Differential Geometry

Based on lectures by Pengzi Miao

Edited/Typeset by
Menaka Lashitha Bandara

School of Mathematical Sciences,
Monash University, Clayton,
Victoria, Australia, 3168

11 July 2007

Contents

1 Preliminaries and Submanifolds

1.1 Einstein's Summation Notation

The theory of Differential Geometry rapidly becomes a beast involving summations and a large number of indices. For this reason, we assume the Einstein's summation convention.

Definition 1.1 (Einstein's Summation Convention) *Whenever we have a raised index multiplied with a lowered index, we are always taking a summation over the appropriate range. That is, we define:*

$$v^i u_i = \sum_{i=1}^{n} v^i u_i$$

Throughout this document, we also assume that a raised index appearing in the denominator is treated as a lowered index. That is, exactly:

$$v^i \frac{\partial}{\partial x^i} = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i}$$

The significance of this will become apparent later when we introduce appropriate notation for basis tangent vectors.

1.2 Submanifolds

Our discussion of Differential Geometry begins with the inverse function theorem. Our setting is $\mathbb{R}^n$. Recall that:

Theorem 1.2 (Inverse Function Theorem) *Let U open in $\mathbb{R}^n$, and let $F : U \to \mathbb{R}^n$, where $F = F(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n))$. Further, suppose that $F \in C^p(\mathbb{R}^n)$ for $p > 0$ and that the Jacobian $\mathrm{D}F$ at $q \in U$ is invertible. Then, there exists a neighbourhood V of q with:*

1. $F(V) = \mathrm{Img}\, F|_V$ *open in* $\mathbb{R}^n$
2. $F|_V : V \to F(V)$ *a* C^p *diffeomorphism*

We define the notion of a submanifold in $\mathbb{R}^{n+k}$:

Definition 1.3 (n-dimensional submanifold) *We say that M is an n-dimensional C^p-submanifold of $\mathbb{R}^{n+k}$ if for all $q \in M$, there exists a U open in $\mathbb{R}^n$ and a C^p mapping $\phi : U \to \mathbb{R}^{n+k}$ satisfying:*

1. $\mathrm{Img}\, \phi = \Omega \cap M$ *for some Ω open in $\mathbb{R}^{n+k}$, with $q \in \mathrm{Img}\, \phi$.*
2. *ϕ is injective*
3. *The matrix:*

$$\mathrm{D}\phi = \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1} & \cdots & \frac{\partial \phi_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_{n+k}}{\partial x_1} & \cdots & \frac{\partial \phi_{n+k}}{\partial x_n} \end{bmatrix}$$

   has rank n.

We call such a ϕ a local parametrisation*.*

Remark The last condition is equivalent to the condition that any $n \times n$ submatrix $D'\phi$ of $D\phi$ has $\det D'\phi \neq 0$.

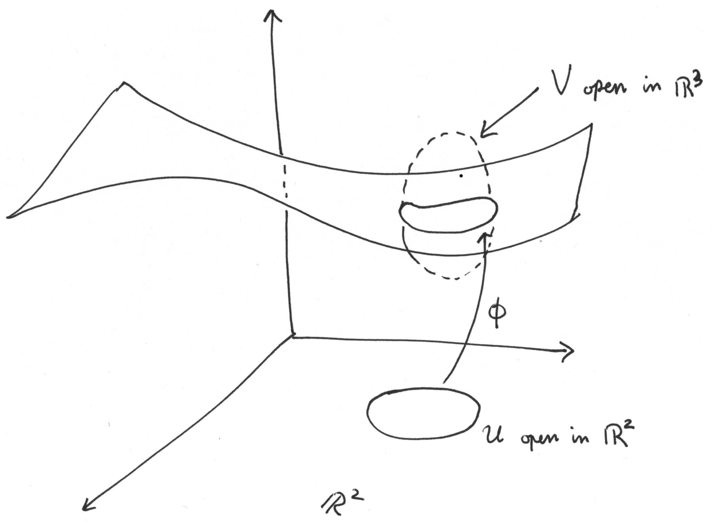

Figure 1: Manifold of dimension 2 in 3 space ($n = 2, k = 1$).

The following result highlights an important property of submanifolds.

Lemma 1.4 *An n-dimensional submanifold $M \subseteq \mathbb{R}^{n+k}$ locally looks like an n-dimensional graph over some n-dimensional coordinate plane.*

Proof Fix $x \in M$ and fix a local-parametrisation $\phi : U \to \mathbb{R}^{n+k}$ around x. By the last condition of our definition, without loss of generality, there exists an $n \times n$ submatrix $\mathrm{D}\phi'(x)$ of $\mathrm{D}\phi(x)$ such that $\det \mathrm{D}\phi'(x) \neq 0$. In fact, we can take:

$$\mathrm{D}\phi'(x) = \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1} & \cdots & \frac{\partial \phi_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_n}{\partial x_1} & \cdots & \frac{\partial \phi_n}{\partial x_n} \end{bmatrix}$$

Consider the map $F : U \to \mathbb{R}^n$ defined by:

$$F(x) = (\phi_1(x), \ldots, \phi_n(x))$$

In light of the Inverse function theorem, F is a C^p-diffeomorphism on some neighbourhood V of x. □

Corollary 1.5 *Suppose $\phi : U \to \mathbb{R}^{n+k}$ and $\psi : V \to \mathbb{R}^{n+k}$ are local parametrisation of a C^p submanifold M around x. Write $\phi(U) \cap \psi(V) = W \neq \emptyset$. Then, $\phi^{-1} \circ \psi : \psi^{-1}(W) \to \phi^{-1}(W)$ is:*

1. *Injective*
2. *C^p-differentiable.*

Proof The first claim is trivial.

To prove the second, fix $x \in \psi^{-1}(W)$. We have $F : U \to \Omega$, where $\Omega = \mathrm{Proj}\,(W)$. By the Lemma, we can assume F is a C^p-diffeomorphism. So, we can write:

$$\phi^{-1} \circ \psi = F^{-1} \circ (\pi \circ \psi)$$

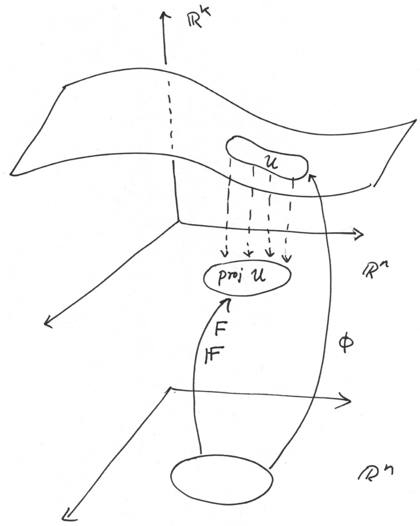

Figure 2: A submanifold locally looks like a graph over some function.

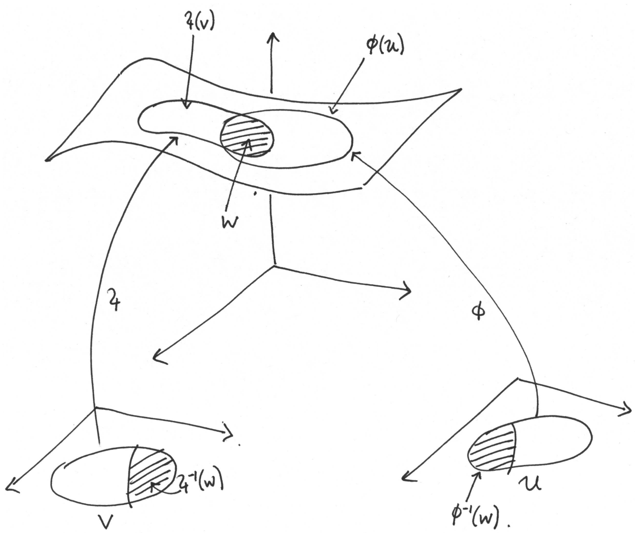

Figure 3: Two local parametrisations with a nonempty image.

where $\pi(x_1, \ldots, x_{n+k}) = (x_1, \ldots, x_n)$.

Also, we have that $F^{-1} \in C^p$. But $\mathrm{D}(\pi \circ \psi)(x)$ is invertible since $\mathrm{D}\psi$ has rank n. So, $\pi \circ \psi \in C^p$, This means exactly that $F^{-1} \circ (\pi \circ \psi)$ is C^p. □

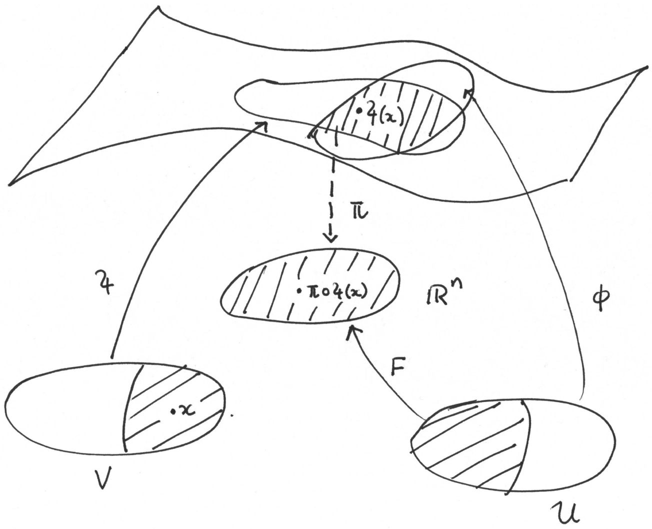

Figure 4: Two parametrisations about a point.

2 Differentiable Structures and Abstract Manifolds

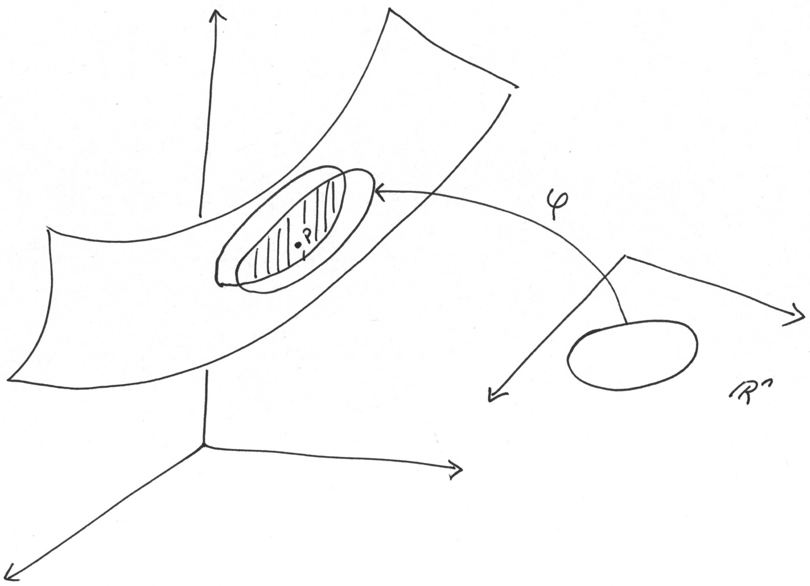

Figure 5: An n-submanifold.

There are different ways of defining abstract manifolds. The one approach is the topological approach, where we define a manifold to be a topological space that is locally diffeomorphic to Euclidean space.

We will take an alternative approach where we start considering an abstract set and induce a topology on it. It's worth noting that the two approaches are equivalent.

Definition 2.1 (Differentiable Structure) *A C^r $(r > 0)$ differentiable structure of dimension n on a set M is a collection of injective maps:*

$$\mathscr{D} = \{\phi_\alpha : U_\alpha \to M\}$$

which each U_α open in $\mathbb{R}^n$ such that the following conditions hold:

1. $M = \bigcup_\alpha \operatorname{Img} \phi_\alpha$
2. *For any pair α, β with $W = \operatorname{Img} \phi_\alpha \cap \operatorname{Img} \phi_\beta \neq \emptyset$,*
   - $\phi_\alpha^{-1}(W)$ *and* $\phi_\beta^{-1}(W)$ *are open in* $\mathbb{R}^n$
   - $\phi_\alpha^{-1} \circ \phi_\beta : \phi_\beta^{-1}(W) \to \phi_\alpha^{-1}(W)$ *is a C^r diffeomorphism.*

Remark It's important to note that M is simply a set. It has no topological structure (yet!). This is the reason we cannot talk about differentiability or continuity of the maps ϕ_α. So, we do the next best thing - talk about the differentiability of $\phi_\alpha^{-1} \circ \phi_\beta$ since this is a map between subsets of n dimensional Euclidean spaces.

To guarentee the uniqueness of the objects which we shall later define, we need to talk about maximal differentiable structures.

Definition 2.2 (Compatible) *Let $\mathscr{D}$ be a C^r differentiable structure on a set M. Let $\phi : U \to M$ with U open in $\mathbb{R}^n$ satisfying:*

1. *ϕ is an injection*

2. *If $W = \operatorname{Img}\psi \cap \operatorname{Img}\phi \neq \varnothing$ for $\psi \in \mathscr{D}$, then:*
   - *$\phi^{-1}(W)$ and $\psi^{-1}(W)$ are open in $\mathbb{R}^n$*
   - *$\phi^{-1} \circ \psi : \psi^{-1}(W) \to \phi^{-1}(W)$ is a C^r diffeomorphism.*

Then we say that ϕ is compatible with $\mathscr{D}$.

Definition 2.3 (Maximal Differentiable Structure) *A C^r differentiable structure $\mathscr{D}$ on M is said to be maximal if $\mathscr{D}$ for every map ϕ compatible with $\mathscr{D}$ we have that $\phi \in \mathscr{D}$.*

Remark It is not enough to simply ask about the containment of extensions for any given compatible ψ. Both the map and the domain are important.

We can now define our abstract manifolds:

Definition 2.4 (C^r differentaible n-Manifold) *The set M together with a C^r differential structure $\mathscr{D}$ on M is called an n-dimensional C^r differentiable manifold.*

Lemma 2.5 *Let $\mathscr{D}$ be a C^r differentiable structure on M. Define:*

$$\mathscr{F} = \{\phi : U \to M : \phi \text{ compatible with } \mathscr{D}\}$$

Then, $\mathscr{F}$ is the unique maximal differential structure containing $\mathscr{D}$.

The proof of this is left as an exercise.

Previously, it was mentioned that we discuss maximal differentiable structures to address a uniqueness issue at hand. Suppose that we have $(M, \mathscr{D})$, and $(M, \mathscr{D}')$, where every map in $\mathscr{D}$ is compatible with $\mathscr{D}'$. Intuitively, these two manifolds should be the same geometric object. We can characterise this with the use of this lemma, we can take this unique maximal extension. Then, the extension of $\mathscr{D}$ and $\mathscr{D}'$ will be the same structure $\mathscr{F}$.

This result is important for another reason. The uniqueness of the maximal extension tells us that we can often omit "maximal."

It is also note noting that we relax the notation and simply call the set M the manifold. In most cases, this suffices because we are only concerned with a single differentiable structure on M.

One immediate consequence of this result allows us to formulate the following important definitions.

Definition 2.6 (Local Parametrisation) *Let $\mathscr{D}$ be a maximal differentiable structure on a manifold M. Then $\phi \in \mathscr{D}$ is called a local parametrisation around $p\operatorname{Img}\phi$.*

Definition 2.7 (Local Coordinate Chart) *A map $\phi^{-1} : U \to \mathbb{R}^n$ with $\phi \in \mathscr{D}$ with $\mathscr{D}$ a maximal differentiable structure on M is called a local coordinate chart of any point in U.*

We give some important examples of abstract manifolds.

Example Submanifolds of $\mathbb{R}^n$.

Example Let $\mathbb{RP}^n$ = {straight lines in $\mathbb{R}^n$ through 0}. Claim: $\mathbb{RP}^n$ can be made into a diff. manifold (C^r diff., whats r?).

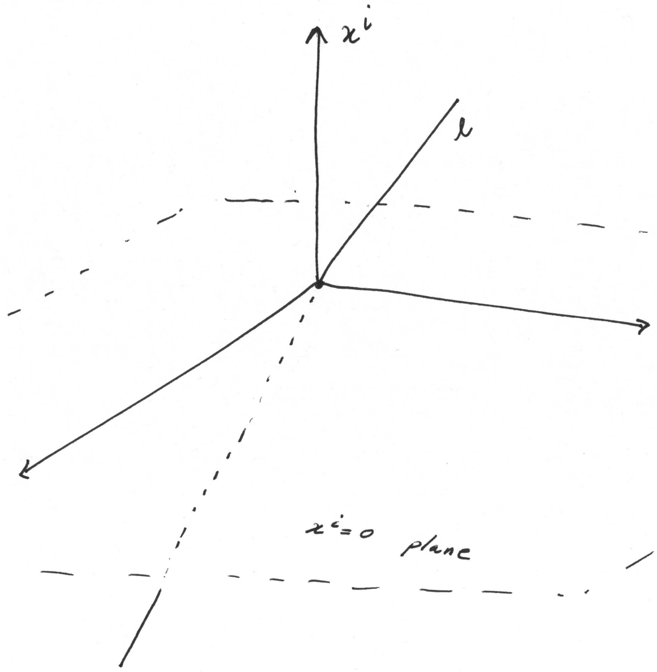

Figure 6: A line not contained in the plane $x^i = 0$.

Consider $V_i = \{l \in \mathbb{RP}^n : l$ is not contained in the coord plane $x_i = 0\}$. See Figure 2.

Trivially $\bigcup_i V_i = \mathbb{RP}^n$.

Define $\phi_i' : V_i \to \mathbb{R}^{n-1}$ given by picking some point $x = (x_1, \ldots, x_n) \in l$ and $\phi_i'(l) = (\frac{x_1}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_n}{x_i})$. Note that we "remove" x_i because if we divide through by x_i, we have made the coordinate i equal to 1.

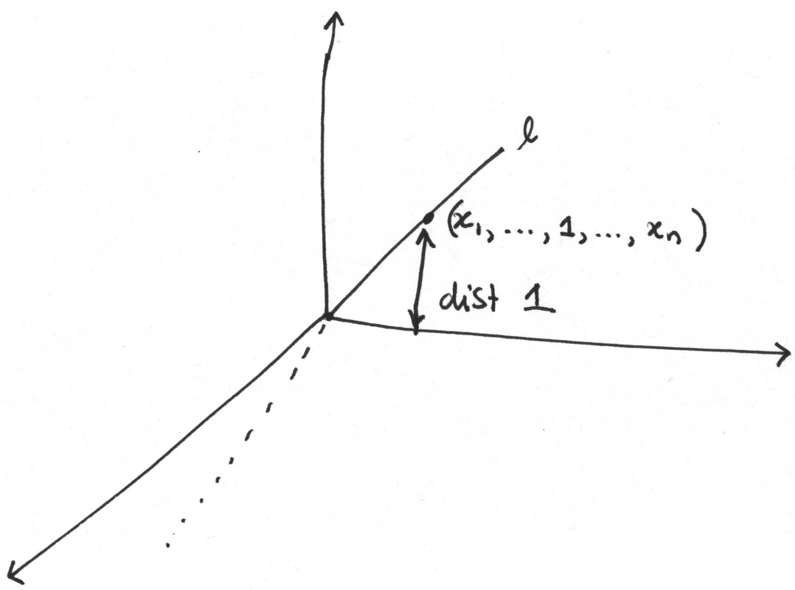

Figure 7: Line under parametrisation ϕ.

