Manifolds (CM437Z/CMMS18)

Neil Lambert
Department of Mathematics
King's College London
Strand
London WC2R 2LS, U. K.
Email: lambert@mth.kcl.ac.uk

Contents

Acknowledgement

These notes are indebted to the notes of A. Rogers. The introduction is largely taken from the book by Messiah.

1 Manifolds

1.1 Elementary Topology and Definitions

This section should be a review of concepts (hence it is all definitions and no theorems).

Definition: A topological space X is set whose elements are called points together with a collection $\mathcal{U} = \{U_i\}$ of subsets of X which are called open sets and satisfy
i) $\mathcal{U}$ is closed under finite intersections and arbitrary unions.
ii) $\emptyset, X \in \mathcal{U}$

Definition: A set is closed if its complement in X is open.

Definition: An open cover of X is a collection of open sets whose union is X.

A topology allows us to define a notion of 'local' meaning that something is true or exists in some open set about a point, but not necessarily the whole space or even all open sets about that point. Some other important concepts are:

Definition: X is connected if it is not the union of disjoint open sets.

Definition: A map $f : X \to Y$ between two topological spaces is continuous iff $f^{-1}(V) = \{x \in X | f(x) \in V\}$ is open for any open set $V \subset Y$.

Definition: If $X \subset Y$ where Y is a topological space then X can be made into a topological space too by considering the induced topology: open sets in X are generated by $U \cap X \subset X$ where U is an open set of Y.

Problem: Show that the induced topology indeed satisfies the definition of a topology.

Definition: A Hausdorff space is a topological space with the additional property that 'points can be seperated': for any two distinct points $x, y \in X$ there exists open sets U_x and U_y such that $x \in U_x$, $y \in U_y$ and $U_x \cap U_y = \emptyset$. Hausdorff spaces are also known as T_2 spaces (as there are infact weaker notions of separability).

Non-Hausdorff spaces have various pathologies that we do not want to consider. Therefore in what follows we will take all topological spaces to be Hausdorff unless otherwise mentioned.

Definition: A function $f : X \to Y$ is one-to-one iff $f(x) = f(y)$ implies $x = y$.

This guarantees the existance of a left inverse $f_L^{-1} : f(X) \subset Y \to X$ such that $f_L^{-1} \circ f(x) = x$ for all $x \in X$, since every element in the image $f(X)$ comes from a unique point in X.

Definition: A function $f : X \to Y$ is onto iff $f(X) = Y$, *i.e.* for all $y \in Y$ there exists an $x \in X$ such that $f(x) = y$.

This guarantees the existance of a right inverse $f_R^{-1} : Y \to X$ such that $f \circ f_R^{-1}(y) = y$ for all $y \in Y$, since every element in Y has some $x \in X$ (not necessarily unique) which is mapped to it by f.

Definition: A bijection is a map which is both one-to-one and onto.

Definition: A homeomorhism is a bijection $f : X \to Y$ which is continuous and whose inverse is continuous.

We often have much more structure. For example if there is a notion of distance then the usual topology is that defined by the open balls.

Definition: A metric space is a point set X together with a map $d : X \times X \to \mathbb{R}$ such that
i) $d(x, y) = d(y, x)$
ii) $d(x, y) \geq 0$ with equality iff $x = y$
iii) $d(x, y) \leq d(x, z) + d(z, y)$

The metric topology is then defined by the open balls

$$U_\epsilon(x) = \{y \in X | d(x, y) < \epsilon\} \tag{1.1}$$

where $\epsilon > 0$, along finite intersections and aribtrary unions of open balls. Note that it follows that X is open since it is a union of open balls

$$X = \bigcup_{x \in X} U_\epsilon(x) \tag{1.2}$$

for any $\epsilon > 0$ of your choosing. These spaces are always Hausdorff (property (ii) ensures that any two distinct points have a finite distance between them and hence open balls with ϵ taken to be half this distance will seperate them) and this also implies that $\emptyset$ is open.

In particular we will heavily use $\mathbb{R}^n$ viewed as a metric topological space (with the usual Pythagorian definition of distance).

1.2 Manifolds

Rougly speaking a manifold is a topological space for which one can locally make charts which piece together in a consistent way.

Definition: An n-dimensional chart on $\mathcal{M}$ is a pair (U, ϕ) where U is an open subset of $\mathcal{M}$ and $\phi : U \to \mathbb{R}^n$ is a homeomorphism onto its image $\phi(U) \subseteq \mathbb{R}^n$

Definition: A n-dimensional differentiable structure on $\mathcal{M}$ is a collection of n-dimensional charts (U_i, ϕ_i), $i \in I$ such that
i) $\mathcal{M} = \bigcup_{i \in I} U_i$

ii) For any pair of charts U_i, U_j with $U_i \cap U_j \neq \emptyset$ the map $\phi_j \cdot \phi_i^{-1} : \phi_i(U_i \cap U_j) \rightarrow \phi_j(U_i \cap U_j)$ is $\mathcal{C}^\infty$, *i.e.* all partial derivatives exist up to any order.

iii) We always take a differentiable structure to be a maximal set of charts, *i.e.* the union over all charts which satisfy (i) and (ii).

N.B.: The functions $\phi_j \circ \phi_i^{-1}$ are called transition functions.

Theorem: If $\mathcal{M}$ is connected then n is well defined, *i.e.* all charts have the same value of n.

Proof: Suppose that two charts had different values of n then it is clear that they can't intersect because the map $\phi_j \circ \phi_i^{-1}$ takes a subset of $\mathbb{R}^{n_i}$ to a subset of $\mathbb{R}^{n_j}$ is a diffeomorphism and hence preserves the dimension. Since this is true for all charts we see that $\mathcal{M}$ must split into disjoint charts, at least one for each different value of n. Since a chart is an open set we can therefore write $\mathcal{M}$ as a union over disjoint open sets, one for each value of n. Thus if $\mathcal{M}$ is connected it must have a unique value of n.

Henceforth we will only consider connected topological spaces.

Definition: A differentiable manifold $\mathcal{M}$ of dimension n is a topologcal space together with an n-dimensional differentiable structure.

N.B.: One can also study differentiable structures where the transition functions are only $\mathcal{C}^k$ for some $k > 0$. Alternatively one could replace $\mathbb{R}^n$ by $\mathbb{C}^n$ and demaned that the transition functions are holomorphic. These are therefore a special case of $2n$-dimensional real manifolds. This leads to the beautiful and rich subject of complex differential geometry which we will not have time to consider here.

Example: Trivially $\mathbb{R}^n$ is a n-dimensional manifold. A single chart that covers the whole of $\mathbb{R}^n$ is $(\mathbb{R}^n, id)$ where id is the identity map $id(x) = x$

Example: Any open subset $U \subset \mathbb{R}^n$ is a n-dimensional manifold. Again the single chart (U, id) is sufficient. In fact any open subset U of a manifold $\mathcal{M}$ with charts (U_i, ϕ_i) is also a manifold since one can use the charts $(U \cap U_i, \phi_i)$.

Problem: Why aren't closed subsets of $\mathbb{R}^n$, *e.g.* a disk with boundary or a line in $\mathbb{R}^2$, along with the identity map charts (note that in its own induced topology any subset of $\mathbb{R}^n$ is an open set)?

Example: The Circle is a one-dimensional manifold.

Let $\mathcal{M} = \{(x,y) \in \mathbb{R}^2 | x^2 + y^2 = 1\}$. We need at least two charts say $U_1 = \{(x,y) \in \mathbb{R}^2 | x^2 + y^2 = 1, y > -1/\sqrt{2}\}$ and $U_2 = \{(x,y) \in \mathbb{R}^2 | x^2 + y^2 = 1, y < 1/\sqrt{2}\}$. We let

$$\begin{aligned} \phi_1 : U_1 &\rightarrow \mathbb{R} \quad \phi_1(x,y) = \theta \in (-\frac{\pi}{4}, \frac{5\pi}{4}) \text{ where } (x,y) = (\cos\theta, \sin\theta) \\ \phi_2 : U_2 &\rightarrow \mathbb{R} \quad \phi_2(x,y) = \theta' \in (-\frac{5\pi}{4}, \frac{\pi}{4}) \text{ where } (x,y) = (\cos\theta', \sin\theta') \end{aligned} \tag{1.3}$$

Now

$$
\begin{aligned}
U_1 \cap U_2 &= \{(x,y) \in \mathbb{R}^2 | -\frac{1}{\sqrt{2}} < y < \frac{1}{\sqrt{2}}\} \\
&= V_L \cup V_R \qquad (1.4)
\end{aligned}
$$

where

$$V_L = \{(x,y) \in \mathbb{R}^2 | -\frac{1}{\sqrt{2}} < y < \frac{1}{\sqrt{2}}, x < 0\} \qquad (1.5)$$

and

$$V_R = \{(x,y) \in \mathbb{R}^2 | -\frac{1}{\sqrt{2}} < y < \frac{1}{\sqrt{2}}, x > 0\} \qquad (1.6)$$

Now on $\phi_1(V_L) = (\frac{3\pi}{4}, \frac{5\pi}{4})$, $\phi_2 \circ \phi_1^{-1}(\theta) = \theta - 2\pi$ whereas on $\phi_1(V_R) = (-\frac{\pi}{4}, \frac{\pi}{4})$, $\phi_2 \circ \phi_1^{-1}(\theta) = \theta$. Similarly $\phi_1 \circ \phi_2^{-1}(\theta') = \theta' + 2\pi$ on $\phi_2(V_L) = (-\frac{5\pi}{4}, -\frac{3\pi}{4})$ and $\phi_1 \circ \phi_2^{-1}(\theta') = \theta'$ on $\phi_2(V_L) = (-\frac{\pi}{4}, \frac{\pi}{4})$. These maps are $\mathcal{C}^\infty$ and hence we have a manifold.

Here on the circle we see that locally we can defined a single coordinate, θ which we think of as an angle. But θ is not defined over the whole circle $\theta = 0$ and $\theta = 2\pi$ are the same.

