first hint at problems is that the maps $\Phi^*_p$ are naturally indexed not by their domains but by their codomains! Indeed, if $p, q \in M$ are such that $z = \Phi(p) = \Phi(q) \in N$ then there are well-defined maps $\Phi^*_z: TN \mapsto T^*_p M$,
$\Phi^*_q: T^*_N \mapsto T^*_q M$ with the same domain, but different codomains (unless this implies $p = q$, i.e., unless $\Phi$ is one-to-one). Nonetheless, in the case that $\Phi$ is one-to-one (i.e. especially if $\Phi$ is a diffeomorphism) define $\Phi^* : T^* N \mapsto T^* M$ pointwise by $\Phi^*(\omega_{\Phi(p)}) = \Phi^*_p(\omega_{\Phi(p)})$ for $p \in M$ and $\omega_{\Phi(p)} \in T^*_{\Phi(p)} N$.

This lack of well-defined cotangent maps between cotangent bundles is a small price to pay for now being able to map sections: Recall, that in general it is not possible to map a vector field $X : M \mapsto TM$ forward to a vector field $\Phi_*X : N \mapsto TN$. However, it is always possible to pull back differential forms (along smooth maps):

**Definition 3.12** If $\Phi \in C^\infty(M, N)$ and $\omega \in \Omega^1(N)$ define the pullback $\Phi^*\omega : M \mapsto T^* M$ of $\omega$ by $\Phi$ to $M$ for $p \in M$ by

$$(\Phi^*\omega)(X_p) = \omega_{\Phi(p)}(\Phi_*X_p).$$

**Exercise 3.27** Suppose $\Phi \in C^\infty(M, N)$ and $\omega \in \Omega^1(N)$. Verify directly that $(\Phi^*\omega) \in \Omega^1(M)$, i.e. that $\Phi^*\omega$ is smooth.

This is a good place to comment about some unfortunate terminology. Associated to a map $\Phi : M \mapsto N$ are two maps, $\Phi_* : TM \mapsto TN$, going in the same direction, and $\Phi^* : T^* N \mapsto T^* M$, going in the opposite direction. Modern language would use the attribute covariant for the first, and the attribute contravariant for the latter. Unfortunately, classical language used the same words for co-tangent and tangent vector fields. Quoting from Spivak vol.I, p.156 “... and no one had the gall or authority to reverse terminology so sanctioned by years of usage. So it’s very easy to remember which kind of vector field is covariant, and which is contravariant – it’s just the opposite of what it logically ought to be. (I.e. sections $X : M \mapsto TM$ are called covariant vector fields, and sections $\omega : M \mapsto T^* M$ are called covariant vector fields ...)

Pullbacks of cotangent vector fields are especially useful when working with imbedded submanifolds. More specifically, suppose that $M \subseteq N$ is a submanifold and consider the inclusion map $\imath : M \hookrightarrow N$. Then every differential form $\omega \in \Omega^1(N)$ immediately gives rise to a differential form $\imath^*(\omega) \in \Omega^1(M)$. Indeed, this is used so often that one routinely even uses the same symbol $\omega$ for $\imath^*(\omega)$. On the side note that there is no equivalent to this for tangent vector fields: Indeed, for any vector field $X \in \Gamma^\infty(M)$ there are in general many extensions to a vector field on $N$. Conversely if $N \subseteq M$ is a submanifold of positive codimension and $\Phi \in C^\infty(M, N)$ then $\Phi$ is necessarily many-to-one and unless something special happens there is little hope that the collection of tangent vectors $\Phi_*X_p$ (with $p \in M$) are the image of a vector field on $N$.

In practical examples one routinely needs to calculate the pullbacks of differential forms in terms of local coordinates. Thus consider a smooth map $\Phi \in C^\infty(M, N)$, local and coordinate charts $(u, U)$ about a point $p \in M$ and $(v, V)$ about $\Phi(p) \in N$. Due to the linearity of $\Phi^*_p$ it suffices to consider the pullbacks $\Phi^* (dv^i)$. As an immediate consequence of the earlier calculations (3.3) of the tangent map in coordinates find

$$(\Phi^*_p(dv^i)) \left. \frac{\partial}{\partial u^j} \right|_p = dv^j \left( \Phi^*_p \left. \frac{\partial}{\partial u^j} \right|_p \right) = dv^j \left( \sum_{\ell=1}^n \left. \frac{\partial (v^i \circ \Phi)}{\partial u^\ell} \right|_p \cdot \frac{\partial}{\partial v^i} \left|_{\Phi(p)} \right. \right) = \frac{\partial (v^i \circ \Phi)}{\partial u^j}$$

and consequently for $\omega_i \in \Omega^1(N)$

$$\Phi^* \left( \sum_{i=1}^n \omega_i\ dv^i \right) = \sum_{j=1}^m \left( \sum_{i=1}^n \omega_i \frac{\partial (v^i \circ \Phi)}{\partial u^j} \right) \cdot dv^j.$$
As expected this means that the coordinates transform by matrix-multiplication. One may look at this in different ways: If assembling the coordinates \( \omega_i \) into column vectors then the coordinates of the image are obtained by left multiplication by the transpose of the usual Jacobian matrix with components \( \frac{\partial(v \circ \Phi)}{\partial u} \). A more elegant way to interpret the sum in equation (69) is in terms of right multiplication of row vectors by the standard Jacobian matrix – no transpose. Thus if we write \( a = (\omega_1, \ldots, \omega_n) \) and \( b = (du^1(\Phi^*\omega), \ldots, du^n(\Phi^*\omega)) \) then \( b = aC \) where \( C \) is the matrix with components \( C_{ij} = \frac{\partial(v \circ \Phi)}{\partial u} \).