To suit our definition, we consider the function $\phi : \mathbb{R}^{n-1} \to V_i$ that takes a point $(x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1}$ to the line passing through the point $(\frac{x_1}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, 1, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_{n-1}}{x_i})$. See Figure 2

Trivially, ϕ is a bijection, and $\mathbb{RP}^n = \bigcup_i V_i = \bigcup_i \mathrm{Img}\, \phi_i$.

3 Vectors, Tangents, Bundles

3.1 The Tangent Space and Tangent Vectors

Tangent vectors are natural in the submanifold setting since we can think of a tangent vector to n surface as simply an $n+k$ vector. However, in the abstract setting, we do not have the privilege of an *ambient* space, so we need to talk in terms of curves.

From here on, unless otherwise stated, we will assume that M is an n-manifold.

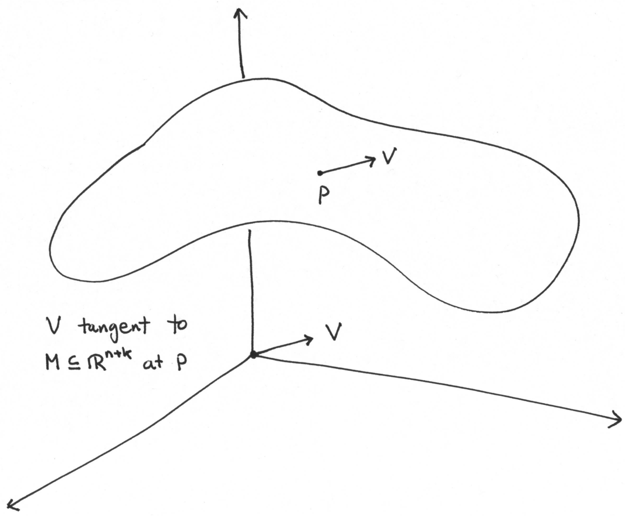

Figure 8: Tangent vector to p in the submanifold setting

Definition 3.1 (Differentiable Curve) *A differentiable curve in M is simply a map $\gamma : I \to M$ where $I \subseteq \mathbb{R}$ is an interval such that for every $t \in \operatorname{Int} I$ and for some ϕ a coordinate chart around $\gamma(t)$, the map $\phi \circ \gamma$ is differentiable.*

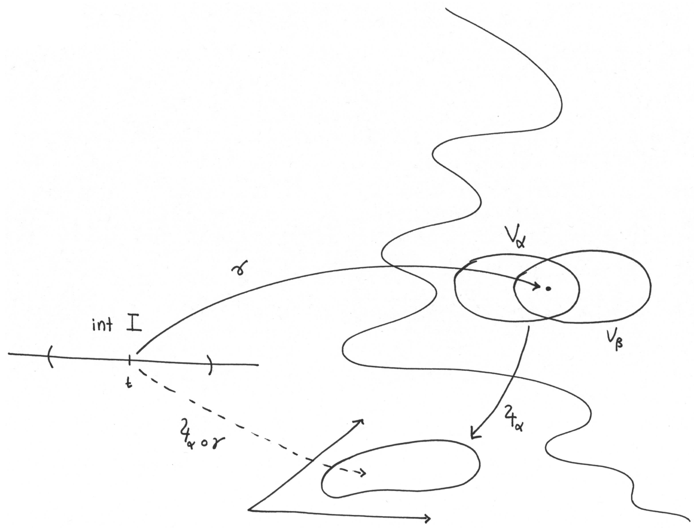

Figure 9: Differentiability of a curve γ

Remark As with all definitions in Differential Geometry, it is important to check that the definition is independent a particular of coordinate chart. That is, if ψ is another coordinate chart around $\gamma(t)$ for some $t \in \operatorname{Int} I$, then $\psi \circ \gamma$ is differentiable at t. This is left as an exercise to the reader.

Now, fix a point $p \in M$. Now, define:

$$\mathscr{M}_p = \{\gamma : (-\varepsilon, \varepsilon) \to M : \gamma(0) = p, \gamma \text{ differentiable at } 0\}$$

Remark The ε may depend on the particular γ - it is not important. The important thing is that γ is differentiable at 0 (that is, at point p on the manifold).

Let $\gamma, \sigma \in \mathscr{M}_p$. Take a coordinate chart ϕ around p. We write $\gamma \sim \sigma$ if:

$$\frac{d}{dt}|_{t=0}\ \phi \circ \gamma = \frac{d}{dt}|_{t=0}\ \phi \circ \sigma$$

Lemma 3.2 *$\sim$ is an equivalence relation*

Proof Trivial. □

Since we have an equivalence relation, we can talk about quotienting by it. Firstly, let:

$$[\sigma] = \left\{\gamma \in \mathscr{M}_p : \sigma \sim \gamma\right\}$$

Then write:

$$\mathscr{M}_p/\!\sim\, = \left\{[\sigma] : \sigma \in \mathscr{M}_p\right\}$$

For any coordinate chart ψ_α around p, we define an important map $A_\alpha : \mathscr{M}/\!\sim\, \to \mathbb{R}^n$ given by:

$$A_\alpha[\sigma] = \frac{d}{dt}|_{t=0}\ (\phi_\alpha \circ \sigma)$$

This map is key in understanding tangent spaces of abstract manifolds. The next few results will illustrate some properties of this map.

Lemma 3.3 *The map A_α is well defined.*

The proof of this is left as an exercise. The following results warrants greater attention.

Lemma 3.4 *The map A_α is a bijection.*

Proof To prove surjectivity, fix $w \in \mathbb{R}^n$. Consider the straight line given by $S_p(t) = tw + \psi_\alpha(p)$. Then, it follows that $S'_p(0) = w$ and $S_p(0) = \psi_\alpha(p)$.

Then, $(\phi_\alpha^{-1} \circ S_p)(t)$ is a curve on M, and by construction, we trivially have that $(\phi_\alpha^{-1} \circ S_p) \in \mathscr{M}_p$.

Now for injectivity:

$$\begin{aligned}
&A_\alpha[\sigma] = A_\alpha[\gamma] \\
&\iff \frac{d}{dt}|_{t=0}\ (\psi_\alpha \circ \sigma) = \frac{d}{dt}|_{t=0}\ (\psi_\alpha \circ \gamma) \\
&\iff \sigma \sim \gamma \\
&\iff [\sigma] = [\gamma]
\end{aligned}$$

□

The fact that $\mathbb{R}^n$ is a vector space means that by the previous result, A_α induces a linear structure on $\mathscr{M}/\sim$.

We define:

$$[\sigma] + [\gamma] = A_\alpha^{-1}(A_\alpha[\sigma] + A_\alpha[\gamma])$$
$$k[\sigma] = A_\alpha^{-1}(kA_\alpha[\sigma]), k \in \mathbb{R}$$

This makes $\mathscr{M}/\sim$ an n-dimensional vector space over $\mathbb{R}$. But we need the following result to confirm that our construction is indeed meaningful geometrically.

Lemma 3.5 *The linear structure defined on $\mathscr{M}_p/\sim$ is independent of the choice of coordinate chart.*

Proof We check that $A_\beta^{-1}(A_\beta[\sigma] + A_\beta[\gamma]) = A_\alpha^{-1}(A_\alpha[\sigma] + A_\alpha[\gamma])$. We leave it up to the reader to check that $A_\beta^{-1}(kA_\beta[\sigma]) = A_\alpha^{-1}(kA_\alpha[\sigma])$.

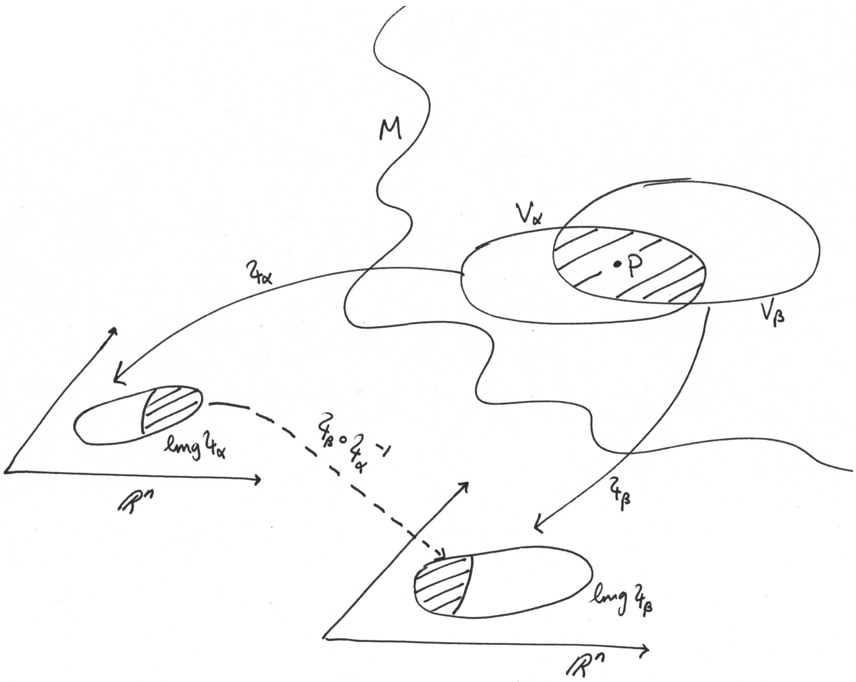

Figure 10: Coordinate transition at p

We note that locally, $\psi_\beta = (\psi_\beta \circ \psi_\alpha^{-1}) \circ (\psi_\alpha \circ \sigma)$. It follows then that:

$$\frac{d}{dt}|_{t=0} (\psi_\beta \circ \sigma) = \mathrm{D}(\psi_\beta \circ \psi_\alpha^{-1})(p)\frac{d}{dt}|_{t=0} (\psi_\alpha \circ \sigma)$$
$$\implies A_\beta[\sigma] + A_\beta[\gamma] = \mathrm{D}(\psi_\beta \circ \psi_\alpha^{-1})(p)(A_\alpha[\sigma] + A_\alpha[\gamma])$$

So, we only need to show that for any $w \in \mathbb{R}^n$, $A_\beta \circ A_\alpha^{-1}(w) = \mathrm{D}(\psi_\beta \circ \psi_\alpha^{-1})(\psi_\alpha(p))w$.

So, let $\hat{\sigma} : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ be a curve that generates w. It follows that $A_\alpha^{-1}(w) = [\psi_\alpha^{-1} \circ \hat{\sigma}]$ and:

$$A_\beta[A_\alpha^{-1} \circ \hat{\sigma}] = \frac{d}{dt}|_{t=0} (\phi_\beta \circ \phi_\alpha \circ \hat{\sigma})$$
$$= \mathrm{D}(\psi_\beta \circ \psi_\alpha^{-1})(\psi_\alpha(p))\frac{d}{dt}|_{t=0} \hat{\sigma}$$
$$= \mathrm{D}(\psi_\beta \circ \psi_\alpha^{-1})(\psi_\alpha(p))w$$

□

For clarity, we list our construction hitherto:

1. $\mathscr{M}_p$ = {curves passing through p, differentiable at p}
2. $\mathscr{M}_p/\sim = \left\{[\sigma] : \sigma \in \mathscr{M}_p\right\}$.
3. ψ_α coordinate chart around p, then A_α a bijection with $\mathbb{R}^n$.
4. We've used ψ_α to pull back the linear structure of $\mathbb{R}^n$ to $\mathscr{M}_p/\sim$.
5. We've shown that $(\mathscr{M}_p/\sim, +)$ is an n dimensional vector space over $\mathbb{R}$ independent of the choice of coordinate chart.

This suggests:

Definition 3.6 (Tangent Space, Tangent Vector) *We define the tangent space to M at p denoted by $\mathrm{T}_p M$:*

$$\mathrm{T}_p M = \mathscr{M}_p/\sim$$

Each tangent vector at p is then given by $[\sigma] \in \mathrm{T}_p M$. We also use the notation $[\sigma] = \frac{d}{dt}|_{t=0}\,\sigma = \sigma'(0)$ to denote a tangent vector.

Remark The notation we introduce in the definition is motivated by the fact that in $\mathbb{R}^n$, since we consider the tangent space at each point $\mathbb{R}^n$ to be a copy of $\mathbb{R}^n$ itself, we can write for $[\sigma] \in \mathrm{T}_p\mathbb{R}^n$ as:

$$[\sigma] = \frac{d}{dt}|_{t=0}\,\sigma$$

Remark Given a curve, we shall write $[\sigma(t+s)]$ to denote the tangent vector at $\sigma(s)$. Also, we shall simply write $\sigma'(t)$ for the tangent vector at $\sigma(t)$.

We now consider finding a basis for $\mathrm{T}_p M$. By the properties of A_α, we should be able to "pull-back" the standard basis for $\mathbb{R}^n$ into $\mathrm{T}_p M$.

We define the coordinate curve in direction i as $X^i(t) = te_i + \psi_\alpha(p)$. From the properties of A_α, it follows that $A_\alpha^{-1}(e_i) = [\psi_\alpha^{-1} \circ X^i]$ and:

$$\operatorname{span}\left\{[\psi_\alpha^{-1} \circ X^i] : 1 \le i \le n\right\} = \mathrm{T}_p M$$

We define an important notation that we will use frequently:

Definition 3.7 (Basis Vectors of $\mathrm{T}_p M$) *Given a coordinate chart ψ_α with coordinates $\{x_i\}$, we write:*

$$\frac{\partial}{\partial x^i} = [\psi_\alpha^{-1} \circ X^i]$$

Remark It is not a coincidence that our choice of notation here coincides with that of the familiar $\frac{\partial}{\partial x}$ of calculus. We will later establish the connection between the two.

Remark It is worth stressing at this point that when we write $\frac{\partial x^i}{\partial y^j}$, are really writing:

$$\psi_\alpha \circ \psi_\beta^{-1}(y) = (x^1(y), \ldots, x^n(y))$$

where $y = (y^1, \ldots, y^n)$. This is the usual calculus operator in $\mathbb{R}^n$, and it is different from $\frac{\partial}{\partial x^i} \in \mathrm{T}_p M$.

Now, suppose we have another coordinate choice around p, say $\left\{y^i\right\}$ given by ϕ_β. Now, we have two basis:

$$\frac{\partial}{\partial x^i} = [\psi_\alpha^{-1} \circ X^i]$$
$$\frac{\partial}{\partial y^i} = [\psi_\beta^{-1} \circ Y^i]$$

where $Y^i(t) = te_i + \psi_\beta(p)$.

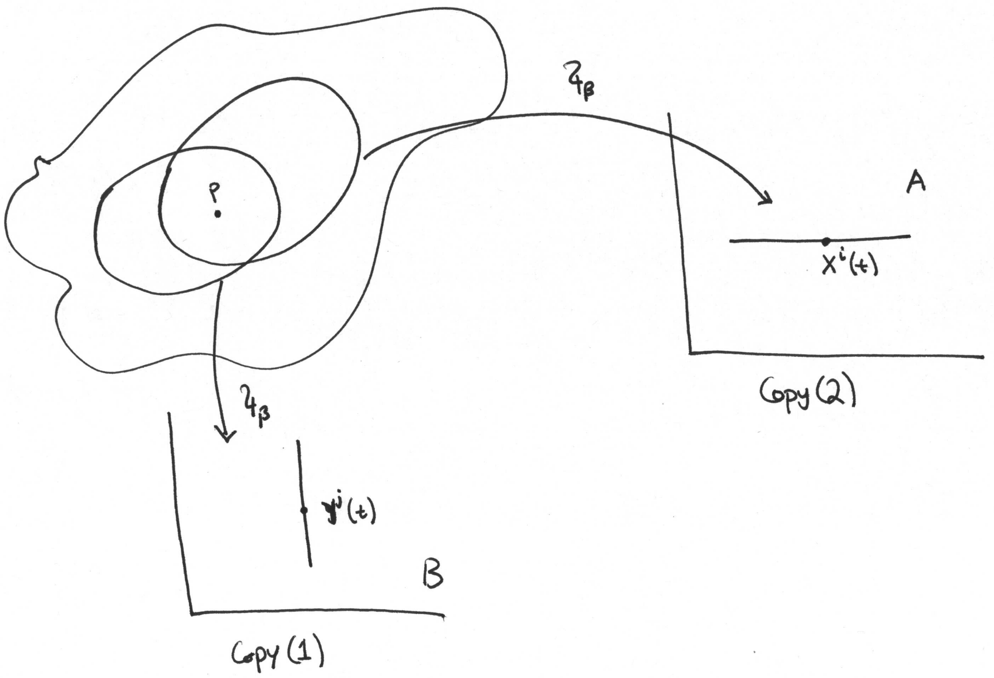

Figure 11: Coordinate Curves in two different charts

If we push the y^j into the second copy of $\mathbb{R}^n$ via ϕ_α, there's no reason to believe that it is a coordinate curve. In fact, in general, it isn't. Instead, we have the following important result:

Theorem 3.8

$$\frac{\partial}{\partial y^j} = \frac{\partial x^i}{\partial y^j}\frac{\partial}{\partial x^i}$$

Proof We note that:

$$\frac{d}{dt}|_{t=0}\,(\psi_\alpha \circ \psi_\beta \circ Y^j) = \mathrm{D}(\psi_\alpha \circ \psi_\beta^{-1})(\psi_\beta(p))\frac{d}{dt}|_{t=0}\,(te_i + \psi_\beta(p)) = \mathrm{D}(\psi_\alpha \circ \psi_\beta^{-1})(\psi_\beta(p))e_j$$

We write: $\psi_\alpha \circ \psi_\beta^{-1}(y) = (x^1(y), \ldots, x^n(y))$ and it follows that:

$$\mathrm{D}(\psi_\alpha \circ \psi_\beta^{-1}) = \begin{bmatrix} \frac{\partial x^1}{\partial y^1} & \cdots & \frac{\partial x^1}{\partial y^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x^n}{\partial y^1} & \cdots & \frac{\partial x^n}{\partial y^n} \end{bmatrix}$$

and so:

$$\mathrm{D}(\psi_\alpha \circ \psi_\beta^{-1})(\psi_\beta(p))e_j = (\frac{\partial x^1}{\partial y^j}(\psi_\beta(p)), \ldots, \frac{\partial x^n}{\partial y^j}(\psi_\beta(p))) = \frac{\partial x^i}{\partial y^j}(\psi_\beta(p))e_i$$

We write $\frac{\partial x^i}{\partial y^j}(\psi_\beta(p)) = \frac{\partial x^i}{\partial y^j}$ since in the context, we're working in local coordinates around the point p. Also, observe that $A_\alpha^{-1}(e_i) = [\phi_\alpha \circ X^i] = \frac{\partial}{\partial x^i}$, so by this and the linearity of A_α, it follows that:

$$\frac{\partial}{\partial y^j} = A_\alpha^{-1}(A_\alpha(\frac{\partial}{\partial y^j})) = A_\alpha^{-1}\left(\frac{d}{dt}|_{t=0}\,(\psi_\alpha \circ \psi_\beta^{-1} \circ Y^j)\right) = A_\alpha^{-1}\left(\frac{\partial x^i}{\partial y^j}e_i\right) = \frac{\partial x^i}{\partial y^j}A_\alpha^{-1}(e_i) = \frac{\partial x^i}{\partial y^j}\frac{\partial}{\partial x^i}$$

□

You will observe that tangent vectors change coordinates very close to the chain rule. In fact, we have used the chain rule in $\mathbb{R}^n$ in the proof when we introduced the Jacobian. This is a glimpse of the connection that tangent vectors have with calculus.