This illustrates a key point: The maps ϕ_i provide coordinates, just like a map in an atlas provides coordinates in the form of longitude and lattitude. However the coordinates will not in general work over the whole of the manifold. For example the surface of the earth can be mapped in an atlas but the notion of latitude and longitude will break down somewhere; at the poles longitude is not defined. Maps that one sees hanging on a wall always break down somewhere (usually at both poles) but an atlas can smoothly cover the whole earth.

Often the U_i are called coordinate neighbourhoods or patches and the ϕ_i are coordinate maps. If $p \in \mathcal{M}$ is some point contained in a given patch U_i then the 'local coordinates' of p are

$$\phi_i(p) = (x^1(p), x^2(p), x^3(p), ..., x^n(p)) \in \mathbb{R}^n \qquad (1.7)$$

Clearly there is a huge amount of choice of local coordinates. Typically in any given patch U_i we could choose from an infinite number of different functions ϕ_i. Furthermore for a given manifold there will be infinitely many choices of open sets U_i which we use to cover it with.

Differential geometry is the study of manifolds and uses tensoriol objects which take into account this huge redundancy in the actual way that we may choose to describe a given manifold. This is the so-called coordinate free approach. Often, especially in older texts, one fixes a covering and coordinate patches and writes any tensor in terms of it values in some given local coordinate system. This may be convenient for some calculational purposes but it obscures the true coordinate independent meaning of the important concepts. In addition it should always be kept in mind when using explicit coordinates that they are almost certainly not valid everywhere. One might often need to change coordinates, either because we prefer to use a different choice of coordinates valid in the same patch, or because we need to transform to a new patch which covers a different portion of the manifold. In this course we will use the coordinate free approach as much as possible.

Example: Let us consider $\mathbb{R}P^n = (\mathbb{R}^{n+1} - \{0\})/\sim$ where $\sim$ is the equivalence relation

$$(x^1, x^2, x^3, ..., x^{n+1}) \sim \lambda(x^1, x^2, x^3, ..., x^{n+1}) \tag{1.8}$$

for any $\lambda \in \mathbb{R} - \{0\}$. We denote an element of the equivalence class by $[x^1, x^2, x^3, ..., x^{n+1}]$. Let us choose for the charts

$$\begin{aligned} U_1 &= \{[x^1, x^2, x^3, ..., x^{n+1}] \in \mathbb{R}P^n | x^1 \neq 0\} \\ U_2 &= \{[x^1, x^2, x^3, ..., x^{n+1}] \in \mathbb{R}P^n | x^2 \neq 0\} \\ U_3 &= \{[x^1, x^2, x^3, ..., x^{n+1}] \in \mathbb{R}P^n | x^3 \neq 0\} \\ &. \\ &. \\ &. \\ U_{n+1} &= \{[x^1, x^2, x^3, ..., x^{n+1}] \in \mathbb{R}P^n | x^{n+1} \neq 0\} \end{aligned} \tag{1.9}$$

with the functions

$$\begin{aligned} \phi_1([x^1, x^2, x^3, ..., x^{n+1}]) &= \left(\frac{x^2}{x^1}, \frac{x^3}{x^1}, \frac{x^4}{x^1}, ..., \frac{x^{n+1}}{x^1}\right) \in \mathbb{R}^n \\ \phi_2([x^1, x^2, x^3, ..., x^{n+1}]) &= \left(\frac{x^1}{x^2}, \frac{x^3}{x^2}, \frac{x^4}{x^2}, ..., \frac{x^{n+1}}{x^2}\right) \in \mathbb{R}^n \\ \phi_3([x^1, x^2, x^3, ..., x^{n+1}]) &= \left(\frac{x^1}{x^3}, \frac{x^2}{x^3}, \frac{x^4}{x^3}, ..., \frac{x^{n+1}}{x^3}\right) \in \mathbb{R}^n \\ &. \\ &. \\ &. \\ \phi_{n+1}([x^1, x^2, x^3, ..., x^{n+1}]) &= \left(\frac{x^1}{x^{n+1}}, \frac{x^2}{x^{n+1}}, \frac{x^3}{x^{n+1}}, ..., \frac{x^n}{x^{n+1}}\right) \in \mathbb{R}^n \end{aligned} \tag{1.10}$$

Clearly $\cup_{i=1}^{n+1} U_i = \mathbb{R}P^n$ and each ϕ_i is a homeomorphism (without loss of generality we can take $x^i = 1$ in U_i). Consider the intersection of U_1 and U_2. Here we can take $x^1 = 1$ and $x^2 = v \neq 0$ so that

$$\phi_1^{-1}(u_1, u_2, u_3, ..., u_n) = [1, u_1, u_2, u_3, ..., u_n] \tag{1.11}$$

and hence

$$\phi_2 \circ \phi_1^{-1}(u_1, u_2, u_3, ..., u_n) = \left(\frac{1}{v}, \frac{u_2}{v}, \frac{u_3}{v}, ..., \frac{u_n}{v}\right) \in \mathbb{R}^n \tag{1.12}$$

Since $v \neq 0$ this map is in $\mathcal{C}^\infty$ on $\phi_1(U_1 \cap U_2)$. All the other intersections follow the same way. Thus $\mathrm{I\!R}P^n$ is an n-dimensional manifold.

Problem: What is $\mathrm{I\!R}P^1$?

Theorem: If $\mathcal{M}$ and $\mathcal{N}$ are m and n-dimensional manifolds respectively then $\mathcal{M} \times \mathcal{N}$ is an $(m+n)$-dimensional manifold.

Proof: Let (U_i, ϕ_i), $i \in I$ be an m-dimensional differential structure for $\mathcal{M}$ and (V_a, ψ_a) $a \in A$ be an n-dimensional differential structure for $\mathcal{N}$. We can construct a differential structure for $\mathcal{M} \times \mathcal{N}$ by taking the following charts:

$$W_{ia} = U_i \times V_a \qquad \chi_{ia} : W_{ia} \to \mathrm{I\!R}^{n+m} \ , \qquad \chi_{ia}(x, y) = (\phi_i(x), \psi_a(y)) \qquad i \in I \ a \in J \tag{1.13}$$

where $x \in \mathcal{M}$ and $y \in \mathcal{N}$. Clearly $\cup_{ia} W_{ia} = \mathcal{M} \times \mathcal{N}$ and χ_{ia} are homeomorphims. It is also clear that the transition functions

$$\chi_{jb} \circ \chi_{ia}^{-1} = (\phi_j \circ \phi_i^{-1}, \psi_b \circ \psi_b^{-1}) \tag{1.14}$$

are $\mathcal{C}^\infty$.

Problem: Show that the following:

$$\begin{aligned}
U_1 &= \{(x,y) \in S^1 | y > 0\} \ , \qquad \phi_1(x,y) = x \\
U_2 &= \{(x,y) \in S^1 | y < 0\} \ , \qquad \phi_2(x,y) = x \\
U_3 &= \{(x,y) \in S^1 | x > 0\} \ , \qquad \phi_3(x,y) = y \\
U_4 &= \{(x,y) \in S^1 | x < 0\} \ , \qquad \phi_4(x,y) = y
\end{aligned} \tag{1.15}$$

are a set of charts which cover S^1.

Problem: Show that the 2-sphere $S^2 = \{(x,y,z) \in \mathrm{I\!R}^3 | x^2 + y^2 + z^2 = 1\}$ is a 2-dimensional manifold.

2 The Tangent Space

An important notion in geometry is that of a tangent vector. This is intuitively familiar for a curve in $\mathrm{I\!R}^n$. But the elementary defintion of a tangent vector, or indeed any vector, relies on special properties of $\mathrm{I\!R}^n$ such a fixed coordinate system and its vector space structure.

Once given a vector $\mathbf{v} = (v^1, v^2, v^3) \in \mathrm{I\!R}^3$ for example, we can consider the derivative in the direction of $\mathbf{v}$:

$$v^1 \frac{\partial}{\partial x^1} + v^2 \frac{\partial}{\partial x^2} + v^3 \frac{\partial}{\partial x^3} \tag{2.16}$$

Here we view this expression as an operator acting on functions $f : \mathbb{R}^3 \to \mathbb{R}$. Changing coordinates, for example by performing a rotation, we also must change the coefficients v^1, v^2, v^3 in an appropriate way however the action on a function remains the same. We need to generalise the notion of a tangent vector to manifolds in a coordinate free way. There are several equiavalent ways to do this but here we will use the identification of a vector field with an operator acting on functions.

Our first step is to introduce differentiable functions on manifolds. We will then proceed to understand vectors as operators on differentiable functions.

2.1 Maps from $\mathcal{M}$ to $\mathbb{R}$

Definition: A function $f : \mathcal{M} \to \mathbb{R}$ is $\mathcal{C}^\infty$ iff for each chart (U_i, ϕ_i) in the differentiable structure of $\mathcal{M}$

$$f \circ \phi_i^{-1} : \phi_i(U_i) \subset \mathbb{R}^n \to \mathbb{R} \tag{2.17}$$

is $\mathcal{C}^\infty$. The set of such functions on a manifold $\mathcal{M}$ is denoted $\mathcal{C}^\infty(\mathcal{M})$.

Note that if $f \circ \phi_i^{-1}$ is $\mathcal{C}^\infty$ and (U_j, ϕ_j) is another chart with $U_i \cap U_j \neq \emptyset$ then $f \circ \phi_j^{-1}$ will be $\mathcal{C}^\infty$ on $U_i \cap U_j$. Thus we need only check that $f \circ \phi_i^{-1}$ is $\mathcal{C}^\infty$ over a set of charts that covers $\mathcal{M}$.

Problem: Consider the circle S^1 as above. Show that $f : S^1 \to \mathbb{R}$ with $f(x, y) = x^2 + y$ is $\mathcal{C}^\infty$.