Consistently using this convention of representing (in local coordinates) tangent vector fields by column vectors and differential forms by row vectors facilitates many calculations. In particular, the evaluation of a differential form on a tangent vector field becomes in coordinates simply the matrix product of a row vector with a column vector (in this order). Moreover, the defining equation \( (\Phi^*\omega)X_p = \omega(\Phi(X_p)) \) is simply interpreted as associativity of matrix multiplication: Let, as before, \( a = (\omega_1, \ldots, \omega_n) \) denote the coordinates of a differential form \( \omega \) on \( N \), \( C \) the Jacobian matrix with components \( C_{ij} = \frac{\partial(v \circ \Phi)}{\partial u} \), and let now \( \xi = (X_p u^1, \ldots, X_p u^n)^T \) denote the column vector of the \( u \)-coordinates of the tangent vector \( X_p \in T_p M \). Then we simply have

\[
(\Phi^*\omega)X_p = \omega(\Phi(X_p)) \quad \longrightarrow \quad (aC)\xi = a(C\xi).
\]

Formally, it is at times convenient to assemble the basis vectors into formal row and column vectors. To be consistent introduce the formal column vectors \( \alpha = (\Phi^*du^1, \ldots, \Phi^*du^n)^T \) and \( \beta = (du^1, \ldots, du^n)^T \). Then \( \alpha = C\beta \) from (68). Together with the notation of the previous paragraph, this provides for such nice shorthand notation as

\[
\Phi^*\omega = a\alpha = a(C\beta) = (aC)\beta = b\beta.
\]

**Exercise 3.28** Suppose \( \Phi \in C^\infty(M^m, N^n) \) and \( \Psi \in C^\infty(N^n, P^r) \) are smooth maps between manifolds, \( p \in M \) and \( X_p \in T_p M \). Furthermore, suppose \( (u, U), (v, V) \) and \( (w, W) \) are local coordinate charts about \( p \in M \), \( \Phi(p) \in N \) and \( (\Psi \circ \Phi)(p) \in P \), respectively. Verify that the matrix representing \( (\Psi \circ \Phi)^* \) with respect to \( (u, U) \) and \( (w, W) \) is the product of the matrices representing \( \Phi^*_p \) (with respect to \( (u, U) \) and \( (v, V) \)) and \( \Psi^*_p \) (with respect to \( (v, V) \) and \( (w, W) \)).

**Exercise 3.29** Revisit the exercise 3.17 with \( M = \{ x \in \mathbb{R}^2 : x^2 > 0 \} \) equipped with rectangular coordinates \( (x, y) \), polar coordinates \( (r, \theta) \), and the mix \( (\xi, \rho) \) defined by \( \xi = x \) and \( \rho = r \).

For each pair of coordinates calculate the Jacobian matrix \( C \) with components \( C_{ij} = \frac{\partial(v \circ \Phi)}{\partial u} \), when \( \Phi = \text{id}_M \) is the identity map, and use this to write each set of basic differential forms \( \{du^1, du^2\} \) as a linear combination of each other set \( \{dv^1, dv^2\} \). In particular sketch these basic co-tangent vector fields using arrows .... Compare to the pictures for the basic tangent vector fields \( \{\frac{\partial}{\partial x^2}, \frac{\partial}{\partial y^2}\} \) from exercise 3.17.

**Exercise 3.30** Consider the imbedded sphere \( S^2 \subseteq \mathbb{R}^3 \) (i.e. \( M = S^2, N = \mathbb{R}^3 \) and the inclusion map \( \Phi = i : S^2 \hookrightarrow \mathbb{R}^3 \) and the standard spherical coordinates \( (u, U) = ((\theta, \phi), U) \), e.g. with \( U = (\theta, \phi)^{-1}(-\pi, \pi) \times (0, \pi) \) on \( M \) and the Cartesian coordinates \( (v, V) = ((x^1, x^2, x^3), \mathbb{R}^3) \) on \( N \).

Explicitly calculate the pullbacks \( i^*dv^i \) for \( i = 1, 2, 3 \) in terms of \( du^1 = d\theta \) and \( du^2 = d\phi \).

Locate all points \( p \in M \) where any of these cotangent vector fields vanish. Describe the vector fields pictorially, both as arrows on the sphere, and as arrows on \( (-\pi, \pi) \times (0, \pi) \) (technically, this means sketching the vector fields \( (u^{-1})^* \circ i^*dv^3 \)).
In a subsequent section we will return to differential forms to investigate when a differential form \( \omega \) is the differential of a smooth function. which will lead to powerful integrability theorems. The reader is encouraged to continue comparing and contrasting the algebraic ease of working with differential form and the more tangible, visual aspects of tangent vector fields.