3.2 The Tangent Bundle

We have so far defined the tangent space for every point on M. It is sometimes important to consider the collection of the tangent spaces.

Definition 3.9 (Tangent Bundle) *The Tangent Bundle is denoted by* $\mathrm{T}M$ *and defined as:*

$$\mathrm{T}M = \left\{(p, v) : v \in \mathrm{T}_pM, p \in M\right\}$$

Remark Sometimes, we relax our notation and consider $\mathrm{T}M = \bigcup_{p \in M} \mathrm{T}_pM$. However, we are implicitly still distinguishing our tangent spaces at different points. In particular, we allow writing $v \in \mathrm{T}M$ and $v \in \mathrm{T}_pM$ instead of $(v, p) \in \mathrm{T}M$. This simplification is used in the proof to follow.

Theorem 3.10 *If M is C^r n-manifold, then TM has a natural differentiable structure which makes it a C^{r-1} $2n$-manifold.*

Proof Firstly, we observe that there is a natural projection map $\pi : \mathrm{T}M \to M$ such that $\pi(v) = p$ for all $v \in \mathrm{T}_pM$.

Let $\mathscr{D} = \{\phi_\alpha : U_\alpha \to M\}$ be the C^r differential structure on M.

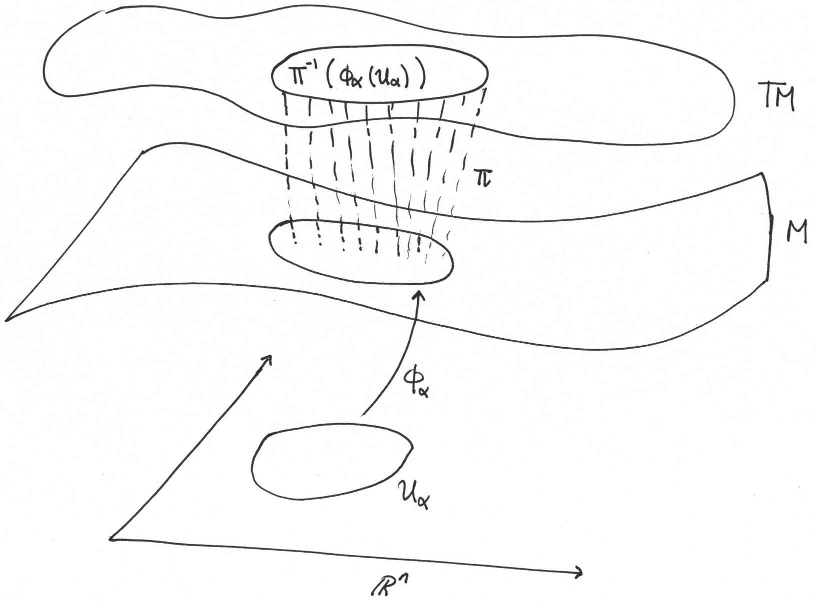

Figure 12: Canonical projection $\pi : \mathrm{T}M \to M$.

For any α, define: $\pi^{-1}(\text{Img } \phi_\alpha) = \{v \in \mathrm{T}M : \phi(v) \in \text{Img } \phi_\alpha\}$

Recall that $\phi_\alpha : U_\alpha \to M$ gives us a basis $\left\{\frac{\partial}{\partial x^{\alpha,1}}, \dots, \frac{\partial}{\partial x^{\alpha,n}}\right\}$ for all $\mathrm{T}_{\phi_\alpha(x^\alpha)}M$.

Now we define $\Phi_\alpha : U_\alpha \times \mathbb{R}^n \to \phi^{-1}(\text{Img } \phi_\alpha) \subseteq \mathrm{T}M$ by:

$$\Phi_\alpha(x^\alpha, t^\alpha) = t_i^\alpha \frac{\partial}{\partial x^{\alpha,i}}|_{\phi_\alpha(x^\alpha)}$$

Clearly, Φ_α is injective, and since U_α open in $\mathbb{R}^n$, $U_\alpha \times \mathbb{R}^n$ open in $\mathbb{R}^{2n}$. Also, we have that $\text{Img } \Phi_\alpha = \pi^{-1}(\text{Img } \phi_\alpha)$ and $\mathrm{T}M = \bigcup_\alpha \text{Img } \Phi_\alpha$.

Define:

$$\mathscr{D}' = \{\Phi_\alpha : U_\alpha \times \mathbb{R}^n \to \mathrm{T}M\}$$

We want to show that this gives a differential structure on $\mathrm{T}M$.

So, suppose that $\text{Img } \Phi_\alpha \cap \text{Img } \Phi_\beta = W \neq \emptyset$. This is if and only if $\text{Img } \phi_\alpha \cap \text{Img } \phi_\beta \neq \emptyset$ and it follows that $\text{Img } \Phi_\alpha \cap \text{Img } \Phi_\beta = \pi^{-1}(\text{Img } \phi_\alpha \cap \text{Img } \phi_\beta) \neq \emptyset$

Firstly, observe that $\Phi_\alpha^{-1}(W)$ and $\Phi_\beta^{-1}(W)$ open in $\mathbb{R}^{2n}$ since $\phi_\alpha^{-1}(\text{Img } \Phi_\alpha \cap \text{Img } \Phi_\beta)$ and $\phi_\beta^{-1}(\text{Img } \Phi_\alpha \cap \text{Img } \Phi_\beta)$ are open in $\mathbb{R}^n$.

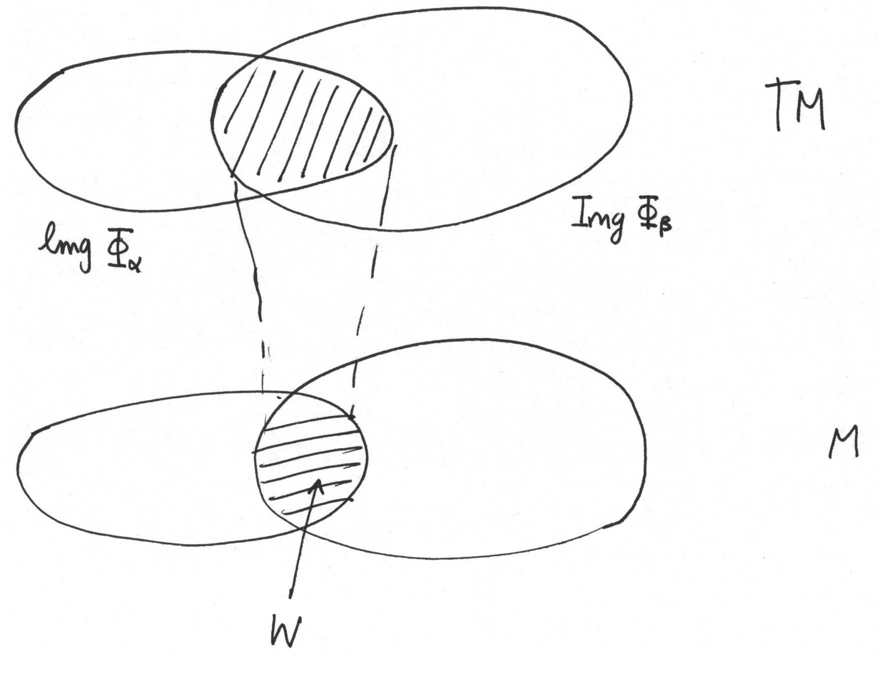

Figure 13: Intersection of $\text{Img } \Phi_\alpha$ and $\text{Img } \Phi_\beta$ projected on M.

Now, we consider the map $\Phi_\beta^{-1} \circ \Phi_\alpha : \Phi_\alpha^{-1}(W) \to \Phi_\beta^{-1}(W)$. Fix $(x^\alpha, t^\alpha) \in \Phi_\alpha^{-1}(W)$. Then,

$$\Phi_\beta^{-1} \circ \Phi_\alpha(x^\alpha, t^\alpha) = \Phi_\beta^{-1}\left(t_i^\alpha \frac{\partial}{\partial x^{\alpha,i}}|_{\phi_\alpha(x^\alpha)}\right) = (x^\beta, t^\beta)$$

So, $\phi_\alpha(x^\alpha) = \phi_\beta(x^\beta)$, and we have:

$$t_i^\beta \frac{\partial}{\partial x^{\beta,i}}|_{\phi_\beta(x^\beta)} = t_i^\alpha \frac{\partial}{\partial x^{\alpha,i}}|_{\phi_\alpha(x^\alpha)}$$

and we have $\left\{\frac{\partial}{\partial x^{\alpha,i}}\right\}$ and $\left\{\frac{\partial}{\partial x^{\beta,j}}\right\}$ a basis for $\mathrm{T}_p M$. By our change of coordinates formula:

$$t_j^\beta \frac{\partial}{\partial x^{\beta,j}}|_{\phi_\beta(x^\beta)} = t_i^\alpha \frac{\partial}{\partial x^{\alpha,i}}|_{\phi_\alpha(x^\alpha)} = t_i^\alpha \frac{\partial x^{\alpha,j}}{\partial x^{\beta,i}}|_{\phi_\beta(x^\beta)}$$

It follows then that:

$$t_j^\beta = t_i^\alpha \frac{\partial x^{\alpha,j}}{\partial x^{\beta,i}}$$

and since $\frac{\partial x^{\alpha,j}}{\partial x^{\beta,i}} \in C^{r-1}$, we have that $\mathscr{D}'$ is a C^{r-1} differentiable structure. The fact that $\mathscr{D}'$ is a $2n$-dimensional structure is trivial. □

Corollary 3.11 *M C^r n-manifold $\mapsto$ $(\mathrm{T}M)$ C^{r-1} $2n$-manifold $\mapsto$ $\mathrm{T}(\mathrm{T}M)$ C^{r-2} $4n$-manifold*

4 The Lie Derivative

4.1 Manifold Topology and Calculus

We would like to talk about concepts of continuity and differentiability of functions $f : M \to N$, where M, N are manifolds. In order to talk about these concepts, we need to topologise a manifold in some appropriate way.

Definition 4.1 (Open Set) *A subset $V \subseteq M$ is called open in M (or simply open) if for any coordinate chart $\psi : U \to \mathbb{R}^n$ with $U \cap V \neq \emptyset$, $\psi(V \cap U)$ open in $\mathbb{R}^n$.*

We leave it as an exercise to verify that this is indeed a topology on M.

Definition 4.2 (Differentiable) *Let V be open in M. Let $f : V \to \mathbb{R}$ be a function. Then, for any $p \in V$, we say that f is differentiable at p if for some coordinate chart $\psi_\alpha : V \to \mathbb{R}^n$ around p, the unique map induced by ψ_α denoted $f_\alpha = f \circ \psi_\alpha^{-1} : \psi(V \cap U) \to \mathbb{R}$ is differentiable at $q = \psi_\alpha(p)$.*

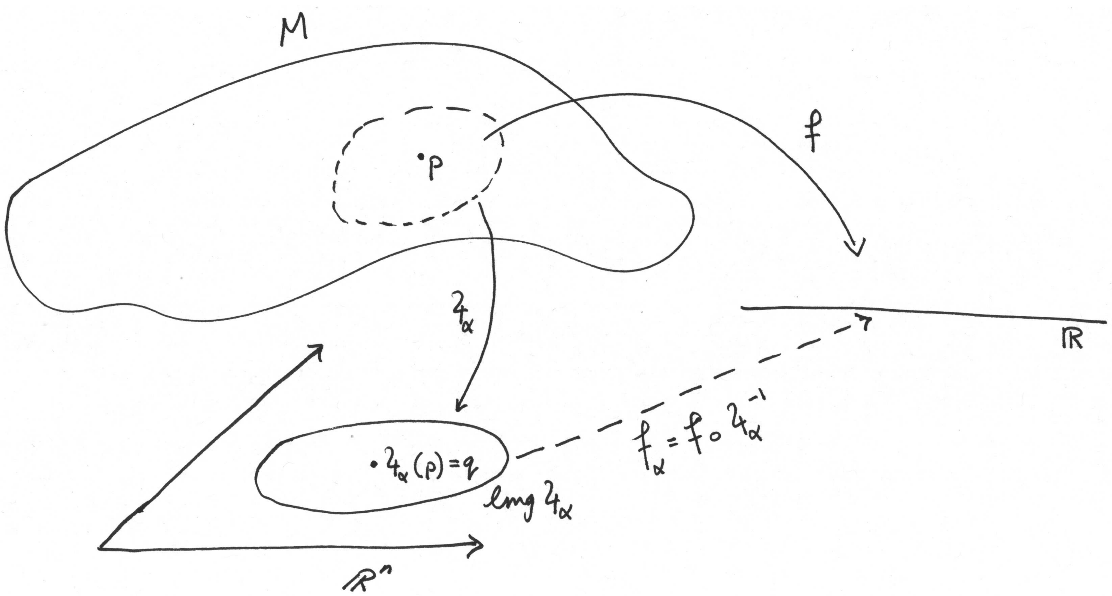

Figure 14: Characterisation of differentiability.

It is a worthwhile exercise to check that this definition holds over all coordinate charts.

We can say that $f \in C^k(V)$ if f is C^k at each point $p \in V$.

Remark If M is C^r, then we can only talk about $f : M \to \mathbb{R}$ being C^k, for $k \leq r$ since we use ψ to talk about differentiability and $\psi \in C^r$.

We now discuss the way in which tangent vectors *act* on functions on manifolds.

Definition 4.3 (Directional Derivative) *Let $v = [\gamma] \in T_pM$, and suppose that $f : V \to \mathbb{R}$ is differentiable at p. Then the directional derivative is denoted $v(f)$ and given by:*

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

Lemma 4.4 *The directional derivative is well defined.*

Proof We compute:

$$\begin{aligned}
\frac{d}{dt}|_{t=0}\ (f\circ\gamma) &= \frac{d}{dt}|_{t=0}\ (f\circ\psi_\alpha^{-1})\circ(\psi_\alpha\circ\gamma)\\
&= \frac{d}{dt}|_{t=0}\ (f_\alpha\circ\beta)\\
&\qquad \textit{(where } \beta=\psi_\alpha\circ\gamma\textit{)}\\
&= \frac{d}{dt}|_{t=0}\ f_\alpha(\beta^1(t),\ldots,\beta^n(t))\\
&= \frac{\partial f}{\partial x^i}|_q\ \frac{d\beta^i}{dt}|_{t=0}\\
&= \dot\beta^i(0)\frac{\partial f}{\partial x^i}
\end{aligned}$$

Now, for any other $\sigma\in[\gamma]$, $(\psi_\alpha\circ\sigma)'(0)=\dot\beta(0)$, and so the result follows. □

We are now prepared to reveal how we can do calculus with the basis induced by coordinate charts.

Lemma 4.5

$$\frac{\partial}{\partial x^i}(f)=\frac{\partial f_\alpha}{\partial x^i}$$

Proof Let ψ_α be the coordinate chart with coordinates $\left\{x^i\right\}$ Recall that T_pM has basis given by $\left\{\frac{\partial}{\partial x^i}=[\psi_\alpha^{-1}\circ X^i]\right\}$. Now take any $f:U\to\mathbb{R}$ differentiable at p, and we compute:

$$\frac{\partial}{\partial x^i}(f)=\frac{d}{dt}|_{t=0}\ (f\circ\psi_\alpha^{-1})\circ(\psi_\alpha\circ\psi_\alpha^{-1}\circ X^i)=\frac{d}{dt}|_{t=0}\ (f_\alpha\circ X^i)=\frac{d}{dt}|_{t=0}\ f_\alpha(q+te_i)=\frac{\partial f_\alpha}{\partial x^i}$$

□

Since this holds for all differentiable f, the basis given by a coordinate system really does act as differential operators.

Lemma 4.6 *If $v\in\mathrm{T}_pM$, and $v=v^i\frac{\partial}{\partial x^i}$ for local coordinates $\left\{x^i\right\}$, then:*

$$v(f)=(v^i\frac{\partial}{\partial x^i})(f)=v^i\frac{\partial}{\partial x^i}(f)$$

Proof We can write $v=[\gamma]$. Then,

$$\begin{aligned}
v(f) &= \frac{d}{dt}|_{t=0}\ f\circ\gamma\\
&= \frac{d}{dt}|_{t=0}\ (f\circ\psi_\alpha^{-1})\circ(\psi_\alpha\circ\gamma)\\
&= \frac{\partial f_\alpha}{\partial x^i}\dot\beta^i(0)
\end{aligned}$$

where $\dot\beta^i(0)=(\psi_\alpha\dot\circ\gamma)^i(0)$. By setting f as mapping of the component i of ψ_α, we get $v(f)=v^i$. This justifies writing $\dot\beta^i(0)=v^i$, and by combining Lemma 4.5:

$$v(f)=v^i\frac{\partial f_\alpha}{\partial x^i}=v^i\frac{\partial}{\partial x^i}(f)$$

□

Corollary 4.7 *, For $v, w \in \mathrm{T}_p M$, if $v(f) = w(f)$ for all differentiable f at p, then $v = w$.*

Proof Taking $f_\alpha = \pi_i$, the projection of the coordinate i, we get $w^i = v^i$ for all i. It follows that $v = w$. □

4.2 Vector Fields

Let V be open in M. We will assume this throughout this section unless otherwise stated.

Definition 4.8 (Vector Field) *A vector field X on V is just a correspondence that associates to each $p \in V$, an element of $\mathrm{T}_p M$. We denote the set of vectorfields in V by $\mathscr{X}(V)$.*

Remark Given a coordinate chart $\psi_\alpha : U_\alpha \to \mathbb{R}^n$, we can consider $\frac{\partial}{\partial x^i}(p) = \frac{\partial}{\partial x^i}|_p$ as a vector field inside U_α.