Definition: An algebra $\mathcal{V}$ is a real vector space along with an operation $\star : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ such that

i) $\mathbf{v} \star \mathbf{0} = \mathbf{0} \star \mathbf{v} = \mathbf{0}$
ii) $(\lambda \mathbf{v}) \star \mathbf{u} = \mathbf{v} \star (\lambda \mathbf{u}) = \lambda(\mathbf{v} \star \mathbf{u})$
iii) $\mathbf{v} \star (\mathbf{u} + \mathbf{w}) = \mathbf{v} \star \mathbf{u} + \mathbf{v} \star \mathbf{w}$
iv) $(\mathbf{u} + \mathbf{w}) \star \mathbf{v} = \mathbf{u} \star \mathbf{v} + \mathbf{w} \star \mathbf{v}$

for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathcal{V}$ and $\lambda \in \mathbb{R}$.

Theorem: $\mathcal{C}^\infty(\mathcal{M})$ is an algebra with addition and multiplication defined pointwise

$$\begin{aligned} (f + g)(p) &= f(p) + g(p) \\ (\lambda f)(p) &= \lambda f(p) \\ (f \star g)(p) &= f(p)g(p) \end{aligned} \tag{2.18}$$

Proof: Let us show that $f \star g$ is in $\mathcal{C}^\infty(\mathcal{M})$. Now

$$(f \star g) \circ \phi_i^{-1} = f \circ \phi_i^{-1} \cdot g \circ \phi_i^{-1} \tag{2.19}$$

where $f \circ \phi_i^{-1}$ and $g \circ \phi_i^{-1}$ are $\mathcal{C}^\infty$. Therefore their pointwise product is too. Hence $(f \star g) \circ \phi_i^{-1}$ is $\mathcal{C}^\infty$ which is what we needed to show.

The other properties can be shown in a similar manor.

Definition: For a point $p \in \mathcal{M}$ we let $\mathcal{C}^\infty(p)$ be the set of functions such that
i) $f : U \to \mathbb{R}$ where $p \in U \subset \mathcal{M}$ and U is an open set.
ii) $f \in \mathcal{C}^\infty(U)$ (recall that an open subset of a manifold is a manifold)

2.2 Tangent Vectors

We can now state our main definition:

Definition: A tangent vector at a point $p \in \mathcal{M}$ is a map $X_p : \mathcal{C}^\infty(p) \to \mathbb{R}$ which satisfies
i) $X_p(f + g) = X_p(f) + X_p(g)$
ii) $X_p(\text{constant map}) = 0$
iii) $X_p(fg) = f(p)X_p(g) + X_p(f)g(p)$ 'Leibniz rule'

The set of tangent vectors at a point $p \in \mathcal{M}$ is called the tangent space to $\mathcal{M}$ at p and is denoted by $T_p\mathcal{M}$.

The union of all tangent spaces to $\mathcal{M}$ is called the tangent bundle

$$T\mathcal{M} = \cup_{p\in\mathcal{M}} T_p\mathcal{M} \tag{2.20}$$

This is an example of a fibre bundle and is itself a $2n$-dimensional manifold.

N.B.: In general objects that satisfy these properties are called derivations.

N.B.: With this definition a tangent vector is a linear map from $\mathcal{C}^\infty(p)$ to $\mathbb{R}$. Since $\mathcal{C}^\infty(p)$ is a vector space (it is an algebra) the tangent vectors are therefore elements of the dual vector space to $\mathcal{C}^\infty(p)$. However $\mathcal{C}^\infty(p)$ and hence its dual are infinite dimensional. The conditions (i),(ii) and (iii) restrict the possible linear maps that we identify as tangent vectors and in fact we will see that they become a finite dimensional vector space.

Example: Consider $\mathbb{R}^n$ as a manifold with its obvious chart $U = \mathbb{R}^n$, $\phi : U \to \mathbb{R}^n$ taken to be the identity. Then

$$X = \sum_{\mu=1}^{n} \lambda^\mu \frac{\partial}{\partial x^\mu} \tag{2.21}$$

is a tangent vector. In fact we will learn that all tangent vectors have this form.

In what follows we will denote

$$\frac{\partial f}{\partial x^\mu} = \partial_\mu f \ , \qquad \frac{\partial^2 f}{\partial x^\mu \partial x^\nu} = \partial_\mu \partial_\nu f \qquad etc.. \tag{2.22}$$

where f is defined on some open set in $\mathbb{R}^n$. We can extend this to manifolds by the following:

Definition: Let $(x^1, ..., x^n)$ be local coordinates about a point $p \in \mathcal{M}$. That is there exists a chart (U_i, ϕ_i) with $p \in U_i$ and $\phi_i(q) = (x^1(q), x^2(q), ..., x^n(q))$ for all $q \in U_i$. We define

$$\frac{\partial}{\partial x^\mu}\Big|_p : \mathcal{C}^\infty(p) \to \mathbb{R} \tag{2.23}$$

by

$$\frac{\partial}{\partial x^\mu}\Big|_p f = \partial_\mu(f \circ \phi_i^{-1})(x^1(p), ..., x^n(p)) = \partial_\mu(f \circ \phi_i^{-1}) \circ \phi_i(p) \tag{2.24}$$

Theorem: $\frac{\partial}{\partial x^\mu}\Big|_p$ is a tangent vector to $\mathcal{M}$ at p.

Proof: Let $f, g \in \mathcal{C}^\infty(p)$ be defined on a common open set U in $\mathcal{M}$ that contains p then

$$\begin{aligned}
\frac{\partial}{\partial x^\mu}\Big|_p(f+g) &= \partial_\mu((f+g) \circ \phi_i^{-1})(x^1(p), ..., x^n(p)) \\
&= \partial_\mu(f \circ \phi_i^{-1} + g \circ \phi_i^{-1})(x^1(p), ..., x^n(p)) \\
&= \partial_\mu(f \circ \phi_i^{-1})(x^1(p), ..., x^n(p)) + \partial_\mu(g \circ \phi_i^{-1})(x^1(p), ..., x^n(p)) \\
&= \frac{\partial}{\partial x^\mu}\Big|_p(f) + \frac{\partial}{\partial x^\mu}\Big|_p(g)
\end{aligned} \tag{2.25}$$

It is clear that $\frac{\partial}{\partial x^\mu}\Big|_p(\text{constant map}) = 0$.

$$\begin{aligned}
\frac{\partial}{\partial x^\mu}\Big|_p(fg) &= \partial_\mu((fg) \circ \phi_i^{-1})(x^1(p), ..., x^n(p)) \\
&= \partial_\mu(f \circ \phi_i^{-1} \cdot g \circ \phi_i^{-1})(x^1(p), ..., x^n(p)) \\
&= \partial_\mu(f \circ \phi_i^{-1})(x^1(p), ..., x^n(p))(g \circ \phi_i^{-1})(x^1(p), ..., x^n(p)) \\
&\quad +(f \circ \phi_i^{-1}(x^1(p), ..., x^n(p))\partial_\mu(g \circ \phi_i^{-1})(x^1(p), ..., x^n(p)) \\
&= \frac{\partial}{\partial x^\mu}\Big|_p(f)g(p) + f(p)\frac{\partial}{\partial x^\mu}\Big|_p(g)
\end{aligned} \tag{2.26}$$

Of course we will show that all tangent vectors arise in this way.

Example: Consider $\mathbb{R}P^1 = (\mathbb{R}^2 - \{\mathbf{0}\})/\sim$ where $(x, y) \sim \lambda(x, y),\ \lambda \neq 0$.
We have two charts

$$\begin{aligned}
U_t &= \{(x, y)|x \neq 0\}\ , \quad t = \phi_t([x^1, x^2]) = \frac{x^2}{x^1} \\
U_s &= \{(x, y)|y \neq 0\}\ , \quad s = \phi_s([x^1, x^2]) = \frac{x^1}{x^2}
\end{aligned} \tag{2.27}$$

thus on the intersection $U_s \cap U_t$, $s = 1/t$. Let $p = [1, 3]$ and consider the tangent vector

$$X : \mathcal{C}^\infty(p) \to \mathbb{R} \; , \qquad X(f) = \frac{\partial}{\partial t}\Big|_p f = \frac{d}{dt}(f \circ \phi_t^{-1})(t = 3) \tag{2.28}$$

How does X act in the other coordinate system (where they overlap)? Recall that

$$\frac{d}{dt}(f(t)) = \frac{ds(t)}{dt}\frac{d}{ds}(f(t(s))) \tag{2.29}$$

Now $s(t) = \phi_s \circ \phi_t^{-1}$ and $t(s) = \phi_t \circ \phi_s^{-1}$ so we have that

$$\begin{aligned} X(f) &= \frac{d}{dt}(f \circ \phi_t^{-1})(t = 3) \\ &= \frac{ds(t)}{dt}\frac{d}{ds}(f \circ \phi_t^{-1}(\phi_t \circ \phi_s^{-1}))(s = \frac{1}{3}) \\ &= -\frac{1}{t^2}\frac{d}{ds}(f \circ \phi_s^{-1})\Big|_{s=\frac{1}{3}} \\ &= -\frac{1}{9}\frac{d}{ds}(f \circ \phi_s^{-1})\Big|_{s=\frac{1}{3}} \end{aligned} \tag{2.30}$$

Thus the tangent vector can look rather different depending on the coordinate system one choses. However its definition as a linear map from $\mathcal{C}^\infty(p)$ to $\mathbb{R}$ is independent of the choice of coordinates, *i.e.* (2.28) and (2.30) will agree on any function in $\mathcal{C}^\infty(p)$.