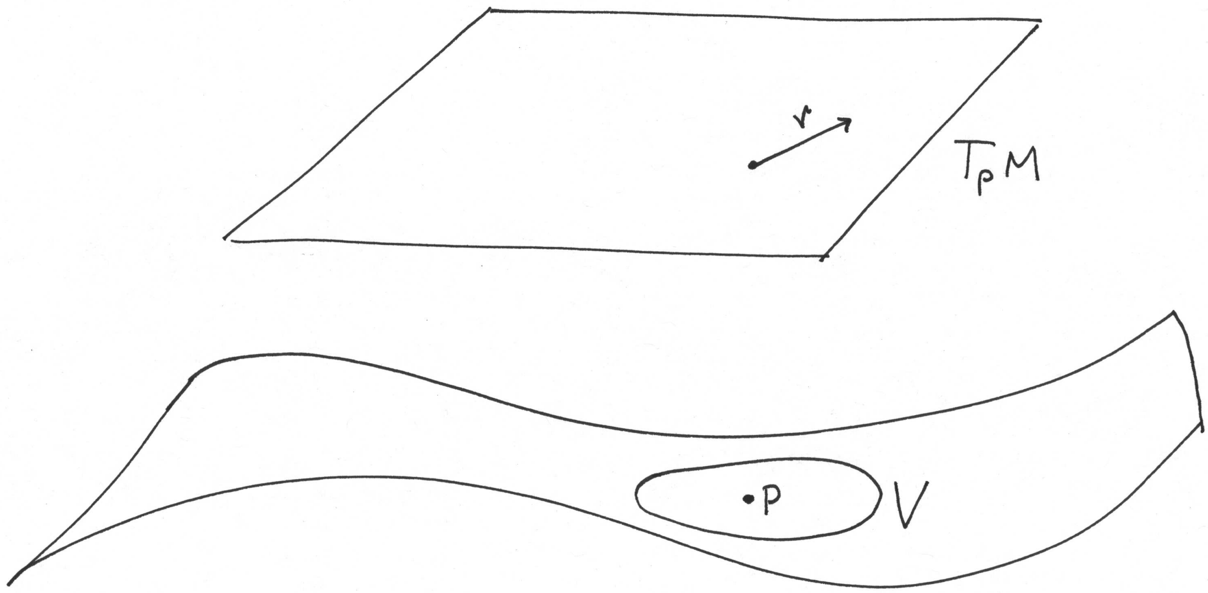

Figure 15: A vectorfield which associates p to $v \in \mathrm{T}_p M$.

We want to firstly address the notion of smoothness of a vector field. Suppose that X is a vector field in V. Then for any $p \in V$, and any coordinate chart ψ_α, we can write $X(x) = X^{\alpha,i}(x)\frac{\partial}{\partial x^i}$ (we write with the raised index α to denote the dependence on ψ_α). This inspires the following definition:

Definition 4.9 (Smoothness of Vector Field) *We say that X is differentiable at p if each $X^{\alpha,i} : V \to \mathbb{R}$ is differentiable at p. We say that $X \in C^k(V)$ if $X \in C^k$ at each point $p \in V$.*

Remark Again, it is worthwhile exercise to show that our definition holds under all choices of coordinates around p. On a C^r manifold, the highest order of derivatives for a vector field is C^{r-1} since the change of basis formula involves first order derivatives.

Theorem 4.10 *Let V open in M, and X, Y vector fields on V. Then there exists a unique vector field Z on V such that:*

$$Zf = X(Yf) - Y(Xf)$$

for all $f \in C^2(V)$.

Proof Let $f \in C^2(V)$. We fix a coordinate choice $\psi_\alpha : U_\alpha \to \mathbb{R}^n$. Then, it follows that:

$$Xf = X^{\alpha,i}\frac{\partial f_\alpha}{\partial x^i}, Yf = Y^{\alpha,i}\frac{\partial f_\alpha}{\partial x^i}$$

Also, note that $Xf : V \to \mathbb{R}^n$, where for $p \in V$:

$$Xf(p) = Y^{\alpha,i}(p)\frac{\partial f_\alpha}{\partial x^i}(\psi_\alpha(p))$$

It follows that:

$$Y(Xf) = Y^{\alpha,j}\frac{\partial}{\partial x^j}(Xf) = Y^{\alpha,j}\frac{\partial}{\partial x^j}(Xf) = Y^{\alpha,j}\frac{\partial\left(X_\alpha^{\alpha,i}\frac{\partial f_\alpha}{\partial x^i}\right)}{\partial x^j} = Y^{\alpha,j}\left(\frac{\partial X_\alpha^{\alpha,i}}{\partial x^j}\frac{\partial f_\alpha}{\partial x^i} + X_\alpha^{\alpha,i}\frac{\partial^2 f_\alpha}{\partial x^j \partial x^i}\right)$$

Similarly,

$$X(Yf) = X^{\alpha,j}\left(\frac{\partial Y_\alpha^{\alpha,i}}{\partial x^j}\frac{\partial f_\alpha}{\partial x^i} + Y_\alpha^{\alpha,i}\frac{\partial^2 f_\alpha}{\partial x^j \partial x^i}\right)$$

Note that we are not applying the summation convention to the α, only to i, j. This will be assumed throughout the rest of the proof.

Also, since $f \in C^2(V)$, we have:

$$\frac{\partial^2 f_\alpha}{\partial x^i \partial x^j} = \frac{\partial^2 f_\alpha}{\partial x^j \partial x^i}$$

and it follows that:

$$\begin{aligned}
X(Yf) - Y(Xf) &= X^{\alpha,j}\frac{\partial Y_\alpha^{\alpha,i}}{\partial x^j}\frac{\partial f_\alpha}{\partial x^i} + X^{\alpha,j}Y_\alpha^{\alpha,i}\frac{\partial^2 f_\alpha}{\partial x^j \partial x^i} - Y^{\alpha,j}\frac{\partial X_\alpha^{\alpha,i}}{\partial x^j}\frac{\partial f_\alpha}{\partial x^i} - Y^{\alpha,j}X_\alpha^{\alpha,i}\frac{\partial^2 f_\alpha}{\partial x^j \partial x^i} \\
&= X^{\alpha,j}\frac{\partial Y_\alpha^{\alpha,i}}{\partial x^j}\frac{\partial f_\alpha}{\partial x^i} - Y^{\alpha,j}\frac{\partial X_\alpha^{\alpha,i}}{\partial x^j}\frac{\partial f_\alpha}{\partial x^i} \\
&= \left(X^{\alpha,j}\frac{\partial Y_\alpha^{\alpha,i}}{\partial x^j} - Y^{\alpha,j}\frac{\partial X_\alpha^{\alpha,i}}{\partial x^j}\right)\frac{\partial f_\alpha}{\partial x^i}
\end{aligned}$$

since at p:

$$X^{\alpha,j}(p)Y_\alpha^{\alpha,i}(\psi_\alpha(p))\frac{\partial^2 f_\alpha}{\partial x^j \partial x^i}(\psi_\alpha(p)) = Y^{\alpha,j}(p)\frac{\partial X_\alpha^{\alpha,i}}{\partial x^j}(\psi_\alpha(p))\frac{\partial f_\alpha}{\partial x^i}(\psi_\alpha(p))$$

We set:

$$Z = \left(X^{\alpha,j}\frac{\partial Y_\alpha^{\alpha,i}}{\partial x^j} - Y^{\alpha,j}\frac{\partial X_\alpha^{\alpha,i}}{\partial x^j}\right)\frac{\partial}{\partial x^i}$$

Now, this expression is valid on all of $V \cap U_\alpha$. We need to show that this is valid on all of V. Let $\psi_\beta : U_\beta \to \mathbb{R}^n$ be another coordinate choice around p, and suppose that $Z'f = X(Yf) - Y(Xf)$ on $V \cap U_\beta$. Now on $U_\alpha \cap V \cap U_\beta$, we have $Zf = Z'f$ for all $f \in C^2$. By Corollary 4.7, this implies that $Z = Z'$. □

4.3 The Lie Derivative

The theorem above gives us a glimpse of one way to take a derivative of a vectorfield with respect to another. We define it formally:

Definition 4.11 (Lie Derivative) *Let $X, Y \in \mathscr{X}(V)$ for V open in M. We define $\mathscr{L}_X Y = [X, Y] = X(Y) - Y(X)$ to be the Lie Derivative of Y with respect to X. If $f \in C^2$, we define $\mathscr{L}_X(f) = X(f)$.*

Our construction above of $\mathscr{L}_X Y$ was functional. Since we expect it to be meaningful geometrically, we build more theory to construct $\mathscr{L}_X Y$ by geometric means.

Definition 4.12 (Continuity) *Let M, N be manifolds. Then, we say that a map $f : M \to N$ is continuous if for every W open in N, $f^{-1}(W)$ is open in M.*

Definition 4.13 (Differentiable) *Let $f : M \to N$, and continuous. Then, we say that f is differentiable at p if for some chart ψ_α around p and ϕ_α around $f(p)$, the map $\phi_\alpha \circ f \circ \psi_\alpha : \mathrm{Img}\ \psi_\alpha \to \mathrm{Img}\ \psi_\beta$ is differentiable at $\psi_\alpha(p)$.*

Again, we make a remark that this definition is independent of our choice of coordinate chart.

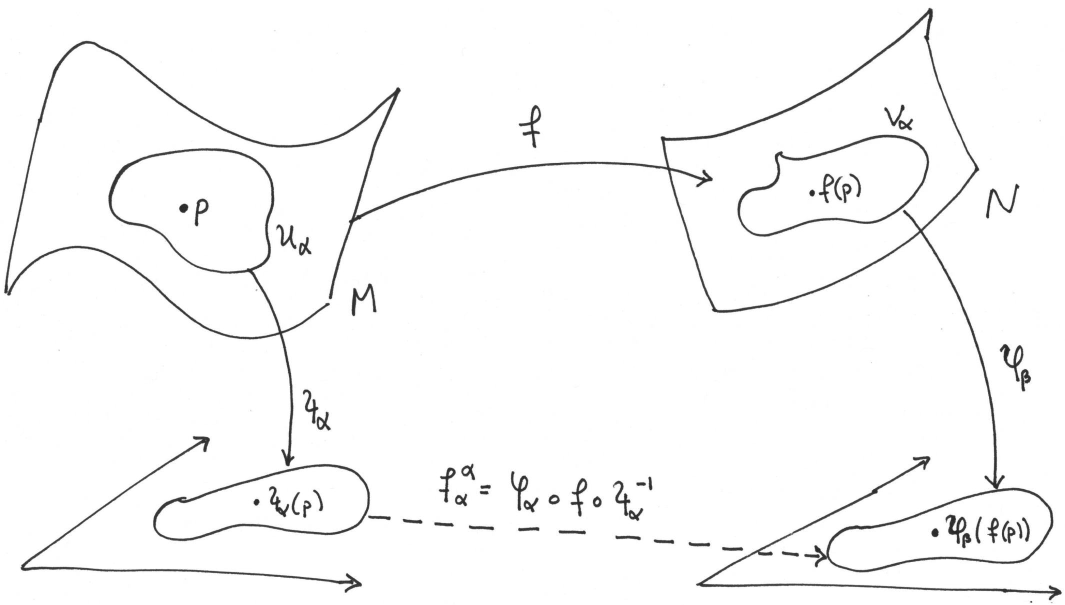

Figure 16: Differentiability of f at p.

Now, given a tangent vector $[\sigma] \in \mathrm{T}_p M$, under a differentiable function $f : M \to N$, we have $[f \circ \sigma] \in \mathrm{T}_{f(p)} N$. This is an important effect induced on the tangent space via differentiable maps.

Definition 4.14 (Differential) *For a differentiable $f : M \to N$, we define the differential $\mathrm{d}f : \mathrm{T}_p M \to \mathrm{T}_{f(p)} N$ by:*

$$\mathrm{d}f([\sigma]) = [f \circ \sigma] = (f \circ \sigma)'(0) = A_\alpha^{-1}(\frac{d}{dt}|_{t=0}\ \phi_\alpha \circ (f \circ \sigma))$$

where ϕ_α is a coordinate chart around $f(p)$, and

We give a first use of the differential. Suppose we have a curve $\sigma : (a, b) \to M$. We want to talk about $[\sigma]$ at each $s \in (a, b)$ being a tangent vector to $\sigma(s)$. What we really characterise here is a way to talk about $\frac{d}{dt}|_{t=s}$ in the manifold setting.

Lemma 4.15

$$\mathrm{d}\sigma(\frac{\partial}{\partial s}) = [\sigma(t + s)]$$

Where $\frac{\partial}{\partial s}$ is the standard basis at $\mathrm{T}_s(a, b)$.

Proof We note that $\frac{\partial}{\partial s} = [t + s]$, and it follows that:

$$\mathrm{d}\sigma(\frac{\partial}{\partial s}) = A_\alpha^{-1}\left(\frac{d}{dt}|_{t=0}\ (\psi_\alpha \circ \sigma \circ (t + s))\right) = [\sigma(t + s)]$$

□

Lemma 4.16 *The differential is linear.*

Proof We check addition. Multiplication by scalar is trivial.

Suppose $[\sigma] + [\gamma] = [\delta]$. Let $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ be a coordinate chart around p, and $\phi_\alpha : V_\alpha \to \mathbb{R}^m$ a coordinate chart around $f(p)$.

Then,

$$\begin{aligned}
A_\alpha(\mathrm{d}f([\delta])) &= \frac{d}{dt}|_{t=0}\ \phi_\alpha \circ f \circ \delta \\
&= \frac{d}{dt}|_{t=0}\ (\phi_\alpha \circ f \circ \psi_\alpha^{-1})(\psi_\alpha \circ \delta) \\
&= \mathrm{D}(\phi_\alpha \circ f \circ \psi_\alpha^{-1})\frac{d}{dt}|_{t=0}\ (\psi_\alpha \circ \delta) \\
&= \mathrm{D}(\phi_\alpha \circ f \circ \psi_\alpha^{-1})(\frac{d}{dt}|_{t=0}\ (\psi_\alpha \circ \gamma) + \frac{d}{dt}|_{t=0}\ (\psi_\alpha \circ \sigma) \\
&= A_\alpha(\mathrm{d}f([\gamma])) + A_\alpha(\mathrm{d}f([\sigma]))
\end{aligned}$$

□

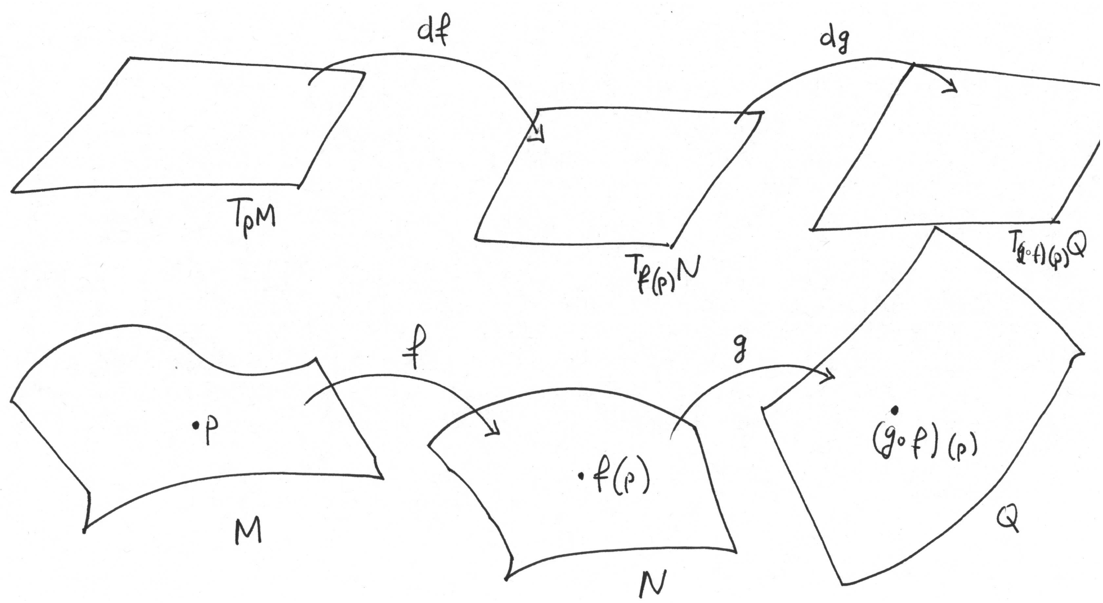

Figure 17: Chain rule on Manifolds.

Lemma 4.17 (Chain Rule on Manifolds)

$$\mathrm{d}(g \circ f) = \mathrm{d}g \circ \mathrm{d}f$$

The proof of this should be attempted as an exercise.

Definition 4.18 (C^k Diffeomorphism) *A C^k map $f : M \to N$ is called a C^k diffeomorphism if there exists a C^k map $g : N \to M$ such that $f \circ g = \mathrm{id}_N$ and $g \circ f = \mathrm{id}_M$.*

Corollary 4.19 *If f is a C^k diffeomorphism, then $\mathrm{d}f : \mathrm{T}_pM \to \mathrm{T}_{f(p)}N$ is a linear isomorphism.*

Proof Firstly, observe:

$$\mathrm{d}f \circ g = \mathrm{d}(\mathrm{id}_N) : \mathrm{T}_{f(p)}N \to \mathrm{T}_{f(p)}N$$
$$\mathrm{d}g \circ f = \mathrm{d}(\mathrm{id}_M) : \mathrm{T}_pM \to \mathrm{T}_pM$$

Also:

$$\mathrm{d}g \circ \mathrm{d}f = \mathrm{id}$$
$$\mathrm{d}f \circ \mathrm{d}g = \mathrm{id}$$

It follows that $(\mathrm{d}f)^{-1}$ exists and is given by $\mathrm{d}g$. □

We now quote an important theorem from the theory of ordinary differential equations.

Theorem 4.20 (Integral Curves in $\mathbb{R}^n$) *Let Ω open in $\mathbb{R}^n$. Suppose that $X : \Omega \to \mathbb{R}^n$ is a C^k vectorfield on Ω. That is exactly $X(x^1, \ldots, x^n) = X^i(x^1, \ldots, x^n)e_i$. Then for any $p \in \Omega$, there exists a neighbourhood U of p in Ω, a $\delta > 0$, and a function $F : U \times (-\delta, \delta) \to \mathbb{R}^n$ satisfying:*

1. *$F(x, 0) = x$ for all $x \in \Omega$*
2. *$F(U, t) \subseteq \Omega$ for all $t \in (-\delta, \delta)$ and $F(\cdot\,, t) : U \to F(U, t)$ is a C^k diffeomorphism*
3. *For any $x \in \Omega$ fixed, for the curve $\gamma(t) = F(x, t)$ we have:*

$$\frac{d}{dt}F(x,t) = (X \circ F)(x,t) \iff \frac{dF^i}{dt} = X^i(F^1(x,t), \ldots, F^n(x,t)), 1 \le i \le n$$

We say that $\left\{F_t^X\right\} = \left\{F^X(\cdot\,, t)\right\} = \left\{F^X(\cdot\,, t)\right\}_{t \in (-\delta,\delta)}$ is a 1-parameter family of diffeomorphisms generated by X.

Corollary 4.21 (Integral Curves on Manifolds) *Let Ω open in M. Suppose that $X : \Omega \to M$ is a C^k vectorfield on Ω. Then for any $p \in \Omega$, there exists a neighbourhood U of p in Ω, a $\delta > 0$, and a function $F : U \times (-\delta, \delta) \to M$ satisfying:*

1. *$F(x, 0) = x$ for all $x \in \Omega$*
2. *$F(U, t) \subseteq \Omega$ for all $t \in (-\delta, \delta)$ and $F(\cdot\,, t) : U \to F(U, t)$ is a C^k diffeomorphism*
3. *For any $x \in \Omega$ fixed, for the curve $\gamma(t) = F(x, t)$, we have*

$$\frac{d}{dt}F(x,t) = [F(x, t+s)] = (X \circ F)(x, s)$$

Proof Fix $p \in \Omega$, and let $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ be a local coordinate chart around p, with coordinates $\left\{x^i\right\}$. We write:

$$X(x) = X^i(x)\frac{\partial}{\partial x^i}$$

for $x \in U_\alpha \cap \Omega$. We define a C^k vectorfield $\tilde{X} : \psi_\alpha(U_\alpha \cap \Omega) \to \mathbb{R}^n$

$$\tilde{X}(y) = A_\alpha((X \circ \psi_\alpha^{-1})(y)) = (X^i \circ \psi_\alpha^{-1})(y)e_i$$

By the theorem, we can find a neighbourhood $\tilde{U} \subseteq \psi_\alpha(U_\alpha \cap \Omega)$ around $\psi_\alpha(p)$ a $\delta > 0$ and a function $\tilde{F} : \tilde{U} \times (-\delta, \delta) \to \mathbb{R}^n$ satisfying:

1. $\tilde{F}(y, 0) = y$
2. $\tilde{F}(\tilde{U}, t) \subseteq \psi_\alpha(U_\alpha \cap \Omega)$ and $\tilde{F}(\cdot\,, t) : U \to \tilde{F}(\tilde{U}, t)$ a C^k diffeomorphism.

3.

$$\frac{d}{dt}|_{t=0}\ \tilde{F}(y,t+s) = \tilde{X}(\tilde{F}(y,s)) \iff \frac{d\tilde{F}^i}{dt}(y,s) = (X^i \circ \psi_\alpha^{-1})(y)$$

We can now use ψ_α to pull this back onto the manifold. Define $U = \psi_\alpha^{-1}(\tilde{U})$, and $F : U \times (-\delta,\delta) \to M$ by:

$$F(x,t) = (\psi_\alpha^{-1} \circ \tilde{F})(\psi_\alpha(x),t)$$

Firstly, we have:

$$F(x,0) = (\psi_\alpha^{-1} \circ \tilde{F}(\psi_\alpha(x),0) = \psi_\alpha^{-1}\psi_\alpha(x) = x$$

Secondly,

$$F(U,t) = (\psi_\alpha^{-1} \circ \tilde{F})(\psi_\alpha(U),t) = \psi_\alpha^{-1} \circ \tilde{F}(\tilde{U},t) \subseteq \psi_\alpha^{-1}(\psi_\alpha(U_\alpha \cap \Omega)) = U_\alpha \cap \Omega$$

and $F(\cdot\,,t) : U \to F(U,t)$ is a C^k diffeomorphism since ψ_α is a diffeomorphism.

Finally, we claim that $[F(x,t+s)] = (X \circ F)(x,s)$. Compute:

$$\begin{aligned}
A_\alpha[F(x,t+s)] &= \frac{d}{dt}|_{t=0}\ (\psi_\alpha \circ F)(x,t+s) \\
&= \frac{d}{dt}|_{t=0}\ \psi_\alpha \circ \psi_\alpha^{-1} \circ \tilde{F}(\psi_\alpha(x),t+s) \\
&= \frac{d}{dt}|_{t=0}\ \tilde{F}^i(\psi_\alpha(x),t+s)e_i \\
&= \frac{d\tilde{F}^i}{dt}(\psi_\alpha(x),s)e_i \\
&= (X^i \circ \psi_\alpha^{-1}) \circ \psi_\alpha(x)e_i \\
&= X^i(x)e_i \\
&= A_\alpha(X(x))
\end{aligned}$$

□

We establish a few auxiliary results that will guide us in the geometric construction of the Lie Derivative:

Lemma 4.22 *Let X be a vectorfield on Ω open in M. Let $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ be a coordinate chart around $p \in \Omega$, and write $X = X^i \frac{\partial}{\partial x^i}$ in $U = U_\alpha \cap \Omega$. If $X(x) = [\gamma_x(t)]$, then*

$$X^i(x) = \frac{d}{dt}|_{t=0}\ (\psi_\alpha \circ \gamma_x(t))^i$$

We leave the proof of this as an exercise.

Lemma 4.23 *Let $F : M \to N$ be a diffeomorphism, and let $p \in M$. Let $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ be a coordinate chart around p with coordinates $\left\{x^i\right\}$ and $\phi_\alpha : V_\alpha \to \mathbb{R}^n$ a coordinate chart around $F(p)$ with coordinates $\left\{y^i\right\}$. Let $F_\alpha^\alpha = \phi_\alpha \circ F \circ \psi_\alpha^{-1}$, and:*

$$(b^i_j) = \left(\frac{\partial F_\alpha^{\alpha,i}}{\partial x^j}\right)^{-1}$$

Then,

$$(\mathrm{d}F)^{-1}(\frac{\partial}{\partial y^j}) = b^i_j \frac{\partial}{\partial x^i}$$

Proof By construction,

$$\delta_{ij} = b^i_k \frac{\partial F^{\alpha,k}_\alpha}{\partial x_j}$$

Also, we have:

$$(\mathrm{d}F)(\frac{\partial}{\partial x^i}) = A^{-1}_\alpha(\mathrm{D}(F^\alpha_\alpha)e_i) = \frac{\partial F^{\alpha,j}_\alpha}{\partial x^i}\frac{\partial}{\partial y^j}$$

Now,

$$\frac{\partial}{\partial x^i} = (\mathrm{d}F)^{-1} \circ (\mathrm{d}F)(\frac{\partial}{\partial x^i}) = \frac{\partial F^{\alpha,j}_\alpha}{\partial x^i}(\mathrm{d}F)^{-1}(\frac{\partial}{\partial y^j})$$

It follows that:

$$b^i_j\frac{\partial}{\partial x^i} = b^i_j\frac{\partial F^{\alpha,j}_\alpha}{\partial x^i}(\mathrm{d}F)^{-1}(\frac{\partial}{\partial y^j}) = \delta_{ii}(\mathrm{d}F)^{-1}(\frac{\partial}{\partial y^j}) = (\mathrm{d}F)^{-1}(\frac{\partial}{\partial y^j})$$

□

Lemma 4.24 *Suppose that* $(a^i_j(t))^{-1} = (b^i_j(t))$*, with each* $b^i_j \in C^1$*. Then,*

$$\frac{db^i_j}{dt} = -b^i_q\frac{da^q_p}{dt}b^p_j$$

Now we have the sufficient material to consider the Lie Derivative geometrically.

Theorem 4.25 (Geometric Construction of the Lie Derivative) *Let* X *be a* C^k $(k \geq 2)$ *vectorfield in* Ω *open in* M *a* C^r *manifold for* $r \geq 3$*. Then, at* p*:*

$$\mathscr{L}_X Y = \frac{d}{dt}|_{t=0}\ (\mathrm{d}F_t)^{-1}(Y \circ F)(p,t)$$

where $F_t = F^X_t$ *is the* 1 *parameter family of diffeomorphisms generated by* X *at* p*.*

Proof Fix a chart $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ at p, with coordinates be $\left\{x^i\right\}$. We can assume that $F_t(U) \subseteq U_\alpha$ by restricting X to U_α in applying the Integral Curves Theorem.

Firstly we write $X = X^i\frac{\partial}{\partial x^i}$ and $Y = Y^i\frac{\partial}{\partial x^i}$. Also, we use the notation:

$$\begin{aligned} F_{t,\alpha} &= F_t \circ \psi^{-1}_\alpha \\ F^\alpha_t &= \psi_\alpha \circ F_t \\ F^\alpha_{t,\alpha} = F^\alpha_\alpha(\cdot\,,t) &= \psi_\alpha \circ F_t \circ \psi^{-1}_\alpha \end{aligned}$$

Let $F_t(p) = q$. Then, by the application of Lemma 4.23,

$$(\mathrm{d}F)^{-1}(Y(q) = Y^j(q)b^i_j(q)\frac{\partial}{\partial x^i}$$

Then,

$$\frac{d}{dt}|_{t=0}\ (\mathrm{d}F)^{-1}(Y(q)) = \frac{d}{dt}|_{t=0}\ \left(Y^j(q)b^i_j(q)\right)\frac{\partial}{\partial x^i} = \left(\frac{dY^j(q)}{dt}|_{t=0}\ b^i_j(q)|_{t=0}\ + Y^j(q)|_{t=0}\ \frac{db^i_j}{dt}|_{t=0}\right)\frac{\partial}{\partial x^i}$$

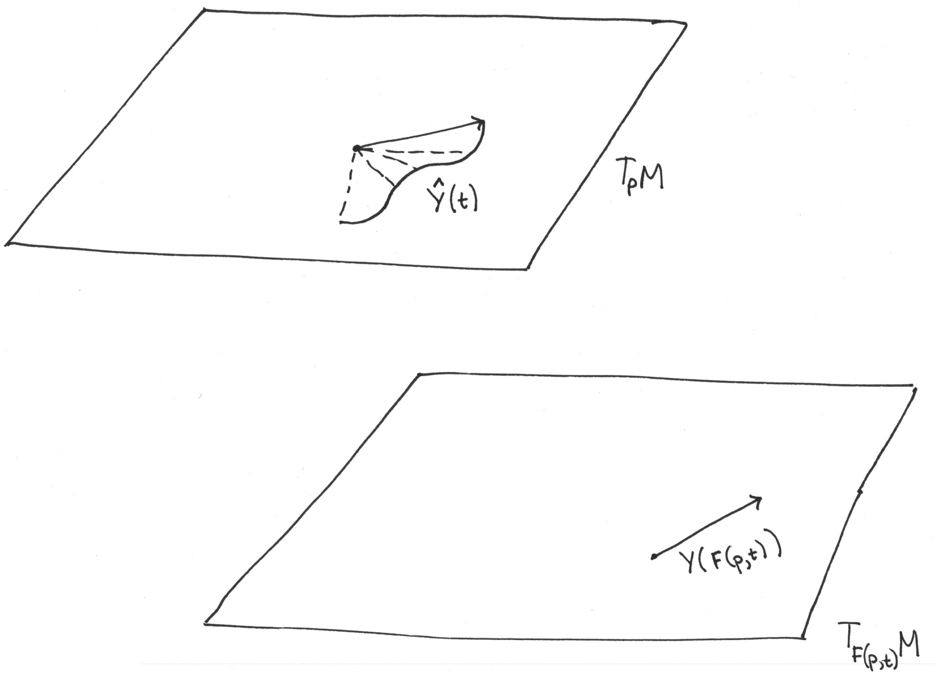

Figure 18: The curve in the tangent space.

We evaluate each of these quantities.

Firstly,

$$\begin{aligned}
\frac{d}{dt}|_{t=0}\ Y^j(q) &= \frac{d}{dt}|_{t=0}\ Y^j(q) \\
&= \frac{d}{dt}|_{t=0}\ Y^j(F(p,t)) \\
&= \frac{d}{dt}|_{t=0}\ (Y^j_\alpha \circ F^\alpha)(p,t) \\
&= \frac{\partial Y^j_\alpha}{\partial x^i}\frac{d}{dt}|_{t=0}\ F^{\alpha,i}(p,t) \\
&= \frac{\partial Y^j_\alpha}{\partial x^i} X^i \\
&\qquad \textit{(By applying Lemma 4.22)} \\
&= X(Y^j)
\end{aligned}$$

Now, notice that:

$$(b^i_j)|_{t=0}\ = \left(\frac{\partial F^{\alpha,i}_{t,\alpha}}{\partial x^j}\right)^{-1}|_{t=0}\ = (\delta^i_j)$$

since:

$$F^\alpha_\alpha(x,0) = (\psi_\alpha \circ F_0 \circ \psi_\alpha^{-1})(x) = \psi_\alpha(F(\psi_\alpha^{-1}(x),0)) = \psi_\alpha(\psi_\alpha^{-1}(x)) = x$$

Trivially,

$$Y^j(q)|_{t=0}\ = Y^j(F(p,0)) = Y^j(p)$$

Now, lastly, by applying Lemma 4.24:

$$\begin{aligned}
\frac{d}{dt}|_{t=0}\, b^i_j &= -b^i_k(0)\frac{d}{dt}|_{t=0}\, \frac{\partial F^{\alpha,k}_{t,\alpha}}{\partial x^l} b^l_j(0) \\
&= -\delta^i_k \frac{d}{dt}|_{t=0}\, \frac{\partial F^{\alpha,k}_{t,\alpha}}{\partial x^l} \delta^l_j \\
&\qquad \textit{(Computed Previously)} \\
&= -\frac{d}{dt}|_{t=0}\, \frac{\partial F^{\alpha,i}_{t,\alpha}}{\partial x^j} \\
&= -\frac{\partial}{\partial x^j}|_{x=p} \left(\frac{dF^{\alpha,i}_{t,\alpha}}{dt}|_{t=0} \right) \\
&\qquad \textit{(Since } F^{\alpha,i}_{t,\alpha} \in C^2\textit{)} \\
&= -\frac{\partial}{\partial x^j}|_{x=\psi_\alpha(p)} \left(\frac{d(\psi_\alpha \circ F_\alpha)^i}{dt}(\psi_\alpha(x), 0) \right) \\
&\qquad \textit{(We evaluate } D(F^\alpha_{t,\alpha}) \textit{ at } \psi_\alpha(x)\textit{))} \\
&= -\frac{\partial}{\partial x^j}|_{x=\psi_\alpha(p)} \left(X^i \circ \psi_\alpha^{-1}(x) \right) \\
&= -\frac{\partial X^i_\alpha}{\partial x^j}
\end{aligned}$$

So, putting these together:

$$\begin{aligned}
\frac{d}{dt}|_{t=0}\, (\mathrm{d}F)^{-1}(Y(q)) &= \left(X(Y^j)\delta^i_j + Y^j \left(-\frac{\partial X^i_\alpha}{\partial x^j} \right) \right) \frac{\partial}{\partial x^i} \\
&= X(Y^j)\frac{\partial}{\partial x^i} - Y^j \frac{\partial X^i_\alpha}{\partial x^j} \frac{\partial}{\partial x^i} \\
&= X(Y^j)\frac{\partial}{\partial x^i} - Y(X^i)\frac{\partial}{\partial x^i} \\
&= \left(X(Y^j) - Y(X^i) \right) \frac{\partial}{\partial x^i} \\
&= [X, Y] \\
&= \mathscr{L}_X Y
\end{aligned}$$

all evaluated at $p \in M$. □

Remark The proof of the geometric construction of the Lie Derivative shows that one cannot avoid the first derivatives of X. So, the Lie Derivative not only depends on points along X, but neighbourhoods around those points. This is unlike the familiar derivative operators in $\mathbb{R}^n$ which only depend on the point.

Remark As we saw before, we can define the Lie Derivative in a functional way for C^1 vector fields on C^2 manifolds. Nevertheless, it is still useful to have a geometric interpretation of the Lie Derivative.

Before we look at some examples, we consider and establish some of the properties of the Lie Derivative.

Theorem 4.26 (Properties of the Lie Derivative) *Let $X, Y \in \mathscr{X}(\Omega)$ and $f, g \in C^2(\Omega)$.*

1. *$\mathscr{L}_\cdot \cdot$ is bilinear*
2. *$\mathscr{L}_X Y = -\mathscr{L}_Y X$*
3. *$\mathscr{L}_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = 0$, where $\left\{x^i\right\}$ are any local coordinates*
4. *$\mathscr{L}_{fX}(gY) = fg\mathscr{L}_X Y + fX(g)Y - gY(f)X$*

5. $\mathscr{L}_{fX}(gY) = \mathscr{L}_{fX}(g)Y + g\mathscr{L}_{fX}Y$
(Leibniz Rule for the Lie Derivative)

Proof *1,2,3.* Trivial

4. Fix $h \in C^2(\Omega)$. Then,

$$\begin{aligned}[fX, gY]h &= fX(gY(h)) - gY(fX(h)) \\ &= f(X(g)Y(h) + gX(Y(h))) - g(Y(f)X(h) + fY(X(h))) \\ &= fX(g)Y(h) - gY(f)X(h) + fg(X(Y(h)) - Y(X(h)))\end{aligned}$$

The result follows since h was arbitrary.

5. From our previous result,

$$\mathscr{L}_{fX}(gY) = fg\mathscr{L}_X + \mathscr{L}_{fX}(g)Y - Y(f)X = \mathscr{L}_{fX}(g)Y + g(f\mathscr{L}_XY - \mathscr{L}_Y(f)X)$$

Re-applying the previous result with $g = 1$,

$$\mathscr{L}_{fX}Y = f[X, Y] + fX(1)Y - 1Y(f)X = f\mathscr{L}_XY - \mathscr{L}_Y(f)X$$

The result follows by combining these two expressions.

□

Example Let $M = \mathbb{R}^2$, and $X(x_1, x_2) = (-x_2, x_1)$.