Lemma: Let $(x^1, ..., x^n)$ be a coordinate system about $p \in \mathcal{M}$. Then for every function $f \in \mathcal{C}^\infty(p)$ there exist n functions $f_1, ..., f_n$ in $\mathcal{C}^\infty(p)$ with

$$f_\mu(p) = \frac{\partial}{\partial x^\mu}\Big|_p f \tag{2.31}$$

and, where defined,

$$f(q) = f(p) + \sum_\mu (x^\mu(q) - x^\mu(p)) f_\mu(q) \tag{2.32}$$

Proof: Let $F = f \circ \phi_i^{-1}$ which is defined on $V = \phi_i(U \cap U_i)$ where U is the domain of f. Let B be an open ball in $V \subset \mathbb{R}^n$ centred on $\mathbf{v} = \phi_i(p)$. Take $\mathbf{y} \in B$

$$\begin{aligned} F(y^1, ..., y^n) &= F(y^1, ..., y^n) - F(y^1, ..., y^{n-1}, v^n) \\ &\quad + F(y^1, ..., y^{n-1}, v^n) - F(y^1, ..., v^{n-1}, v^n) + ... \\ &\quad ... + F(y^1, v^2 ..., v^n) - F(v^1, ..., v^n) + F(v^1, ..., v^n) \\ &= F(v^1, ..., v^n) + \sum_\mu \Big(F(y^1, ..., y^\mu, v^{\mu+1}, ..., v^n) - F(y^1, ..., v^\mu, v^{\mu+1}, ..., v^n) \Big) \\ &= F(v^1, ..., v^n) + \sum_\mu F(y^1, ..., v^\mu + t(y^\mu - v^\mu), v^{\mu+1}, ..., v^n)\Big|_{t=0}^{t=1} \end{aligned}$$

$$
\begin{aligned}
&= F(v^1,...,v^n) + \sum_\mu \int_0^1 \frac{d}{dt} F(y^1,...,v^\mu + t(y^\mu - v^\mu), v^{\mu+1},...,v^n)dt \\
&= F(v^1,...,v^n) + \sum_\mu \int_0^1 \partial_\mu F(y^1,...,y^{\mu-1}, v^\mu + t(y^\mu - v^\mu), v^{\mu+1},...,v^n)(y^\mu - v^\mu)dt
\end{aligned}
\tag{2.33}
$$

If we let

$$
F_\mu(y^1,...,y^n) = \int_0^1 \partial_\mu F(y^1,...,v^\mu + t(y^\mu - v^\mu), v^{\mu+1},...,v^n)dt \tag{2.34}
$$

then we have that

$$
F(y^1,...,y^n) = F(v^1,...,v^n) + \sum_\mu (y^\mu - v^\mu) F_\mu(y^1,...,y^n) \tag{2.35}
$$

Finally we recall that $F = f \circ \phi_i^{-1}$ so that if we let $(y^1,...,y^n) = \phi_i(q) = (x^1(q),...,x^n(q))$ then we have

$$
f \circ \phi_i^{-1} \circ \phi_i(q) = f \circ \phi_i^{-1} \circ \phi_i(p) + \sum_\mu (x^\mu(q) - x^\mu(p)) F_\mu \circ \phi_i(q) \tag{2.36}
$$

thus

$$
f(q) = f(p) + \sum_\mu (x^\mu(q) - x^\mu(p)) f_\mu(q) \tag{2.37}
$$

where we identify $f_\mu = F_\mu \circ \phi_i$.

It also follows that

$$
\begin{aligned}
\left.\frac{\partial}{\partial x^\mu}\right|_p f &= \frac{\partial (f \circ \phi_i^{-1})}{\partial x^\mu} \circ \phi_i(p) \\
&= \frac{\partial F}{\partial y^\mu} \circ \phi_i(p) \\
&= \frac{\partial}{\partial y^\mu} \left(F(v^1,...,v^n) + \sum_\nu (y^\nu - v^\nu) F_\nu(y^1,...,y^n) \right)_{\mathbf{y}=\phi_i(p)} \\
&= \left(F_\mu(y^1,...,y^n) + \sum_\nu (y^\nu - v^\nu) \partial_\mu F_\nu(y^1,...,y^n) \right)_{\mathbf{y}=\phi_i(p)} \\
&= F_\mu(\phi_i(p)) \\
&= f_\mu(p)
\end{aligned}
\tag{2.38}
$$

Theorem: $T_p(\mathcal{M})$ is an n-dimensional real vector space and a set of basis vectors is

$$
\left\{ \left.\frac{\partial}{\partial x^\mu}\right|_p ,\ \mu = 1,...,n \right\} \tag{2.39}
$$

i.e. a general element of $T_p(\mathcal{M})$ can be written as

$$X_p = \sum_\mu \lambda^\mu \frac{\partial}{\partial x^\mu}\Big|_p \tag{2.40}$$

Proof:

First we note that if X_p and Y_p are two tangent vectors at a point $p \in \mathcal{M}$ then we can add them or multiply them by a number $\lambda \in \mathbb{R}$:

i) $(X_p + Y_p) : \mathcal{C}^\infty \to \mathbb{R}$, $\quad f \to X_p f + Y_p f$
ii) $(\lambda X_p) : \mathcal{C}^\infty \to \mathbb{R}$, $\quad f \to \lambda(X_p f)$

(Convince yourself of this.)

Thus $T_p\mathcal{M}$ is a real vector space.

From the above lemma we have that

$$X_p(f) = X_p\left(f(p) + \sum_\mu (x^\mu - x^\mu(p)) f_\mu\right) \tag{2.41}$$

Now $X_p(f(p)) = X_p(x^\mu(p)) = 0$ since $f(p)$ and $x^\mu(p)$ are constants. Thus we have that

$$\begin{aligned} X_p(f) &= \sum_\mu \left((x^\mu(q) - x^\mu(p)) X_p(f_\mu) + X_p(x^\mu) f_\mu(q)\right)|_{q=p} \\ &= \sum_\mu X_p(x^\mu) f_\mu(p) \\ &= \sum_\mu X_p(x^\mu) \frac{\partial}{\partial x^\mu}\Big|_p f \end{aligned} \tag{2.42}$$

where we used (2.38). This shows that the elements (2.39) span the tangent space. We must now show that they are linearly independent. To this end suppose that

$$\sum_\mu \lambda^\mu \frac{\partial}{\partial x^\mu}\Big|_p = 0 \tag{2.43}$$

The coordinate functions are in $\mathcal{C}^\infty(p)$ so we may consider

$$0 = \sum_\mu \lambda^\mu \frac{\partial}{\partial x^\mu}\Big|_p x^\nu = \sum_\mu \lambda^\mu \frac{\partial}{\partial x^\mu}(x^\nu) = \sum_\mu \lambda^\mu \delta_\mu^\nu = \lambda^\nu \tag{2.44}$$

Thus $\lambda^\mu = 0$ for all μ.

So why are they called tangent vectors? First we consider $\mathbb{R}^n$ consider a curve $C : (0,1) \to \mathbb{R}^n$. We recall from elementary geometry that the tangent vector to a point $p = C(t_1)$ is a line through p in the direction (*i.e.* with the slope)

$$\left(\frac{dC^1}{dt}\Big|_{t=t_1}, ..., \frac{dC^n}{dt}\Big|_{t=t_1} \right) \tag{2.45}$$

where $C(t) = (C^1(t), ..., C^n(t)) \in \mathbb{R}^n$ is $\mathcal{C}^\infty$.

Now if $f \in \mathcal{C}^\infty(p)$ then, by our definition, $X_p : \mathcal{C}^\infty(p) \to \mathbb{R}$ defined by

$$X_p(f) = \frac{d}{dt} f(C(t))\Big|_{t=t_1} \tag{2.46}$$

is a tangent vector to $\mathbb{R}^n$ at $p = C(t_1)$, *i.e.* it acts linearly on the function f, vanishes on constant functions and satisfies the Leibniz rule. On the other hand we also have

$$X_p(f) = \frac{dC^\mu}{dt}\Big|_{t=t_1} \partial_\mu (f(C(t))|_{t=t_1} \tag{2.47}$$

Thus dC^μ/dt are the components of X_p in the basis

$$\frac{\partial}{\partial x^\mu}\Big|_p \tag{2.48}$$

So we have suceeded in generalising the definition (2.16)

2.3 Curves and their Tangents

We can now discuss curves on manifolds and their tangents.

Definition: Consider an open interval $(a, b) \in \mathbb{R}$ a map $C : (a, b) \to \mathcal{M}$ to a manifold $\mathcal{M}$ is called a smooth curve on $\mathcal{M}$ if $\phi_i \circ C$ is $\mathcal{C}^\infty$ whereever it is defininned for any chart (U_i, ϕ_i) of $\mathcal{M}$ (*i.e.* with $U_i \cap C((a, b)) \neq \emptyset$).

Definition: For a point $t_1 \in (a, b)$ with $C(t_1) = p$ we can define the tangent vector $T_p(C) \in T_p(\mathcal{M})$ to the curve C at p by

$$T_p(C)(f) = \frac{d}{dt} f(C(t))\Big|_{t=t_1} \tag{2.49}$$

It should be clear that $T_p(C) \in T_p\mathcal{M}$.

Let $p \in \mathcal{M}$ be a point on a curve C at $t = t_1$ which is covered by a chart (U_i, ϕ_i). Then there is some $\epsilon > 0$ such that

$$C((t_1 - \epsilon, t_1 + \epsilon)) \subset U_i \tag{2.50}$$

We can express the tangent to C at p as

$$\begin{aligned} T_p(C)(f) &= \frac{df(C(t))}{dt}\Big|_{t=t_1} \\ &= \frac{d}{dt}\left(f\circ\phi_i^{-1}\circ\phi_i\circ C(t)\right)\Big|_{t=t_1} \\ &= \sum_{\mu=1}^{n}\frac{d(\phi^\mu\circ C(t))}{dt}\Big|_{t=t_1}\partial_\mu(f\circ\phi_i^{-1})(\phi_i(C(t_1)) \end{aligned} \tag{2.51}$$

Here we have split up $f\circ C : (t_1-\epsilon, t_1+\epsilon)\to \mathbb{R}$ as the composition of a function $\phi_i\circ C : (t_1-\epsilon, t_1+\epsilon)\to\mathbb{R}^n$ and a function $f\circ\phi_i^{-1}:\mathbb{R}^n\to\mathbb{R}$ and used the chain rule. Thus we have

$$T_p(C)(f) = \sum_{\mu=1}^{n}\frac{d(\phi^\mu\circ C(t))}{dt}\Big|_{t=t_1}\frac{\partial}{\partial x^\mu}\Big|_p f \tag{2.52}$$

so that

$$T_p(C) = \sum_{\mu=1}^{n}\frac{d(\phi^\mu\circ C(t))}{dt}\Big|_{t=t_1}\frac{\partial}{\partial x^\mu}\Big|_p \tag{2.53}$$

Conversely suppose that T_p is a tangent vector to $\mathcal{M}$ at p. We will now show that we can construct a curve through p such that T_p is its tangent at p. Let $(x^1,...,x^n)=\phi_i(q)$ be local coordinates about p defined on an open set U_i. Hence we may write

$$T_p = \sum_{\mu=1}^{n}T^\mu\frac{\partial}{\partial x^\mu}\Big|_p \tag{2.54}$$

for some numbers T^μ. We define $C:(-\epsilon,\epsilon)\to\mathcal{M}$ by

$$\phi_i^\mu\circ C(t) = x^\mu(p)+tT^\mu \tag{2.55}$$

where we pick ϵ sufficiently small so that $\phi_i(C(-\epsilon,\epsilon))\in U_i$. It follows that

$$\begin{aligned} \frac{d}{dt}f(C(t))\Big|_{t=0} &= \frac{d}{dt}\left((f\circ\phi_i^{-1})\circ(\phi_i\circ C)(t)\right)\Big|_{t=0} \\ &= \frac{d(\phi_i^\mu\circ C)}{dt}\Big|_{t=0}\partial_\mu(f\circ\phi_i^{-1})(\phi_i(C(0)) \\ &= T^\mu\frac{\partial}{\partial x^\mu}\Big|_p f \end{aligned} \tag{2.56}$$

We have therefore shown that all curves through $p\in M$ define a tangent vector to $\mathcal{M}$ at p and that conversely all tangent vectors to $\mathcal{M}$ at p can be realised as the tangent vector to some curve C. However this correspondence is not unique. Clearly many distinct curves may have the same tangent vector at p and conversely the construction of the curve through p with tangent T_p was not unique. But this does lead to the following theorem:

Theorem: $T_p\mathcal{M}$ is isomorphic to the set of all curves through $p \in M$ modulo the equivalence relation that $C(t) \sim C'(t)$ iff

$$\frac{d(f \circ C)}{dt}\Big|_{t=t_1} = \frac{d(f \circ C')}{dt}\Big|_{t=t'_1} \tag{2.57}$$

for all $f \in \mathcal{C}^\infty(p)$ where $C(t_1) = C'(t'_1) = p$.

Proof: It is easy to check that the construction above provides a bijection between these two spaces. The more difficult part, which we won't go into here, is to show that the vector space structure is presesved. Indeed to do this we'd need to give a vector space structure to the equivalance class of curves through p.

Theorem: Let $(x^1, ..., x^n) = \phi_1^\mu$ and $(y^1, .., y^n) = \phi_2^\mu$ be two coordinate systems at a point $p \in M$ with $U_1 \cap U_2 \neq \emptyset$ and suppose that $X_p \in T_p\mathcal{M}$. If

$$X_p = \sum_{\mu=1}^{n} A^\mu \frac{\partial}{\partial x^\mu}\Big|_p \qquad \text{and} \qquad X_p = \sum_{\mu=1}^{n} B^\mu \frac{\partial}{\partial y^\mu}\Big|_p \tag{2.58}$$

then

$$B^\mu = \sum_{\nu=1}^{n} A^\nu \frac{\partial}{\partial x^\nu}\Big|_p y^\mu \tag{2.59}$$

where $y^\mu(x^1, ..., x^n) = \phi_2 \circ \phi_1^{-1}$ is a smooth function from $\phi_1(U_1) \subset \mathbb{R}^n \to \phi_2(U_2) \subset \mathbb{R}^n$

Proof: We have in the second coordinate system that

$$X_p(y^\mu) = \sum_{\nu=1}^{n} B^\nu \frac{\partial}{\partial y^\nu}\Big|_p y^\mu = B^\mu \tag{2.60}$$

but in the first coordinate system we see that

$$X_p(y^\mu) = \sum_{\nu=1}^{n} A^\nu \frac{\partial}{\partial x^\nu}\Big|_p y^\mu \tag{2.61}$$

since these must agree we prove the theorem.

N.B.: This formula is often simply written as

$$B^\mu = \sum_{\nu=1}^{n} A^\nu \frac{\partial y^\mu}{\partial x^\nu} \tag{2.62}$$

or even

$$A'^\mu = A^\nu \frac{\partial x'^\mu}{\partial x^\nu} \tag{2.63}$$

with a sum over ν understood and a prime denoting quantities in the new coordinate system.

3 Maps Between Manifolds

3.1 Diffeomorhisms

Definition: Suppose that $f : \mathcal{M} \to \mathcal{N}$ where $\mathcal{M}, (U_i, \phi_i), i \in I$ and $\mathcal{N}, (V_a, \psi_a), a \in A$ are two manifolds. We say that f is $\mathcal{C}^\infty$ iff

$$\psi_a \circ f \circ \phi_i^{-1} : \phi_i(f^{-1}(V_a)) \to \mathbb{R}^n \tag{3.64}$$

is $\mathcal{C}^\infty$ for all $i \in I$ and $a \in A$.

Problem: Show that $f : S^1 \to S^1$ defined by $f(e^{2\pi i\theta}) = e^{2\pi in\theta}$ is $\mathcal{C}^\infty$ for any n.

Theorem: Suppose that $\mathcal{M}$,$\mathcal{N}$ and $\mathcal{P}$ are manifolds with $f : \mathcal{M} \to \mathcal{N}$ and $g : \mathcal{N} \to \mathcal{P}$ $\mathcal{C}^\infty$. Then $g \circ f : \mathcal{M} \to \mathcal{P}$ is also $\mathcal{C}^\infty$.

Proof: this follows from the chain rule.

Definition: If $f : \mathcal{M} \to \mathcal{N}$ is a bijection with both f and f^{-1} C^∞ then f is called a diffeomorphism.

Two manifolds with a diffeomorphic, *i.e.* for which there exists a difformorphism between them, are equivalent geometrically.

Problem: Show that the charts of two diffeomorphic manifolds are in a one to one correspondence.

Problem: Show that the set of diffeomorphisms from a manifold to itself forms a group under composition.

3.2 Push Forward

Theorem: Let $f : \mathcal{M} \to \mathcal{N}$ be $\mathcal{C}^\infty$. If $X_p \in T_p\mathcal{M}$ then

$$f_\star X : \mathcal{C}^\infty(f(p)) \to \mathbb{R} \quad \text{defined by} \quad g \to X_p(g \circ f) \tag{3.65}$$

is a tangent vector to $\mathcal{N}$ at $f(p)$.

Proof: Let $g_1, g_2 \in \mathcal{C}^\infty(f(p))$ and $\lambda \in \mathbb{R}$.

Then

$$\begin{aligned} f_\star X_{f(p)}(g_1 + g_2) &= X_p((g_1 + g_2) \circ f) \\ &= X_p(g_1 \circ f + g_2 \circ f) \\ &= X_p(g_1 \circ f) + X_p(g_2 \circ f) \\ &= f_\star X_{f(p)}(g_1) + f_\star X_{f(p)}(g_2) \end{aligned} \tag{3.66}$$

and

$$f_\star X_{f(p)}(\lambda) = X_p(\lambda \circ f) = X_p(\lambda) = 0 \tag{3.67}$$

Finally we have

$$\begin{aligned} f_\star X_{f(p)}(g_1 \cdot g_2) &= X_p((g_1 \cdot g_2) \circ f) \\ &= X_p((g_1 \circ f) \cdot (g_2 \circ f)) \\ &= X_p(g_1 \circ f)(g_2 \circ f) + (g_1 \circ f) X_p(g_2 \circ f) \\ &= f_\star X_{f(p)}(g_1) g_2(f(p)) + g_1(f(p)) f_\star X_{f(p)}(g_2) \end{aligned} \tag{3.68}$$

Definition: $f_\star X_p$ is called the push forward of X_p.

Theorem: Suppose that $f : \mathcal{M} \to \mathcal{N}$ is $\mathcal{C}^\infty$, $(x^1, ..., x^m)$ are local coordinates for a point $p \in \mathcal{M}$ and $y^1, ..., y^n$ are local coordinates for the image $f(p) \in \mathcal{N}$. If

$$X_p = \sum_{\mu=1}^{m} A^\mu \frac{\partial}{\partial x^\mu}\Big|_p \tag{3.69}$$

is in $T_p\mathcal{M}$ then

$$f_\star X = \sum_{\mu=1}^{m} \sum_{\nu=1}^{n} A^\mu \frac{\partial}{\partial x^\mu}\Big|_p (y^\nu \circ f) \cdot \frac{\partial}{\partial y^\nu}\Big|_{f(p)} \tag{3.70}$$

Proof: Let $g \in \mathcal{C}^\infty(f(p))$. Then

$$\begin{aligned} f_\star X(g) &= X_p(g \circ f) \\ &= \sum_{\mu=1}^{m} A^\mu \frac{\partial}{\partial x^\mu}\Big|_p (g \circ f) \\ &= \sum_{\mu=1}^{m} A^\mu \frac{\partial(g \circ f \circ \phi_i^{-1})}{\partial x^\mu} \circ \phi_i(p) \\ &= \sum_{\mu=1}^{m} A^\mu \frac{\partial(g \circ \psi_a^{-1} \circ \psi_a \circ f \circ \phi_i^{-1})}{\partial x^\mu} \circ \phi_i(p) \\ &= \sum_{\mu=1}^{m} \sum_{\nu=1}^{n} A^\mu \frac{\partial(\psi_a^\nu \circ f \circ \phi_i^{-1})}{\partial x^\mu}(\phi_i(p)) \frac{\partial(g \circ \psi_a^{-1})}{\partial y^\nu}(\psi_a(f(p)) \\ &= \sum_{\mu=1}^{m} \sum_{\nu=1}^{n} A^\mu \frac{\partial}{\partial x^\mu}\Big|_p (y^\nu \circ f) \cdot \frac{\partial}{\partial y^\nu}\Big|_{f(p)} (g) \end{aligned} \tag{3.71}$$

4 Vector Fields

4.1 Vector Fields

Next we consider vector fields

Definition: A vector field is a map $X : \mathcal{M} \to T\mathcal{M}$ such that $X(p) = X_p \in T_p\mathcal{M}$ and for all $f \in \mathcal{C}^\infty(\mathcal{M})$ the mapping

$$p \to X(f)(p) \tag{4.72}$$

is $\mathcal{C}^\infty$. For vector fields we will drop the explicit supscript p.