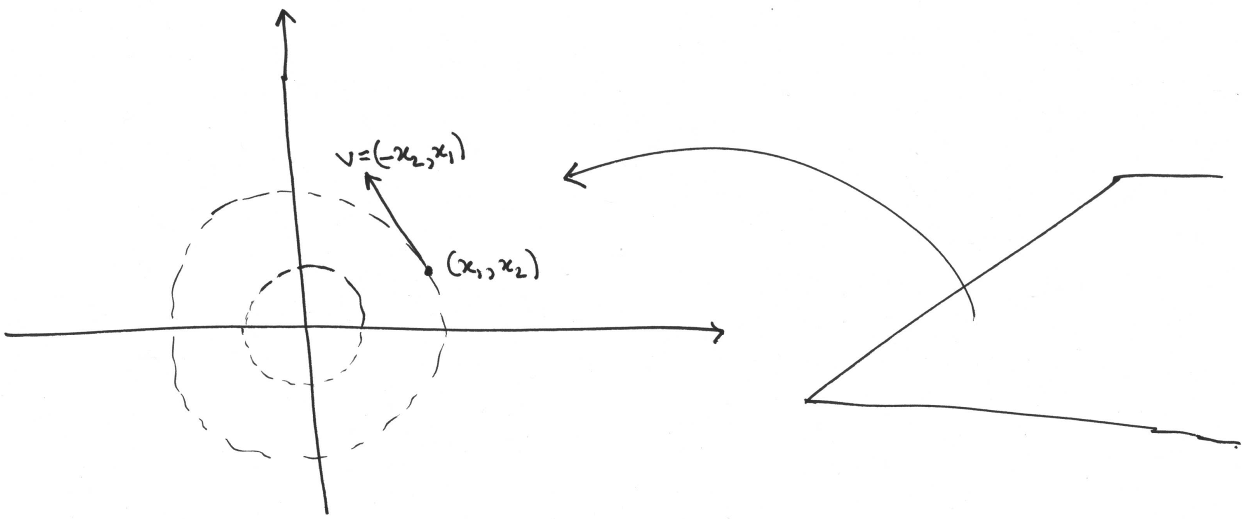

Figure 19: The vector field $X(x_1, x_2 = (-x_2, x_1)$.

Claim: X generates the 1 parameter family of counter clockwise rotation of $\mathbb{R}^2$:

$$F(x, t) = F(x_1, x_2, t) = (x_1'(t), x_2'(t)) = (x_1 \cos t - x_2 \sin t, x_1 \sin t + x_2 \cos t)$$

and:

$$\frac{dF}{dt}|_{t=0} \, (x_1, x_2, t) = (-x_2, x_1) = X$$

Note: that since $(1,0) = \frac{\partial}{\partial x_1}$ and $(0,1) = \frac{\partial}{\partial x_2}$, then,

$$X = -x_2 \frac{\partial}{\partial x_1} + x_1 \frac{\partial}{\partial x_2}$$

We can solve for $(\mathrm{d}F_t)^{-1}(\frac{\partial}{\partial x_1})$ by elementary geometric methods to find

$$\frac{d}{dt = 0}(\mathrm{d}F_t)^{-1}(\frac{\partial}{\partial x_1}) = -\frac{\partial}{\partial x_2}$$

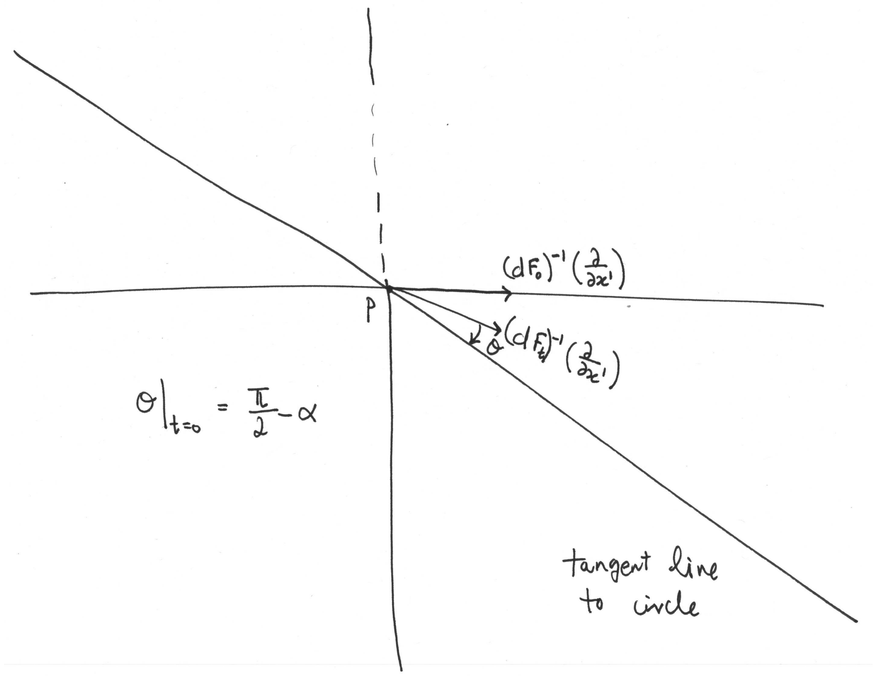

Figure 20: Elementary geometric solution to $(\mathrm{d}F_t)^{-1}\left(\frac{\partial}{\partial x^1}\right)$.

Now, set $Y = \frac{\partial}{\partial x_1}$. Then,

$$[X, Y] = [-x_2 \frac{\partial}{\partial x_1} + x_1 \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_1}]$$

5 Tensors

The motivation behind tensors is to construct meaningful algebraic objects on manifolds that are invariant under coordinate transforms.

5.1 Tensor Spaces

We refresh some familiar details about Vector Spaces. Let V be an n dimensional vector space over $\mathbb{R}$. We define the dual.

Definition 5.1 (Dual of a Vector Space) *Given a vector space V, we denote its Dual as V^* and define it as:*

$$V^* = \{L : V \to \mathbb{R}, L \text{ Linear}\}$$

We note (without proof), some facts about V^*.

Theorem 5.2 (Properties of Finite Dimensional Vector Spaces) *Let V be a vector space over $\mathbb{R}$ with $\dim V < \infty$. Then:*

1. *V^* is also a vector space over $\mathbb{R}$ with $\dim V^* = \dim V$.*
2. *$(V^*)^* = \{\phi : V^* \to \mathbb{R}\} \cong V$ in the sense that $\varepsilon_v : V^* \to \mathbb{R}$ for $v \in V$ defined by $\varepsilon_v(\phi) = \phi(v)$ for each $\phi \in V^*$. Here ε_v is called the evaluational functional of v, and $F : V \to (V^*)^*$ defined by $F(v) = \varepsilon_v$ defines a linear isomorphism.*
3. *If $\{v_1, \ldots, v_n\}$ forms a basis for V, then $\left\{\eta^1, \ldots, \eta^n\right\}$ defined by $\eta^i(v_j) = \delta^i_j$ is called the dual basis.*

If we set $V = \mathrm{T}_p M$, and $V^* = \mathrm{T}^*_p M$, then these quantities are independent of our coordinate system. We want to generalise V and V^*.

Definition 5.3 (Tensor of type (r, s)) *Associated to V, a multilinear functional:*

$$f : \left(\bigtimes_{i=1}^{r} V^* \right) \times \left(\bigtimes_{i=1}^{s} V \right) \to \mathbb{R}$$

is called a tensor of type (r, s).

Definition 5.4 (Tensor Space) *The set of all (r, s) type tensors over V forms a vector space over $\mathbb{R}$, denoted by:*

$$V_{(r,s)} = \left(\bigotimes^{r} V \right) \otimes \left(\bigotimes^{s} V^* \right)$$

Suppose that $V = \operatorname{span}\{v_i\}$ and let $V^* = \operatorname{span}\left\{\eta^i\right\}$ the dual basis. Then $f \in V_{(r,s)}$ is determined by it's value by:

$$(\eta^{i_1}, \ldots, \eta^{i_r}, v_{j_1}, \ldots, v_{j_s})$$

Now, note that:

$$(v_{i_1} \otimes \ldots \otimes v_{i_r} \otimes \eta^{j_1} \otimes \ldots \otimes \eta^{j_s})(\eta^{k_1}, \ldots, \eta^{k_r}, v_{l_1}, \ldots, v_{l_s}) = \delta^{k_1}_{i_1} \cdots \delta^{k_r}_{i_r} \cdot \delta^{j_1}_{l_1} \cdots \delta^{j_s}_{l_s}$$

Given this notation, it is obvious that every $f \in V_{(r,s)}$ can be written as:

$$f = f^{i_1,\ldots,i_r}_{j_1,\ldots,j_s} v_{i_1} \otimes \ldots \otimes v_{i_r} \otimes \eta^{j_1} \otimes \ldots \otimes \eta^{j_s}$$

This motivates:

Definition 5.5 (Basis of (r, s) Tensor Space) *We define the basis of $V_{(r,s)}$ by:*

$$\left\{v_{i_1} \otimes \ldots \otimes v_{i_r} \otimes \eta^{j_1} \otimes \ldots \otimes \eta^{j_s} : 1 \leq i_k, j_k \leq n\right\}$$

Now, we can begin relating this back to the geometry. Now we know that given a local coordinate chart with coordinates x^i at p, T_pM has a basis $\left\{\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right\}$. We want to find out the dual basis for T^*_pM.

We want the following result for the rest of the discussion.

Lemma 5.6 *Let $f : V \to \mathbb{R}^n$, with V open in M. Then for $u \in \mathrm{T}_pM$,*

$$\mathrm{d}f(u) = u(f)$$

Proof Let $u = [\sigma]$. Then, from our definition, $\mathrm{d}f(u) = [f \circ \sigma]$. Now, note that $f \circ \sigma : I \to \mathbb{R}^n$. By the remark following Definition 3.6, we regard $v = [\delta] \in \mathrm{T}_pU$ for $U \subseteq \mathbb{R}^n$ as $\frac{d}{dt}|_{t=0}\ \delta$. Now $[f \circ \sigma] \in \mathrm{T}_{f(p)}\mathbb{R}^n$, so it follows that:

$$\mathrm{d}f(u) = [f \circ \sigma] = \frac{d}{dt}|_{t=0}\ f \circ \sigma = u(f)$$

□

We want to consider functions $f : V \to \mathbb{R}$, so that our differential of f is $\mathrm{d}f : \mathrm{T}_pM \to \mathbb{R}$, since $\mathrm{T}_p\mathbb{R} = \mathbb{R}$. In particular, we are concerned with the following type of functions.

Definition 5.7 (Coordinate Functions) *Let $\left\{x^i\right\}$ be a coordinate choice with chart $\psi_\alpha : U_\alpha \to \mathbb{R}^n$. We define the coordinate functions $x^i : U \to \mathbb{R}$ by*

$$(x^i \circ \psi_\alpha^{-1})(p^1, \ldots, p^n) = p^i$$

We now give a formulation of the dual basis:

Theorem 5.8 (Dual Basis of T^*_pM) *The functionals $\mathrm{d}x^i : \mathrm{T}_pM \to \mathbb{R}$ form a dual basis for T^*_pM.*

Proof We've previously established that $\mathrm{d}x^i$ is a linear functional. We only need to check that $\mathrm{d}x^i(\frac{\partial}{\partial x^j}) = \delta^i_j$. We compute at $p \in V$:

$$\mathrm{d}x^i(\frac{\partial}{\partial x^j}) = \frac{d}{dt}|_{t=0}\ x^i \circ (\psi_\alpha^{-1} \circ X^j) = \frac{d}{dt}|_{t=0}\ (x^i \circ \psi_\alpha^{-1}) \circ (\psi_\alpha(p) + te_j) = \frac{d}{dt}|_{t=0}\ (\psi^i_\alpha(p) + t\delta^i_j) = \delta^i_j$$

□

For the sake of completeness, we introduce the following important terminology.

Definition 5.9 (Vectors and Forms) *We call the elements of T_pM* Vectors*, and elements of T^*_pM* Forms*.*

Now, we note that for $f \in \mathrm{T}_p^{(r,s)}M$, then locally we can write:

$$f = f^{i_1,\ldots,i_r}_{j_1,\ldots,j_s} \frac{\partial}{\partial x^1} \otimes \ldots \otimes \frac{\partial}{\partial x^n} \otimes \mathrm{d}x^{j_1} \otimes \ldots \otimes \mathrm{d}x^{j_s}$$

We introduce some important notation:

Definition 5.10 (Cotangent Bundle) *The Cotangent Bundle is given by* T^*M*:*

$$\mathrm{T}^*M = \left\{(p, \eta) : \eta \in \mathrm{T}_p^*M), p \in M\right\}$$

Definition 5.11 (Bundle of (r, s) Tensors) *The Bundle of (r, s) is denoted by $\mathrm{T}^{(r,s)}M$ and given by:*

$$\mathrm{T}^{(r,s)}M = \left\{(p, T) : T \in \mathrm{T}_p^{(r,s)}M, p \in M\right\}$$

Definition 5.12 (Tensor Field) *A tensor field of type (r, s) on M is a function $T : M \to \mathrm{T}^{(r,s)}M$ such that for each $p \in M$, $T(p) \in \mathrm{T}_p^{(r,s)}M$.*

5.2 Lie Derivatives of (r, s) type tensor fields

We begin our discussion considering pairings of vectors with dual vectors. This notation will be important in later discussions.

Definition 5.13 (Vector-Dual Vector Pairings) *Let $v \in V$ and $A \in V^*$. Then, we define $\langle \cdot, \cdot \rangle : V^* \times V \to \mathbb{R}$ by*

$$\langle A, v\rangle = A(v)$$

Definition 5.14 (Adjoint) *Let V, W be vector spaces. If $A : V \to W$ is a linear map, we define the adjoint $A^* : W^* \to V^*$ define by:*

$$\langle A^*(\omega), v\rangle = \langle w, A(v)\rangle$$

for all $\omega \in W^*$ and $v \in V$.

Now, we relate this back to geometry:

Definition 5.15 (Push Forward) *Let $f : M \to N$ be differentiable at $p \in M$ We define the push forward f_* at p as:*

$$f_* = \mathrm{d}f : \mathrm{T}_pM \to \mathrm{T}_{f(p)}N$$

Definition 5.16 (Pull Back) *Let $f : M \to N$ be differentiable at $p \in M$. Then, we define the pull back f^* at p as:*

$$f^* = (\mathrm{d}f)^* : \mathrm{T}^*_{f(p)}N \to \mathrm{T}^*_pM$$

Remark To emphasise for clarity: For any $\eta \in \mathrm{T}^*_{f(p)}N$, then $f^*(\eta) \in \mathrm{T}^*_pM$ via:

$$\langle f^*(\eta), v\rangle = \langle \eta, f_*(v)\rangle$$

for every $v \in \mathrm{T}_pM$.

We firstly establish what the pull back looks like under local coordinate charts.

Lemma 5.17 *Let $f : M \to N$ be differentiable at p, and let $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ be a coordinate chart around p with coordinates $\left\{x^i\right\}$, and $\phi_\alpha : U_\alpha \to \mathbb{R}^n$ be a coordinate chart around $f(p)$ with coordinates $\left\{y^j\right\}$. Then:*

$$f_*(\frac{\partial}{\partial x^i}) = \frac{\partial f_\alpha^{\alpha,j}}{\partial x^i}\frac{\partial}{\partial y^j}$$

and

$$f^*(\mathrm{d}y^j) = \frac{\partial f_\alpha^{\alpha,j}}{\partial x^i}\mathrm{d}x^i$$

Proof The first result is immediate from the definition of $\mathrm{d}f$.

For the second, fix $v^i \frac{\partial}{\partial x^i} \in \mathrm{T}_p M$. Then,

$$\langle f^*(\mathrm{d}y^j), v^i\frac{\partial}{\partial x^i}\rangle = \langle \mathrm{d}y^j, v^i\frac{\partial}{\partial x^i}{}_*\rangle = \langle \mathrm{d}y^j, v^i\frac{\partial f^\alpha_{\alpha,k}}{\partial x^i}\frac{\partial}{\partial y^k}\rangle = v^i\frac{\partial f^\alpha_{\alpha,j}}{\partial x^i}$$

This implies:

$$f^*(\mathrm{d}y^j) = \frac{\partial f^\alpha_{\alpha,j}}{\partial x^i}\mathrm{d}x^i$$

$\square$

Remark Note that this result shows that:

$$f_* \in \mathrm{T}_p^*M \otimes \mathrm{T}_{f(p)}N$$
$$f^* \in \mathrm{T}_{f(p)}N \otimes \mathrm{T}_p^*M$$

via the observation:

$$f_* = \frac{\partial f^\alpha_{\alpha,j}}{\partial x^i}\mathrm{d}x^i \otimes \frac{\partial}{\partial y^j}$$
$$f^* = \frac{\partial f^\alpha_{\alpha,j}}{\partial x^i}\frac{\partial}{\partial y^j} \otimes \mathrm{d}x^i$$

Remark If we write $\mathrm{D}f_* = \left(\frac{\partial f^\alpha_{\alpha,j}}{\partial x^i}\right)$, then it follows that $\mathrm{D}{f_*}^T = \mathrm{D}f^*$, the transpose.

Remark This result also gives a rigorous interpretation of ODEs:

$$\mathrm{d}y = f(x,y)\mathrm{d}x$$

We also want to consider pull backs and push forwards of tensor fields:

Definition 5.18 (Push Forward, Pull Back of Tensor Fields) *Let $F : M \to N$ be differentiable at p. Let $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ be a local chart around p, and $\phi_\alpha : V_\alpha \to \mathbb{R}^n$ a chart around $F(p)$. Let*

$$S = s^{i_1,\ldots,i^k}\frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_k}} \in \mathrm{T}_p^{(k,0)}M$$
$$T = t_{j_1,\ldots,j_k}\mathrm{d}y^{j_1} \otimes \ldots \otimes \mathrm{d}y^{j_k} \in \mathrm{T}_p^{(0,k)}N$$

Then, we define the push forward of S by F as:

$$F_*(S) = s^{i_1,\ldots,i^k}F_*(\frac{\partial}{\partial x^{i_1}}) \otimes \ldots \otimes F_*(\frac{\partial}{\partial x^{i_k}}) \in \mathrm{T}_p^{(k,0)}N$$

If F is surjective, we can define the pull back of T by F as:

$$F^*(T) = t_{j_1,\dots,j_k} F^*(\mathrm{d}y^{j_1}) \otimes \dots \otimes F^*(\mathrm{d}y^{j_k}) \in \mathrm{T}_p^{(0,k)} M$$

If F is a diffeomorphism, and:

$$C = c^{i_1,\dots,i^k}_{j_1,\dots,j_k} \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_k}} \otimes \mathrm{d}x^{j_1} \otimes \dots \otimes \mathrm{d}x^{j_k} \in \mathrm{T}_p^{(k,l)} M$$
$$D = d^{i_1,\dots,i^k}_{j_1,\dots,j_k} \frac{\partial}{\partial y^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial y^{i_k}} \otimes \mathrm{d}y^{j_1} \otimes \dots \otimes \mathrm{d}y^{j_k} \in \mathrm{T}_p^{(k,l)} N$$

then, we can define the push forward and pull back as:

$$F_*(C) = c^{i_1,\dots,i^k}_{j_1,\dots,j_k} F_*(\frac{\partial}{\partial x^{i_1}}) \otimes \dots \otimes F_*(\frac{\partial}{\partial x^{i_k}}) \otimes (F^*)^{-1}(\mathrm{d}x^{j_1}) \otimes \dots \otimes (F^*)^{-1}(\mathrm{d}x^{j_k})$$
$$F^*(C) = c^{i_1,\dots,i^k}_{j_1,\dots,j_k} (F_*)^{-1}(\frac{\partial}{\partial x^{i_1}}) \otimes \dots \otimes (F_*)^{-1}(\frac{\partial}{\partial x^{i_k}}) \otimes F^*(\mathrm{d}x^{j_1}) \otimes \dots \otimes F^*(\mathrm{d}x^{j_k})$$

Now we have the sufficient machinery to move onto considering Lie Derivatives of tensor fields by vector fields. Recall that given a vector field X, we can find a 1-parameter family of diffeomorphisms $\{F_t\}$ and that $\mathrm{d}F_t : \mathrm{T}_p M \to \mathrm{T}_{F_t(p)} M$ is a linear isomorphism. This also means that the adjoint map $(\mathrm{d}F_t)^* : \mathrm{T}_{f(p)} M^* \to \mathrm{T}_p^* M$ is also a linear isomorphism.