Thus a vector field assigns, in a smooth way, a vector in $T_p\mathcal{M}$ of to each point $p \in \mathcal{M}$.

N.B.: As we defined it a vector field is valid over all of $\mathcal{M}$ however it can also be defined over an open subset $U \subset \mathcal{M}$.

Is the product of two vector fields a vector feild? To check this we consider two vector fields X, Y and $f, g \in \mathcal{C}^\infty(\mathcal{M})$. With multiplication of vectors taken to mean $X \cdot Y(f) = X(Y(f))$ then we see that indeed

$$\begin{aligned} X(Y(f+g)) &= X(Y(f) + X(g)) = X(Y(f)) + X(Y(g)) \\ X(Y(\text{constant map})) &= X(\mathbf{0}) = 0 \end{aligned} \tag{4.73}$$

where $\mathbf{0}$ is the constant map that takes all points to zero. However we find

$$\begin{aligned} X(Y(f \cdot g)) &= X(Y_p(f)g + fY(g)) \\ &= X(Y(f))g + Y(f)X(g) + X(f)Y(g) + fX(Y(g)) \end{aligned} \tag{4.74}$$

and this is not a vector field (due to the middle two terms). But we can construct a vector field by taking

$$[X, Y](f) = X(Y(f)) - Y(X(f)) \tag{4.75}$$

since we see that

$$\begin{aligned} [X, Y](f \cdot g) &= X(Y(f))g + Y(f)X(g) + X(f)Y(g) + fX(Y(g)) \\ &\quad -Y(X(f))g - X(f)Y(g) - Y(f)X(g) - fY(X(g)) \\ &= [X, Y](f)g + f[X, Y](g) \end{aligned} \tag{4.76}$$

so that $[X, Y]$ is a vector field.

Definition: $[X, Y]$ is called the commutator of two vector fields.

Problem: What goes wrong if try to define $(X \cdot Y)(f) = X(f) \cdot Y(f)$?

Problem: Show that, if in a particular coordinate system,

$$X = \sum_\mu X^\mu(x)\frac{\partial}{\partial x^\mu}\Big|_p , \qquad Y = \sum_\mu Y^\mu(x)\frac{\partial}{\partial x^\mu}\Big|_p \tag{4.77}$$

then

$$[X,Y] = \sum_\mu\sum_\nu (X^\mu\partial_\mu Y^\nu - Y^\mu\partial_\mu X^\nu)\frac{\partial}{\partial x^\nu}\Big|_p \tag{4.78}$$

Theorem: With the product of two vector fields defined as the commutator the space of vector fields is an algebra.

Proof: We have already seen that $T_p\mathcal{M}$ is a vector space for a particular $p \in \mathcal{M}$. This clearly ensures that the space of vector fields is a vector space (with addition and scalar mutliplication defined pointwise). We need to check the conditions (i) - (iv) in the definition of an algebra. Conditions (i) and (ii) are obviously satisfied. In addition since $[X,Y] = -[Y,X]$ we need only check that

$$\begin{aligned}
[X, Y+Z](f) &= X(Y+Z)(f) - (Y+Z)(X)(f) \\
&= X(Y)(f) + X(Z(f) - Y(X)(f) - Z(X)(f) \\
&= [X,Y](f) + [X,Z](f)
\end{aligned} \tag{4.79}$$

Problem: Show that for three vector fields X, Y, Z on $\mathcal{M}$ the jacobi identity holds:

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0 \tag{4.80}$$

Problem: Consider a manifold with a local coordinate system $U_i, \phi_i = (x^1, ..., x^n)$. Therefore, in U_i we can simply write

$$\frac{\partial}{\partial x^\mu}\Big|_p = \frac{\partial}{\partial x^\mu} \tag{4.81}$$

i) Show that $\left[\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right] = 0$
ii) Evaluate $\left[\frac{\partial}{\partial x^1}, \varphi(x^1,x^2)\frac{\partial}{\partial x^2}\right]$ where $\varphi(x^1,x^2)$ is a $\mathcal{C}^\infty$ function of x^1, x^2.

4.2 Integral and Local Flows

Given a vector field X we can construct curves that pass through $p \in \mathcal{M}$ for which the tangent vector at p is X.

Definition: Let X be a vector field on $\mathcal{M}$ and consider a point $p \in \mathcal{M}$. An integral curve of X passing through p is a curve $C(t)$ in $\mathcal{M}$ such that

$$C(0) = p \qquad \text{and} \qquad C_\star\left(\frac{d}{dt}\right) = X_{C(t)} \tag{4.82}$$

for all t in some open interval $(-\epsilon, \epsilon) \subset \mathbb{R}$. Here we are viewing d/dt as a vector field on $\mathbb{R}$ so that

$$C_\star \left(\frac{d}{dt}\right)(f) = \frac{d}{dt} f(C(t)) = T_p(C) \tag{4.83}$$

is just the tangent vector to $C(t)$ at $p = C(t_1)$.

If we introduce a local coordinate system so that, in a open set about $p \in \mathcal{M}$,

$$X = \sum_\mu X^\mu(x) \frac{\partial}{\partial x^\mu}\Big|_{\phi^{-1}(x)} \tag{4.84}$$

then we find the integral curve is

$$\begin{aligned} C_\star \left(\frac{d}{dt}\right)(f) &= \frac{d}{dt} f(C(t)) \\ &= \frac{d}{dt}(f \circ \phi_i^{-1} \circ \phi_i \circ C(t)) \\ &= \sum_\mu \frac{dC^\mu}{dt}(t) \frac{\partial}{\partial x^\mu}\Big|_{C(t)}(f) \end{aligned} \tag{4.85}$$

where, as before, we have $C^\mu = \phi_i^\mu \circ C$. On the other hand we have

$$X(C(t))(f) = \sum_\mu X^\mu(C(t)) \frac{\partial}{\partial x^\mu}\Big|_{C(t)}(f) \tag{4.86}$$

Thus we see that the condition for an integral curve is a first order differential equation for the coordinates of the curve $C^\mu(t)$

$$\frac{d}{dt} C^\mu(t) = X^\mu(C(t)) \tag{4.87}$$

with the initial condition $x^\mu(C(0)) = x^\mu(p)$. This is a first order differential equation and, as such it has a unique solution with the given initial condition. However it is not at all clear whether or not the solution can be extended to all values of t. In partcular even if there is a solution to the differential equation (4.87) for al t one must worry about patching solutions together over the different coordinate patches. This leads to

Definition: A vector field X on $\mathcal{M}$ is complete if for every point $p \in \mathcal{M}$ the integral curve of X can be extended to a curve on $\mathcal{M}$ for all values of t.

Theorem: If $\mathcal{M}$ is compact, *i.e.* all open covers have a finite subcover, then all vector fields on $\mathcal{M}$ are complete.

Proof: We will not take the time to prove this theorem in this course.

Let $\sigma(t,p)$ be an integral curve of a vector field X that passes through p at $t=0$. This therefore satisfies

$$\frac{d}{dt}\sigma^\mu(t,p) = X^\mu(\sigma(t,p)) \tag{4.88}$$

along with the initial condition $\sigma(0,p)=p$. Thus we have found the following, at least locally (and globally for complete vector fields):

Definition: The flow generated by a vector field X is a differentiable map

$$\sigma : \mathbb{R} \times \mathcal{M} \to \mathcal{M} \tag{4.89}$$

such that

i) at each point $p \in \mathcal{M}$ the tangent to the curve $C_p(t) = \sigma(t,p)$ at p is X

ii) $\sigma(0,p) = p$

iii) $\sigma(t+s,p) = \sigma(t,\sigma(s,p))$

We must show that the third property holds. To do this we simply note that since (4.88) is a first order ordinary differential equation it has a unique solution for a fixed initial condition. So consider

$$\begin{aligned}\frac{d}{dt}\sigma^\mu(t+s,p) &= \frac{d\sigma^\mu(t+s,p)}{d(t+s)}\\ &= X^\mu(\sigma(t+s,p))\end{aligned} \tag{4.90}$$

which satisfies the intial condition $\sigma(0+s,p)=\sigma(s,p)$. On the other hand we have

$$\frac{d}{dt}\sigma^\mu(t,\sigma(s,p)) = X^\mu(\sigma(t,\sigma(s,p))) \tag{4.91}$$

with the intial condition $\sigma(0,\sigma(s,p)) = \sigma(s,p)$. Thus both $\sigma(t+s,p)$ and $\sigma(t,\sigma(s,p))$ satisfy (4.88) with the same boundary condition and hence they must be equal.

We see that each point $p \in \mathcal{M}$ a flow defines a curve $C_p(t) = \sigma(t,p)$ in $\mathcal{M}$ whose tangent is X and such that $C_p(0) = p$.

Thinking of things the other way we have that, for each t, the flow defines a map $\psi_t : \mathcal{M} \to \mathcal{M}$ such that $\psi_{t+s} = \psi_t \circ \psi_s$ Now for $t=\epsilon$ small we have, from (4.88),

$$\psi^\mu_\epsilon(p) = \sigma^\mu(\epsilon,p) = x^\mu(p) + \epsilon X^\mu(p) + \mathcal{O}(\epsilon^2) \tag{4.92}$$

Thus at least for a small enough value of t $\psi_t(p)$ is 1-1 and $\mathcal{C}^\infty$. Hence it is a diffeomorphism onto its image (at least for an open set in $\mathcal{M}$). This prompts another defintion

Definition: A one parameter group of diffeomorphism of $\mathcal{M}$ is collection of diffeomorhisms $\sigma_t : \mathcal{M} \to \mathcal{M}$ with $t \in \mathbb{R}$ such that

i) $\sigma_t(\sigma_s) = \sigma_{t+S}$

ii) σ_0 is the identity map
iii) $\sigma_{-t} = \sigma_t^{-1}$

So in effect we can think of vector fields as generating infinitessimal diffeomorphism, through their flows. In this sense vector fields can be identified with the Lie algebra of the diffeomorphism group.

4.3 Lie Derivatives

A vector field allows us to introduce a notion of a derivative on a manifold. The problem with the usual derivative

$$\frac{\partial f}{\partial p} = \lim_{\epsilon\to 0} \frac{f(p+\epsilon) - f(p)}{\epsilon} \tag{4.93}$$

is that we don't know how to add two points on a manifold, *i.e.* what is $p+\epsilon$? However we saw that, at least locally, a vector field generates a unique integral flow about any given point p. Therefore we can use the flow to take us to a nearby point and hence form a derivative. This is the notion of a Lie deriviative

Definition: Let X be a vector field on $\mathcal{M}$ and $f \in \mathcal{C}^\infty(\mathcal{M})$ we define the Lie derivative of f along X to be

$$\mathcal{L}_X f(p) = \lim_{\epsilon\to 0} \frac{f(\sigma(\epsilon, p)) - f(p)}{\epsilon} \tag{4.94}$$

where $\sigma(\epsilon, p)$ is the flow generated by X at p.

Theorem: $\mathcal{L}_X f = X(f)$

Proof: We have from the definition that

$$\begin{aligned}
\mathcal{L}_X f(p) &= \lim_{\epsilon\to 0} \frac{f(\sigma(\epsilon, p)) - f(\sigma(0,p))}{\epsilon} \\
&= \frac{d}{dt} f((\sigma(t,p))|_{t=0} \\
&= \sigma(t,p)_\star \left(\frac{d}{dt}\right)_{t=0} (f) \\
&= X_p(f)
\end{aligned} \tag{4.95}$$

where we used the defining property of the flow, namely that its tangent at p is X_p.

We can also define the Lie derivative of a vector field Y along X by

Definition:

$$\mathcal{L}_X Y = \lim_{\epsilon\to 0} \frac{\sigma(-\epsilon)_\star Y_{\sigma(\epsilon)} - Y}{\epsilon} \tag{4.96}$$

where again $\sigma(\epsilon)$ is the flow generated by X and we have suppressed the dependence on the point p.

Theorem:

$$\mathcal{L}_X Y = [X, Y] \tag{4.97}$$

Proof: Let us introduce coordinate about $p \in M$ such that

$$X = \sum_\mu X^\mu(x) \frac{\partial}{\partial x^\mu}\Big|_p , \qquad Y = \sum_\mu Y^\mu(x) \frac{\partial}{\partial x^\mu}\Big|_p \tag{4.98}$$

We start with the fact that

$$\sigma^\nu(\epsilon, p) = x^\nu(p) + \epsilon X^\nu(p) + \mathcal{O}(\epsilon^2) \tag{4.99}$$

and note that, for any f,

$$\begin{aligned} Y_{\sigma(\epsilon)} f &= \sum_\mu Y^\mu(\sigma(\epsilon)) \frac{\partial}{\partial x^\mu}\Big|_{\sigma(\epsilon)} f \\ &= \sum_\mu (Y^\mu + \epsilon X^\lambda \partial_\lambda Y^\mu) \frac{\partial}{\partial x^\mu}\Big|_{\sigma(\epsilon)} f + \mathcal{O}(\epsilon^2) \\ &= \sum_\mu (Y^\mu + \epsilon X^\lambda \partial_\lambda Y^\mu)(\partial_\mu f + \epsilon X^\lambda \partial_\mu \partial_\lambda f) + \mathcal{O}(\epsilon^2) \\ &= \sum_\mu (Y^\mu \partial_\mu + \epsilon X^\lambda \partial_\lambda Y^\mu \partial_\mu + \epsilon Y^\mu X^\lambda \partial_\mu \partial_\lambda) f + \mathcal{O}(\epsilon^2) \end{aligned} \tag{4.100}$$

Therefore

$$\begin{aligned} \sigma(-\epsilon)_\star Y_{\sigma(\epsilon)}(f) &= Y_{\sigma(\epsilon)}(f \circ \sigma(-\epsilon)) \\ &= \sum_\mu Y^\mu(f \circ \sigma(\epsilon)) \frac{\partial}{\partial x^\mu}\Big|_p (f \circ \sigma(-\epsilon)) \\ &= \sum_\mu \Big(Y^\mu \partial_\mu + \epsilon X^\lambda \partial_\lambda Y^\mu \partial_\mu + \epsilon Y^\mu X^\lambda \partial_\mu \partial_\lambda\Big) (f - \epsilon \partial_\nu f X^\nu) + \mathcal{O}(\epsilon^2) \\ &= \sum_\mu Y^\mu \partial_\mu f + \epsilon (\partial_\lambda Y^\mu X^\lambda - Y^\lambda \partial_\lambda X^\mu) \partial_\mu f + \mathcal{O}(\epsilon^2) \end{aligned} \tag{4.101}$$

From which it follows that

$$\begin{aligned} \frac{\sigma(\epsilon)_\star Y_{\sigma(\epsilon)}(f) - Y(f)}{\epsilon} &= (\partial_\lambda Y^\mu X^\lambda - Y^\lambda \partial_\lambda X^\mu) \partial_\mu f + \mathcal{O}(\epsilon) \\ &= [X, Y]^\mu \partial_\mu f + \mathcal{O}(\epsilon) \\ &= [X, Y](f) + \mathcal{O}(\epsilon) \end{aligned} \tag{4.102}$$

and we are done.

5 Tensors

5.1 Co-Tangent Vectors

Recall that the tangent space $T_p\mathcal{M}$ at a point $p \in \mathcal{M}$ is a vector space. For any vector space there is a natural notion of a dual vector space which is defined as the space of a linear maps from the vector space to $\mathbb{R}$.

Problem: Show that the dual space is a vector space and has the same dimension as the original vector space (if it is finite dimensional).

Thus we can define

Definition: The co-tangent space to $\mathcal{M}$ at $p \in \mathcal{M}$ is the dual vector space to $T_p\mathcal{M}$ and is denoted by $T_p^\star\mathcal{M}$.

In other words $\omega_p \in T_p^\star\mathcal{M}$ iff $\omega_p : T_p\mathcal{M} \to \mathbb{R}$ is a linear map. We denote the action of ω_p on a vector $X_p \in T_p\mathcal{M}$ by

$$\omega_p(X_p) = \langle \omega_p, X_p \rangle \tag{5.103}$$

Since ω_p is a linear map we have

$$\langle \omega_p, X_p + \lambda Y_p \rangle = \omega_p(X_p + \lambda Y_p) = \omega_p(X_p) + \lambda\omega_p(Y_p) = \langle \omega_p, X_p \rangle + \lambda\langle \omega_p, Y_p \rangle \tag{5.104}$$

and we may also take, in effect by definition,

$$\langle \omega_p + \lambda\eta_p, X_p \rangle = (\omega_p + \lambda\eta_p)(X_p) = \omega_p(X_p) + \lambda_p\eta(X_p) = \langle \omega_p, X_p \rangle + \lambda\langle \eta_p, X_p \rangle \tag{5.105}$$

Thus $\langle\ ,\ \rangle$ is linear in each of its entries.

Now the dual of the dual of a vector space is just the orginal space itself. Why? Well for a fixed vector X_p we can construct the map:

$$\omega_p \to \omega_p(X_p) \in \mathbb{R} \tag{5.106}$$

The properties of dual vectors ensure that this is a linear map. Thus we can view vectors X_p as linear maps acting on co-vectors ω_p via

$$X_p(\omega_p) = \langle \omega_p, X_p \rangle \tag{5.107}$$

Just as for the tangent bundle we defined the co-tangent bundle to be

$$T^\star\mathcal{M} = \bigcup_{p\in\mathcal{M}} T_p^\star\mathcal{M} \tag{5.108}$$

which is a $2n$-dimensional manifold.

Definition: A smooth co-vector field is a map $\omega : \mathcal{M} \to T^\star\mathcal{M}$ such that

i) $\omega(p) \in T_p^\star\mathcal{M}$
ii) $\omega(X) : \mathcal{M} \to {\rm I\!R}$ is in $\mathcal{C}^\infty(\mathcal{M})$ for all smooth vector fields X

We saw that a chart (U_i, ϕ_i) defines a natural basis of $T_p\mathcal{M}$ for $p \in U_i$:

$$\frac{\partial}{\partial x^\mu}\Big|_p \tag{5.109}$$

where $\phi_i(p) = (x^1(p), ..., x^n(p))$. This allows us to define a natural basis for $T_p^\star\mathcal{M}$ by

$$\langle dx^\mu|_p, \frac{\partial}{\partial x^\nu}\Big|_p\rangle = \delta^\mu_\nu \tag{5.110}$$

i.e. $dx^\mu|_p$ is a linear map from $T_p\mathcal{M}$ to ${\rm I\!R}$ that maps the vector

$$\sum_{\nu=1}^n v^\nu \frac{\partial}{\partial x^\nu}\Big|_p \tag{5.111}$$

to v^μ.