This allows us to make the following definition.

Definition 5.19 (Lie Derivative of Tensor Fields) *Let X be a vector field, and let T be a (r, s) tensor field. Then we define the lie derivative at p by:*

$$\mathscr{L}_X(T)(p) = \frac{d}{dt}|_{t=0}\, (F_t^*)^{-1}(T(F_t(p)))$$

6 The Affine Connection

As mentioned previously, the Lie Derivative of Y with respect to a vectorfield X depends not only on X, but at each point, a small neighbourhood around X. This is unlike our directional derivatives, where the derivative is dependent only at each point.

Our goal here is a way to differentiate vector fields by vector fields such that the derivative of the vector field is only dependent on each point.

6.1 Connection on Manifolds

Definition 6.1 (Affine Connection) *Suppose M is C^r, $r > 1$. An affine connection on M is a map,*

$$\nabla : \mathscr{X}^0(M) \times \mathscr{X}^1(M) \to \mathscr{X}^0(M)$$

such that:

1. *For $X, Y \in \mathscr{X}^0(M), Z \in \mathscr{X}^1(M)$, and $f, g : M \to \mathbb{R}$,*
$$\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$$
2. *For $X \in \mathscr{X}^0(M), Y, Z \in \mathscr{X}^1(M)$, and constants $\alpha, \beta \in \mathbb{R}$,*
$$\nabla_X(\alpha Y + \beta Z) = \alpha\nabla_X Y + \beta\nabla_X Z$$
3. *For $X \in \mathscr{X}^1(M), Y \in \mathscr{X}^0(M)$, and f differentiable,*
$$\nabla_X(fY) = X(f)Y + f\nabla_X Y$$

Such a map is also called a covarient derivative.

Remark If f is a differentiable function, we could define

$$\nabla_X f = X(f)$$

Then, the last rule indeed “looks” like a Leibniz rule.

Remark Notice that $\nabla : \mathscr{X}^0(M) \times \mathscr{X}^1(M) \to \mathscr{X}^0(M)$. It does not make sense (yet!) to compute: $\nabla_X Y$, where $X \in \mathscr{X}^0(U)$, or $Y \in \mathscr{X}^1(U)$ where $U \subsetneq M$. This is because our connection ∇ is a global structure on M.

Remark Unlike the Lie Derivative, there are many such connections, and every affine connection does not arise from some deep property of the manifold. Rather, connections are external structures imposed on the manifold.

Lemma 6.2 *The Lie Derivative is not a connection.*

Proof Firstly, notice that

$$\mathscr{L} : \mathscr{X}^1(M) \times \mathscr{X}^1(M) \to \mathscr{X}^0(M)$$

That is, it doesn't make sense to consider $\mathscr{L}_X Y$, where $X \in \mathscr{X}^0(M)$. But even if we altered our definition of the affine connection as a map from and into $\mathscr{X}^\infty(M)$, we can show that the first condition fails by considering:

$$\mathscr{L}_{fX} Y = f\mathscr{L}_X Y - X(f)Y \neq f\mathscr{L}_X Y$$

in general. □

We establish a few auxiliary results that will enable us to show why such a connection only depends on only the points along the vectorfield. Our general approach is to "restrict" vectorfields to 0 outside an open set containing p, and show that they are still equal to the original vectorfield inside the neighbourhood.

Lemma 6.3 (Cutoff in $\mathbb{R}^n$) *There exists a smooth function $\xi : \mathbb{R}^n \to [0,1]$ such that:*

$$\xi(x) = \begin{cases} 1 & x \in B_1(0) \\ 0 & x \in \mathbb{R}^n \setminus B_2(0) \end{cases}$$

Corollary 6.4 (Cutoff on Manifolds) *Let $p \in M$. Then for any neighbourhood U of p, there exists a neighbourhood K around p such that $K \Subset U$, and a differentiable $\eta : M \to [0,1]$ such that:*

$$\eta(x) = \begin{cases} 1 & x \in K \\ 0 & x \in M \setminus U \end{cases}$$

Proof Since we have a maximal differentiable structure $\mathscr{D}$ associated with M, we can find a coordinate chart $\psi_\alpha : U_\alpha \to B_3(0)$. with $U_\alpha \subseteq U$. Then, define

$$\hat{\eta}(x) = (\xi \circ \psi_\alpha)(x)$$

So, we have that:

$$\hat{\eta}(x) = \begin{cases} 1 & x \in \psi_\alpha^{-1}(B_1(0)) \\ 0 & x \in U_\alpha \setminus \psi_\alpha^{-1}(B_2(0)) \end{cases}$$

So, extend $\hat{\eta}$ to η by:

$$\eta(x) = \begin{cases} \hat{\eta}(x) & x \in U_\alpha \\ 0 & x \in M \setminus U_\alpha \end{cases}$$

□

Now, we prove our first result regarding the local properties of ∇.

Lemma 6.5 *Let $X \in \mathscr{X}^0(M)$, $Y \in \mathscr{X}^1(M)$. Let U open in M, and suppose $X|_U = 0$ or $Y|_U = 0$. Then,*

$$\nabla_X Y(p) = 0$$

for all $p \in U$.

Proof For the first case, fix $p \in U$. By Corollary 6.4, we can find a $K \Subset U$, and a function $\xi : M \to \mathbb{R}$ such that $\xi = 1$ in K, and $\xi = 0$ outside of U.

Since f differentiable, the vectorfield $\hat{X} = fX$ is also differentiable. Also, $\hat{X} = 0$ everywhere. Now, it follows that:

$$0 = \nabla_{fX} Y = f \nabla_X Y$$

But by construction $f(p) = 1$, so

$$0 = f(p)\nabla_X Y(p) = \nabla_X Y(p)$$

For the second case, again, fix $p \in M$. We follow the same trick, and this time define $\hat{Y} = fY$. Then,

$$0 = \nabla_X(fY) = X(f)Y + f\nabla_X Y$$

Now, since $Y|_U = 0$, in particular $Y(p) = 0$, and as before $f(p) = 1$. So,

$$0 = X(f)Y(p) + f(p)\nabla_X Y(p) = \nabla_X Y(p)$$

□

Corollary 6.6 *Let $X_1, X_2 \in \mathscr{X}^1(M)$ and $Y_1, Y_2 \in \mathscr{X}^0(M)$, satisfying $X_1 = X_2$ and $Y_1 = Y_2$ on some U open in M. Then $\nabla_{X_1} Y_1 = \nabla_{X_2} Y_2$ on U.*

Proof Apply the lemma to $X = (X_1 - X_2)|_U = 0$, and $Y = (Y_1 - Y_2)|_U = 0$. □

These results highlight an important fact. They allow us to localise any affine connection ∇, motivating the following definition.

Definition 6.7 (Local Affine Connection) *Let U be open in M, and $X \in \mathscr{X}^0(U)$ and $Y \in \mathscr{X}^1(U)$. Then, we define $\nabla : \mathscr{X}^0(U) \times \mathscr{X}^1(U) \to \mathscr{X}^0(U)$ as:*

$$\nabla_X Y(p) = \nabla_{\hat{X}} \hat{Y}(p)$$

for any $\hat{X} \in \mathscr{X}^1(M), \hat{Y} \in \mathscr{X}^0(M)$ satisfying $\hat{X}|_W = X|_W$, and $\hat{Y}|_W = Y|_W$ for any open W containing p.

Remark This definition is well defined by the previous corollary because given $\tilde{X}$, and $\tilde{Y}$ satisfying $\tilde{X}|_W = X$, $\tilde{Y}|_W = Y$, and by applying the result, we get $\nabla_{\hat{X}} \hat{Y}(p) = \nabla_{\tilde{X}} \tilde{Y}(p)$

By localisation, we introduce an important geometric object.

Definition 6.8 (Christoffel Symbols) *Let $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ be a coordinate chart with coordinates $\left\{x^i\right\}$. We define the Christoffel Symbols $\left\{\Gamma^k_{ij}\right\}$ of ∇ associated with $\left\{x^i\right\}$ by:*

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma^k_{ij} \frac{\partial}{\partial x^k}$$

Lemma 6.9 *Let $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ have coordinates $\left\{x^i\right\}$. Then ∇ in U_α is completely determined by the Christoffel symbols.*

Proof Let $X = X^i \frac{\partial}{\partial x^i}, Y = Y^j \frac{\partial}{\partial x^j}$. Then,

$$\nabla_X Y = \nabla_{X^i \frac{\partial}{\partial x^i}} \left(Y^i \frac{\partial}{\partial x^j} \right) = X^i \left(Y^j \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} + \frac{\partial}{\partial x^i}(Y^j) \frac{\partial}{\partial x^i} \right) = X^i Y^j \Gamma^k_{ij} \frac{\partial}{\partial x^k} + \frac{\partial}{\partial x^i}(Y^j) \frac{\partial}{\partial x^i}$$

and the result follows. □

6.2 Parallel Transport

In this section, we give a useful theoretical application of an affine connection on a Manifold. We illustrate how such a structure can be used to "connect" two different tangent spaces.

Note that from this section onwards we begin to liberally use the notation $\sigma'(t)$ to denote tangent vectors.

Definition 6.10 (Vector Field Along a Curve) *Let $\sigma : I \to M$ be a differentiable curve. A vector field along σ is a map $V : I \to \mathrm{T}M$ such that $V(t) \in \mathrm{T}_{\sigma(t)}M$ and $\pi \circ V = \sigma$, where $\pi : \mathrm{T}_b M \to M$ is the canonical projection.*

Remark Given such a vectorfield along a curve, we could assert the existence of a vectorfield $\tilde{V}$ on M such that $\tilde{V} \circ \gamma = V$.

Definition 6.11 (Parallel Vectorfield) *We say that a vectorfield V along a curve σ is parallel if:*

$$\nabla_{\sigma'(t)} V = 0$$

Remark We need to take a moment to fully appreciate this formulation. It is deep and non trivial. We began with a connection ∇ which is a global property of a manifold, and we showed it depends only at each point $p \in M$. This is why we can formulate this definition at all.

Ideally, we want to compute a parallel vectorfield inside some chart. First, we require the following important definition.

Definition 6.12 *Let V be a vectorfield along σ, and write $V = V^i \frac{\partial}{\partial x^i}$ in U_α. Then,*

$$\sigma'(t)(V^i) = \sigma'(t)(\tilde{V}^i)$$

where $\tilde{V}$ is a vectorfield on M such that $\tilde{V}|_{\gamma(t)} = V$ with $\tilde{V} = \tilde{V}^i \frac{\partial}{\partial x^i}$ in U_α.

Lemma 6.13

$$\sigma'(t)(V^j) = \frac{d}{dt} V^j$$

Proof We compute:

$$\sigma'(t)(V^j) = \sigma'(t)(\tilde{V}^j) = \frac{d}{dt}(\tilde{V}^j \circ \sigma) = \frac{d}{dt} V^j$$

□

We have the following important result.

Theorem 6.14 *Let V be a vectorfield along σ. Then V is parallel if and only if*

$$\dot{V} + V^j \dot{\sigma}^{\alpha,i} \Gamma^k_{ij} = 0$$

for all k and every coordinate chart U_α.

Proof Fix $t \in I$, and a chart ψ_α around t. Then, $\sigma'(t) = \dot{\sigma}^{\alpha,i}(t) \frac{\partial}{\partial x^i}$, and $V = V^j \frac{\partial}{\partial x^j}$

$$\begin{aligned}
\nabla_{\sigma'(t)} V &= \nabla_{\sigma'(t)} \left(V^j \frac{\partial}{\partial x^j} \right) \\
&= \sigma'(t)(V^j) \frac{\partial}{\partial x^j} + V^j(t) \nabla_{\dot{\sigma}^{\alpha,i} \frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \\
&= \dot{V}^j(t) \frac{\partial}{\partial x^j} + V^j(t) \dot{\sigma}^{\alpha,i}(t) \Gamma^k_{ij} \frac{\partial}{\partial x^k} \\
&= \left(\dot{V}^k(t) + V^j(t) \dot{\sigma}^{\alpha,i}(t) \Gamma^k_{ij} \right) \frac{\partial}{\partial x^k}
\end{aligned}$$

Then, it follows that $\nabla_{\sigma'(t)} V = 0$ if and only if

$$\left(\dot{V}^k(t) + V^j(t) \dot{\sigma}^{\alpha,i}(t) \Gamma^k_{ij} \right) \frac{\partial}{\partial x^k} = 0$$

and the result follows. □

We note the following important lemma without proof.

Lemma 6.15 *Let $\sigma : [0,1] \to M$. Then the ODE*

$$\left(\dot{V}^k(t) + V^j(t)\dot{\sigma}^{\alpha,i}(t)\Gamma^k_{ij}\right)\frac{\partial}{\partial x^k} = 0$$

has a there exists a unique solution to V along σ given an initial condition $V(0) = v$.

Remark ODE theory guarantees us there exists a unique solution to V inside some fixed coordinate chart. By coordinate transforms, we can extend this solution to the entire domain of $\operatorname{Img} \sigma$.

Example Put $M = \mathbb{R}^n$, with standard coordinates $(x^1, \ldots, x^n)$. Then, any connection ∇ is determined by the Christoffel symbols in $\mathbb{R}^n$. So, we take $\Gamma^k_{ij} \equiv 0$.

Now, it follows that:

$$\nabla_{\sigma'(t)} V = 0 \iff \dot{V}^k(t)\forall k \iff V^k(t) = \xi^k$$

So, parallel actually means that the vector is parallel along our curve (since V is constant).

We can now illustrate how the affine connection can be used to connect two tangent spaces along a curve.

Definition 6.16 (Parallel Transport) *Let $\sigma : [0,1] \to M$. Then, we define parallel transport by the map $\Phi_\sigma : \mathrm{T}_{\sigma(0)}M \to \mathrm{T}_{\sigma(1)}M$ where*

$$\Phi_\sigma(v) = \tilde{v}$$

where V is the unique solution to the ODE with the initial condition $V(0) = v$ and $V(1) = \tilde{v}$.

We note importantly that:

Lemma 6.17 *$\Phi_\sigma : \mathrm{T}_{\sigma(0)}M \to \mathrm{T}_{\sigma(1)}M$ is a linear isomorphism.*

Remark Without a connection, we have no way of associating two tangent spaces together. We can in fact define ∇ backwards by looking maps ∇ giving rise to Parallel Transport maps.

Remark The differentiability of σ implies that the underlying points of the tangent spaces being parallel transported need to lie in the same connected component.

Remark Given another curve γ, we have that in general $\Phi_\sigma \neq \Phi_\gamma$.

7 Curvature and Metrics

So far, we have considered an abstract connection. We can get quite far without imposing further structure. Parallel transport is one immediate consequence, but so is the curvature tensor. We will introduce this initially to highlight that further structure is unnecessary. We shall see how we can do quite a lot more by imposing extra structure on our connection.

7.1 Curvature

Definition 7.1 (Reimannian Curvature Operator) *Let $X, Y, Z \in \mathscr{X}(M)$. We define the curvature operator $\mathscr{R}(\cdot,\cdot)(\cdot) : \mathscr{X}(M) \times \mathscr{X}(M) \times \mathscr{X}(M) \to \mathscr{X}(M)$:*

$$\mathscr{R}(X,Y)Z = \nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) - \nabla_{\mathscr{L}_X Y} Z$$

Remark The curvature operator is a $(1,3)$ tensor field since $\mathscr{R}(X,Y)Z$ is a $(0,1)$ tensor. In fact, this is often referred to as the Curvature Tensor.

Theorem 7.2 (Function Linearity of the Curvature Operator) *$\mathscr{R}(\cdot,\cdot)(\cdot)$ is function linear.*

Proof Compute:

$$\begin{aligned}
\mathscr{R}(X,Y)(fZ) &= \nabla_X(\nabla_Y(fZ)) - \nabla_Y(\nabla_X(fZ)) - \nabla_{\mathscr{L}_X Y}(fZ) \\
&= \nabla_X(Y(f)Z + f\nabla_Y Z) - \nabla_Y(X(f)Z + f\nabla_X Z) - \mathscr{L}_X Y(f)Z - f\nabla_{\mathscr{L}_X Y} Z \\
&= XY(f)Z + Y(f)\nabla_X Z + X(f)\nabla_Y Z + f\nabla_X(\nabla_Y Z) \\
&\qquad - YX(f)Z - X(f)\nabla_Y Z - Y(f)\nabla_X Z - f\nabla_Y(\nabla_X Z) \\
&\qquad - (X(Yf) - Y(Xf))Z - f\nabla_{\mathscr{L}_X Y} Z \\
&= f\nabla_X(\nabla_Y Z) - f\nabla_Y(\nabla_X Z) - f\nabla_{\mathscr{L}_X Y} Z \\
&= f\mathscr{R}(X,Y)Z
\end{aligned}$$

Similarly for $\mathscr{R}(fX,Y)Z, \mathscr{R}(X,fY)Z$. □

Corollary 7.3 *$\mathscr{R}(\cdot,\cdot)(\cdot)$ can be localised to any U open in M.*

The proof of this is similar to the localisation of the affine connection and is left as an exercise.

Corollary 7.4 *$\mathscr{R}(X,Y)(Z)$ only depends on $X(p), Y(p), Z(p)$ for every $p \in M$.*

Proof Let $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ be a local coordinate chart with coordinates $\left\{x^i\right\}$. We write $X = X^i \frac{\partial}{\partial x^i}$, $Y = Y^j \frac{\partial}{\partial x^k}$, and $Z = Z^k \frac{\partial}{\partial x^k}$. Then, by localisation and function linearity, we have

$$\mathscr{R}(X,Y)Z = X^i Y^j Z^k \mathscr{R}\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)\left(\frac{\partial}{\partial x^k}\right)$$

and the result follows. □

To use this curvature tensor to construct other curvatures on a manifold requires us to impose further restrictions on our connection.

7.2 The Metric

The usual inner product in $\mathbb{R}^n$ allows us to measure the lengths and angles of vectors. Our goal is to consider the way in which a metric can interact with a specific type of connection.

Definition 7.5 (Reimannian Metric) *A Riemannian Metric g on M is a $(0, 2)$ tensor field such that at each $p \in M$, g is an inner product on $\mathrm{T}_p M$.*

Another important product we consider is:

Definition 7.6 (Lorenzian Product) *Let V be a vector space over $\mathbb{R}$. A Lorentzian product is a symmetric bilinear functional $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ such that there exists a basis $\{v_1, \ldots, v_n\}$ of V and:*

$$\langle v_i, v_j \rangle = \begin{cases} -1 & i = j = 1 \\ \delta_{ij} & \textit{otherwise} \end{cases}$$

We can now define an important a slightly different type of metric on a manifold.

Definition 7.7 (Lorenzian Metric) *A Lorentzian Metric g on M is a $(0, 2)$ tensor field such that at each $p \in M$, g is a Lorentzian product on $\mathrm{T}_p M$.*

Remark To show the existence of such a structure on *any* manifold, we need a Partition of Unity. This is done later in §8.

Definition 7.8 (Reimannian/Lorenzian Manifold) *If M is a manifold equipped with a Riemannian/Lorentzian metric g, then we say that (M, g) is a Riemannian/Lorentzian Manifold.*

Now, we require some more notation to make a link between connections and metrics.

Definition 7.9 (Torsion Free) *We say that a connection ∇ is Torsion Free if for all $X, Y \in \mathscr{X}(M)$,*

$$\nabla_X Y - \nabla_Y X = \mathscr{L}_X Y$$

Lemma 7.10 *∇ is torsion free if and only if $\Gamma^k_{ij} = \Gamma^k_{ji}$ for all coordinate charts.*

The proof is left as an exercise.

Definition 7.11 (Metric Compatible) *Let $\langle \cdot, \cdot \rangle$ be a metric on M. Then, we say that a connection ∇ is metric compatible with $\langle \cdot, \cdot \rangle$ if for all $X, Y, Z \in \mathscr{X}(M)$,*

$$X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle$$

Definition 7.12 (Levi-Civita Connection) *Let $\langle \cdot, \cdot \rangle$ be a metric on M, and ∇ a connection. If ∇ is torsion free and metric compatible with $\langle \cdot, \cdot \rangle$, then we say that ∇ is a Levi-Civita connection.*

Definition 7.13 (Non Degenerate $(0, 2)$ Tensor Field) *Let V be a vector space over $\mathbb{R}$, and $T \in V^* \otimes V^*$. If for all $v \in V$, $T(u, v) = 0$ implies $u = 0$, then we say that T is non degenerate.*

Theorem 7.14 *Let $\langle\cdot,\cdot\rangle$ be a symmetric non degenerate $(0,2)$ tensor field on M. Then there exists a unique ∇ such that for all $X, Y, Z \in \mathscr{X}(M)$,*

1. $\nabla_X Y - \nabla_Y X = \mathscr{L}_X Y$
2. $X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$

Proof Fix $X, Y, Z \in \mathscr{X}(M)$. By the non degeneracy hypothesis on $\langle\cdot,\cdot\rangle$, it suffices to define $\nabla_X Y$ by defining $\langle \nabla_X Y, Z\rangle$. We proceed the proof by assuming such a connection exists to compute an expression for $\langle \nabla_X Y, Z\rangle$.

We note that:

$$\begin{aligned} X\langle Y, Z\rangle &= \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle \\ Y\langle Z, X\rangle &= \langle \nabla_Y Z, X\rangle + \langle Z, \nabla_Y X\rangle \\ Z\langle X, Y\rangle &= \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle \end{aligned}$$

Then, it follows that:

$$\begin{aligned} X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle &= \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle + \langle \nabla_Y Z, X\rangle + \langle Z, \nabla_Y X\rangle - \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle \\ &= \langle \nabla_X Y, Z\rangle + \langle \nabla_Y X, Z\rangle + \langle \nabla_X Z, Y\rangle - \langle \nabla_Z X, Y\rangle + \langle \nabla_Y Z, X\rangle - \langle \nabla_Z Y, X\rangle \\ &= 2\langle \nabla_X Y, Z\rangle - \langle \nabla_X Y, Z\rangle + \langle \nabla_Y X, Z\rangle + \langle \mathscr{L}_X Z, Y\rangle + \langle \mathscr{L}_Y Z, X\rangle \\ &= 2\langle \nabla_X Y, Z\rangle - \langle \mathscr{L}_X Y, Z\rangle + \langle \mathscr{L}_X Z, Y\rangle + \langle \mathscr{L}_Y Z, X\rangle \end{aligned}$$

So, we define:

$$\langle \nabla_X Y, Z\rangle = \frac{1}{2}\left(\langle \mathscr{L}_X Y, Z\rangle - \langle \mathscr{L}_X Z, Y\rangle - \langle \mathscr{L}_Y Z, X\rangle + X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle\right)$$

We leave it as an exercise to check that defining ∇ in this way indeed gives the result. □

Corollary 7.15 (Existence of Levi-Civita Connection) *For every Riemannian Manifold $(M, \langle\cdot,\cdot\rangle)$, there exists a unique Levi-Civita connection.*

The proof of this is immediate, since it is easy to check that a Riemannian metric is indeed a $(0,2)$ non degenerate tensor field.

Example Take $M = \mathbb{R}^n$. Let $X, Y \in \mathscr{X}(M)$. So, with the standard coordinates $\left\{x^i\right\}$, we can write $X = X^i \frac{\partial}{\partial x^i}$, $Y = Y^j \frac{\partial}{\partial x^j}$. We equip M with the usual metric $\langle X, Y\rangle = X^i Y_i$.

Define $\nabla_X Y = \mathrm{D}_X Y^i \frac{\partial}{\partial x^i}$. We check that this is indeed metric compatible:

$$X\langle Y, Z\rangle = X(Y^i Z_j) = \mathrm{D}_X(Y^i) Z_i + Y^i \mathrm{D}_X(Z_i) = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$$

Example Let $M \subseteq \mathbb{R}^{n+k}$ an n-submanifold. Assume that we have the usual metric $\langle\cdot,\cdot\rangle$ as defined previously in $\mathbb{R}^{n+k}$.

Let $v, w \in \mathrm{T}_p M$, and define:

$$\langle v, w\rangle_M = \langle v, w\rangle$$

Let $X, Y, Z \in \mathscr{X}(M)$. We have:

$$Z\langle X, Y\rangle_M = \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle$$

We want:

$$\langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle = \langle \nabla^M_Z X, Y \rangle_M + \langle X, \nabla^M_Z Y \rangle_M$$

where ∇^M is the connection on M. We look for something stronger:

$$\langle \nabla_Z X, Y \rangle = \langle \nabla^M_Z X, Y \rangle_M$$
$$\langle X, \nabla_Z Y \rangle = \langle X, \nabla^M_Z Y \rangle_M$$

We define:

$$\nabla^M_X Y = (\nabla_Z X)^{\perp}$$

where $\perp$: $\mathrm{T}_p\mathbb{R}^{n+k} \to \mathrm{T}_p M$ giving us the connection we want.

7.3 More Curvature

In this section, we give a brief survey of other useful curvatures on a manifold. Although we can define the curvature tensor with an abstract connection, we can construct many other curvatures with desirable properties by considering $\mathscr{R}(\cdot, \cdot)(\cdot)$ given by the unique Levi-Civita ∇ on a manifold $(M, \langle \cdot, \cdot \rangle)$.

Definition 7.16 (Curvature 4-Tensor) *We define $\mathscr{R}(\cdot, \cdot, \cdot, \cdot) : \mathscr{X}(M) \times \mathscr{X}(M) \times \mathscr{X}(M) \times \mathscr{X}(M) \to \mathbb{R}$ by:*

$$\mathscr{R}(X, Y, Z, W) = \langle \mathscr{R}(X, Y) Z, W \rangle$$

Theorem 7.17 (Properties of the Curvature 4-Tensor) *Let $(M, \langle \cdot, \cdot \rangle)$ with Levi-Civita connection ∇. Then, for all $X, Y, Z, W \in \mathscr{X}(M)$,*

1. $\mathscr{R}(X, Y, Z, W) + \mathscr{R}(Y, Z, X, W) + \mathscr{R}(Z, X, Y, W) = 0$
2. $\mathscr{R}(X, Y, Z, W) = -\mathscr{R}(Y, X, Z, W)$
3. $\mathscr{R}(X, Y, Z, W) = -\mathscr{R}(X, Y, W, Z)$
4. $\mathscr{R}(Z, W, X, Y) = \mathscr{R}(X, Y, Z, W)$

The proof of this is left as an exercise.

We now move onto briefly consider some other useful curvatures on a Manifold.

Definition 7.18 (Perpendicular Tangent Vectors) *We say that $u, v \in \mathrm{T}_p M$ are perpendicular if $\langle u, v \rangle = 0$ and we write $u \perp v$.*

Definition 7.19 (Length of a Tangent Vector) *We define the length of a tangent vector $u \in \mathrm{T}_p M$ to be:*

$$\|u\| = \sqrt{\langle u, u \rangle}$$

Definition 7.20 (Sectional Curvature) *Let $u, v \in \mathrm{T}_p M$ with $u \perp v$ and $\|u\| = \|v\| = 1$. Then, the sectional curvature of $P = \mathrm{span}\ \{u, v\}$ is given by:*

$$k(u, v) = -\mathscr{R}(u, v, u, v) = \mathscr{R}(u, v, v, u)$$

Definition 7.21 (Ricci Curvature) *Let $\{e_i\}$ be an orthonormal basis for T_pM. Let $X, Y \in \mathscr{X}(M)$. Then, we define Ricci Curvature as:*

$$\mathrm{Ric}\,(X, Y) = \Sigma_{i=1}^{n} \mathscr{R}\,(X, e_i, e_i, Y)$$

Remark The Ricci Curvature is sometimes referred to as the Ricci Tensor, because in fact $\mathrm{Ric}\,(\cdot,\cdot)$ is a $(0, 2)$ tensor field.

Lemma 7.22 $\mathrm{Ric}\,(X, Y) = \mathrm{Ric}\,(Y, X)$

The proof of this follows trivially from the definition of the 4 Tensor.

Lemma 7.23 *Let $X \in \mathscr{X}(M)$ and choose an orthonormal basis at p to T_pM such that e_1 is in the direction of X. Then,*

$$\mathrm{Ric}\,(X, X) = \Sigma_{i=2}^{n} \|x\|^2 \mathrm{Ric}\,(e_1, e_i, e_i, e_1)$$

Again the proof of this fact follows trivially from the definition.

Definition 7.24 (Scalar Curvature) *Given an orthonormal basis $\{e_i\}$ to T_pM, the scalar curvature $\mathscr{R}_S$ is given by:*

$$\mathscr{R}_S = \Sigma_{i=1}^{n} \mathrm{Ric}\,(e_i, e_i)$$

Remark Trivially, we can see that $\mathscr{R}_S = 2\Sigma_{1\leq i\leq j\leq n} k(e_i, e_j)$

8 Existence of Metrics

We have mentioned metrics previously and used them in various calculations. However, the question of whether such structures exist was not dealt with previously. Here, we prove that every manifold M has an associated metric (in fact, many metrics).

8.1 Partitions of Unity

We firstly need some topological notions.

Definition 8.1 (Locally Finite) *Let $\mathscr{A} = \{A_\alpha : \alpha \in \Lambda\} \subseteq \mathscr{P}(M)$, where Λ is some index set. If for any $p \in M$, there exists a neighbourhood U of p such that $U \cap A_\alpha \neq \emptyset$ for only finitely many $\alpha \in \Lambda$, then we say that $\mathscr{A}$ is locally finite.*

Definition 8.2 (Partition of Unity) *A Partition of Unity on M is a collection $\mathscr{F} = \{f_\alpha : \alpha \in \Lambda\} \subseteq C_C^\infty(M)$ such that:*

1. *$f_\alpha \geq 0$ for all $\alpha \in \Lambda$*
2. *$\{\text{spt } f_\alpha : \alpha \in \Lambda\}$ is locally finite*
3. *$\Sigma_{\alpha\in\Lambda} f_\alpha(x) = 1$ for all $x \in M$*

Remark It is really due to the locally finite condition that we can make sense of the sum.

Definition 8.3 (Subordinate Parition of Unity) *Let $\mathscr{U} = \{U_\alpha : \alpha \in \Xi\}$ and open cover of M. Then, a partition of unity $\mathscr{F} = \{f_\alpha : \alpha \in \Lambda\}$ is said to be subordinate to $\mathscr{U}$ if for every $\alpha \in \Lambda$, there exists a $\delta \in \Xi$ such that $\text{spt } f_\alpha \subseteq U_\delta$.*

We quote the following important result without proof.

Theorem 8.4 (Existence of Countable Subordinate Partition of Unity) *Let $\mathscr{U}$ be an open covering of M. Then there exists a countable partition of unity $\mathscr{F}$ subordinate to $\mathscr{U}$.*

Corollary 8.5 *Let G open in M. Let $A \subseteq G$ closed in G. Then, there exists a smooth $\phi : M \to [0, 1]$ such that:*

1. *$\phi \equiv 1$ for all $x \in A$*
2. *$\phi \equiv 0$ for all $x \in M \setminus G$.*

Proof Let $\mathscr{C} = \{G, M \setminus A\}$. Trivially $\mathscr{C}$ is an open cover for M. By the theorem, we have a countable partition of unity $\mathscr{F} = \{f_i : i \in \mathbb{N}\}$ subordinate to $\mathscr{C}$. By this subordinate condition, we have that for each $i \in \mathbb{N}$, either $\text{spt } f_i \subseteq G$ or $\text{spt } f_i \subseteq M \setminus A$, but not both.

Define $\phi : M \to [0, 1]$ by:

$$\phi(x) = \Sigma_{\text{spt } f_i \subseteq G} f_i(x)$$

Now, if $x \in A$, then since $A \subseteq G$, we have $\phi(x) = 1$. Again, by construction of ϕ, we have that $\phi(x) = 0$ for all $x \in M \setminus G$ □

We need one further auxiliary result without proof.

Lemma 8.6 *Let V be a vector space. Then the space of positive definite bilinear forms on V is convex.*

Now we can finally prove the existence result for metrics on manifolds.

Corollary 8.7 (Existence of a Metric) *There exists a metric g on a manifold M.*

Proof Let $\mathscr{U} = \{U_\alpha : \psi_\alpha : U_\alpha \to \mathbb{R}^n\}$. By the theorem, we are guaranteed a countable partition of unity $\mathscr{F}$ subordinate to $\mathscr{U}$. That is exactly, for every $i \in \mathbb{N}$, there exists an α_i such that $\mathrm{spt}\, f_i \subseteq U_{\alpha_i}$.

Since U_{α_i} is associated to ψ_{α_i}, we have a coordinate chart. Using ψ_{α_i} we can pull back any metric in $\mathrm{Img}\, \psi_{\alpha_i}$. Let this metric be g_i.

Now, we have that $f_i g_i$ is a $(0, 2)$ tensor field that vanishes outside U_{α_i}. Define:

$$g(x) = \Sigma_i f_i(x) g_i(x)$$

Since f_i come from a partition of unity, at each $x \in M$, we have that $\Sigma_i f_i(x) = 1$, and since the space of metrics is convex, g is a positive definite metric. □

Remark Since we pull back each g_i from $\mathbb{R}^n$, and since there are infinitely many metrics on $\mathbb{R}^n$, for every manifold, there are infinitely many metrics.

Notation

$(M, \mathscr{D})$	Manifold given by set M and differentiable structure $\mathscr{D}$
(M, g)	Riemannian/Lorentzian Manifold M with metric g
$[\sigma(t+s)]$	Tangent vector at $\sigma(s)$
$[\sigma]$	Tangent vector at $\sigma(0)$
Γ_{ij}^k	Christoffel symbols of ∇ in a given coordinate system
$\nabla_X Y$	Covarient Derivative of Y with respect to X
δ_{ij}	Kronecker delta
$\mathrm{D}F(x)$	Jacobian of F evaluated at x
$\mathrm{T}^* M$	Cotangent Bundle
V^*	Dual of the vector space V
$\mathrm{d}f$	The differential of f
$\mathrm{Img}\, F$	Image of F
$\mathrm{Int}\, D$	The interior of the set D
$\mathscr{L}_X Y$	Lie Derivative of Y with respect to X
$\frac{\partial}{\partial x^i}$	Basis Vector for $\mathrm{T}_p M$ for coordinates $\left\{x^i\right\}$
f^*	Pull back by differentiable f
f_*	Push forward by differentiable f
$\mathscr{R}(\cdot, \cdot, \cdot, \cdot)$	Same as $\langle \mathscr{R}(\cdot, \cdot)(\cdot), \cdot \rangle$, the curvature 4-Tensor
$\mathscr{R}(\cdot, \cdot)(\cdot)$	Riemannian Curvature Operator
$\mathbb{R}^n$	n-dimensional Euclidean Space
$\mathbb{RP}^n$	n dimensional Real Projective Plane
$\mathrm{Ric}(X, Y)$	Ricci Curvature
$\mathscr{R}_S$	Sectional Curvature
$\left\{F_t^X\right\}, \left\{F^X(\cdot\,, t)\right\}$	1-parameter family of diffeomorphisms generated by vectorfield X
$\sigma'(t)$	For a curve σ, tangent vector at $\sigma(t)$
$\mathrm{T}^{(r,s)} M$	Bundle of (r, s) Tensors
$\mathrm{T}M$	Tangent Bundle of M
$\mathrm{T}_p M$	Tangent Space to M at p
$\mathscr{X}^k(V)$	Set of C^k vectorfields on V
$C^r(\mathbb{R}^n), C^r$	Same as $C^r(\mathbb{R}^n, \mathbb{R})$
$C^r(\mathbb{R}^n, \mathbb{R}^m)$	r-differentiable $(r > 0)$/continuous $(r = 0)$ functions from $\mathbb{R}^n$ to $\mathbb{R}^m$
$F\vert_V$	Restriction of F to V
f_α	The function $f_\alpha = f \circ \psi_\alpha^{-1}$, for coordinate chart ψ_α
$K \Subset U$	K open in U and $\overline{K} \subseteq U$ and compact
$u \otimes v$	Tensor product of u and v

$V_{(r,s)}$	Vector Space of (r, s) type tensors over vector space V
$X^{\perp}$	Projection of X to the subspace in context
Diffeomorphism	Differentiable bijection with differentiable inverse

Index