Thus if in a local coordinate chart we have a vector

$$V(p) = \sum_{\mu=1}^n V^\mu(p)\frac{\partial}{\partial x^\nu}\Big|_p \tag{5.112}$$

and a co-vector

$$\omega(p) = \sum_{\nu=1}^n \omega_\nu(p) dx^\nu|_p \tag{5.113}$$

Then

$$\begin{aligned}
\langle\omega(p), V(p)\rangle &= \langle\sum_{\nu=1}\omega_\nu(p)dx^\nu|_p, \sum_{\mu=1} V^\mu(p)\frac{\partial}{\partial x^\mu}\Big|_p\rangle \\
&= \sum_{\mu=1}\sum_{\nu=1}\omega_\nu(p)V^\mu(p)\langle dx^\nu|_p, \frac{\partial}{\partial x^\mu}\Big|_p\rangle \\
&= \sum_{\mu=1}\sum_{\nu=1}\omega_\nu(p)V^\mu(p)\delta^\nu_\mu \\
&= \sum_{\mu=1}\omega_\mu(p)V^\mu(p)
\end{aligned} \tag{5.114}$$

Theorem: Let $(x^1, ..., x^n) = \phi^\mu_1$ and $(y^1, .., y^n) = \phi^\mu_2$ be two coordinate systems at a point $p \in M$ with $U_1 \cap U_2 \neq \emptyset$ and suppose that $\omega_p \in T_p^\star\mathcal{M}$. If

$$\omega_p = \sum_{\mu=1}^n A_\mu dx^\mu\Big|_p \qquad \text{and} \qquad \omega_p = \sum_{\mu=1}^n B_\mu dy^\mu\Big|_p \tag{5.115}$$

then

$$A_\mu = \sum_{\nu=1}^{n} B_\nu \frac{\partial}{\partial x^\mu}\Big|_p y^\nu \tag{5.116}$$

where $y^\mu(x^1, ..., x^n) = \phi_2 \circ \phi_1^{-1}$ is a smooth function from $\phi_1(U_1) \subset \mathbb{R}^n \to \phi_2(U_2) \subset \mathbb{R}^n$.

Proof: We have in the first coordinate system that

$$\omega_p\left(\frac{\partial}{\partial x^\mu}\Big|_p\right) = \sum_{\nu=1}^{n} A_\nu \langle dx^\nu\Big|_p, \frac{\partial}{\partial x^\mu}\Big|_p\rangle = A_\mu \tag{5.117}$$

but in the second coordinate system we see that

$$\omega_p\left(\frac{\partial}{\partial x^\mu}\Big|_p\right) = \sum_{\nu=1}^{n} B_\nu \langle dy^\nu\Big|_p, \frac{\partial}{\partial x^\mu}\Big|_p\rangle \tag{5.118}$$

now we saw in the case of vectors that

$$\frac{\partial}{\partial x^\mu}\Big|_p = \sum_\lambda \left(\frac{\partial}{\partial x^\mu}\Big|_p y^\lambda\right) \frac{\partial}{\partial y^\lambda}\Big|_p \tag{5.119}$$

Thus we find

$$\begin{aligned}
\omega_p\left(\frac{\partial}{\partial x^\mu}\Big|_p\right) &= \sum_{\nu=1}^{n} B_\nu \langle dy^\nu\Big|_p, \sum_\lambda \left(\frac{\partial}{\partial x^\mu}\Big|_p y^\lambda\right) \frac{\partial}{\partial y^\lambda}\Big|_p\rangle \\
&= \sum_{\nu=1}^{n} B_\nu \left(\frac{\partial}{\partial x^\mu}\Big|_p y^\lambda\right) \langle dy^\nu\Big|_p, \sum_\lambda \frac{\partial}{\partial y^\lambda}\Big|_p\rangle \\
&= \sum_{\nu=1}^{n} B_\nu \frac{\partial}{\partial x^\mu}\Big|_p y^\nu
\end{aligned} \tag{5.120}$$

since these must agree we prove the theorem.

N.B.: We can also think in terms of $x^\mu(y) = \phi_1 \circ \phi_2^{-1}$. We can see that

$$\frac{\partial}{\partial x^\mu}\Big|_p y^\nu \qquad \text{and} \qquad \frac{\partial}{\partial y^\mu}\Big|_p x^\nu \tag{5.121}$$

are inverses of each other (when view as matrices):

$$\begin{aligned}
\delta^\nu_\mu &= \frac{\partial x^\nu}{\partial x^\mu} \\
&= \frac{\partial}{\partial x^\mu}\left(\phi_1^\nu \circ \phi_1^{-1}\right)(x) \\
&= \frac{\partial}{\partial x^\mu}\left(\phi_1^\nu \circ \phi_2^{-1} \circ \phi_2 \circ \phi_1^{-1}\right)(x) \\
&= \sum_\lambda \frac{\partial}{\partial y^\lambda}\left(\phi_1^\nu \circ \phi_2^{-1}\right)(y) \frac{\partial}{\partial x^\mu}\left(\phi_2^\lambda \circ \phi_1^{-1}\right)(x) \\
&= \sum_\lambda \left(\frac{\partial}{\partial y^\lambda}\Big|_p x^\nu\right)\left(\frac{\partial}{\partial x^\mu}\Big|_p y^\lambda\right)
\end{aligned} \tag{5.122}$$

In this case the formula becomes

$$B_\mu = \sum_{\nu=1}^{n} A_\nu \frac{\partial}{\partial y^\mu}\Big|_p x^\nu \tag{5.123}$$

N.B.: This formula is often simply written as

$$B_\mu = \sum_{\nu=1}^{n} A_\nu \frac{\partial x^\nu}{\partial y^\mu} \tag{5.124}$$

or even

$$A'_\mu = A_\nu \frac{\partial x^\nu}{\partial x'^\mu} \tag{5.125}$$

with a sum over ν understood and a prime denoting quantities in the new coordinate system. Note the different positions of the prime and unprimed coordinates as compared to the analogous formula for a vector!

5.2 Pull-back and Lie Derivative of a co-vector

Suppose we have smooth map $f : \mathcal{M} \to \mathcal{N}$. We saw that we could push-forward a vector $X_p \in T_p\mathcal{M}$ to a vector $f_\star X_{f(p)} \in T_{f(p)}\mathcal{N}$ by

$$f_\star X_{f(p)}(g) = X_p(g \circ f) \tag{5.126}$$

We therefore can consider the dual map $f^\star : T^\star_{f(p)}\mathcal{N} \to T^\star_p\mathcal{M}$ defined by

$$f^\star_p\omega(X_p) = \langle f^\star\omega, X_p\rangle = \langle \omega, f_\star X_{f(p)}\rangle \tag{5.127}$$

Note that for each co-vector $\omega \in T^\star_{f(p)}\mathcal{N}$ this defines a linear map $f^\star\omega : T_p\mathcal{M} \to \mathbb{R}$ and hence an element of $T^\star_p\mathcal{M}$.

Theorem: Let $f : \mathcal{M} \to \mathcal{N}$, $(y^1, ..., y^n)$ be local coordinates on $V \subset \mathcal{N}$ and $(x^1, ..., x^m)$ local coordinates on $U \cap f^{-1}(V) \subset \mathcal{M}$. If

$$\omega = \sum_{\nu=1}^{n} \omega_\nu dy^\nu\Big|_{f(p)} \tag{5.128}$$

then

$$f^\star\omega = \sum_{\mu=1}^{m}\sum_{\nu=1}^{n} \omega_\nu \frac{\partial}{\partial x^\mu}(y^\nu \circ f)dx^\mu\Big|_p \tag{5.129}$$

Proof: We have that

$$\begin{aligned}
\langle \omega, f_\star X_{f(p)}\rangle &= \sum_{\nu=1}^{n} \omega_\nu (f_\star X_{f(p)})^\nu \\
&= \sum_{\nu=1}^{n}\sum_{\mu=1}^{m} \omega_\nu X^\mu \frac{\partial}{\partial x^\mu}\Big|_p (y^\nu \circ f)
\end{aligned} \tag{5.130}$$

where we used our earlier result for the components of the push forward of a vector. On the other hand we have

$$\langle f^\star\omega, X_p\rangle = \sum_{\mu=1}^{m}(f^\star\omega)_\mu X^\mu \tag{5.131}$$

and since, by definition, $\langle f^\star\omega, X_p\rangle = \langle\omega, f_\star X_{f(p)}\rangle$ we prove the theorem.

We can also extend the definition of the Lie Derivitive to co-vector fields (and in fact all tensor fields).

Definition: If X is a smooth vector field and ω a smooth co-vector field on $\mathcal{M}$ then

$$\mathcal{L}_X\omega = \frac{d}{dt}\left(\sigma(t,p)^\star\omega\Big|_{t=0}\right) \tag{5.132}$$

where $\sigma(t,p)$ is the flow generated by X.

Theorem: If, in a local coordinate system $(x^1, ..., x^n)$,

$$X = \sum_\mu X^\mu(x)\frac{\partial}{\partial x^\mu}\Big|_p \ , \qquad \omega = \sum_\mu \omega_\mu dx^\mu\Big|_p \tag{5.133}$$

then

$$\mathcal{L}_X\omega = \sum_\mu\sum_\nu\left(\partial_\nu\omega_\mu X^\nu + \omega_\nu\partial_\mu X^\nu\right)dx^\mu\Big|_p \tag{5.134}$$

Problem: Prove this theorem.

5.3 Tensors

We can now definition the notition of a (r,s)-tensor. First we need to recall the definition of the tensor product. If $\mathcal{V}$ and $\mathcal{W}$ are two vector spaces, with basis $\{\mathbf{v}_i|i=1,...,n\}$ and $\{\mathbf{w}_a|a=1,...,m\}$ respectively, then the vector space sum is an $n+m$ dimensional vector space which is spanned by $\{\mathbf{v}_i, \mathbf{w}_a|i=1,...,n, a=1,...,n\}$ *i.e.*

$$\mathcal{V}\oplus\mathcal{W} = \mathrm{Span}_{i,a}\{\mathbf{v}_i, \mathbf{w}_a\} \tag{5.135}$$

so a general element is

$$\sum_i^n a^i\mathbf{v}_i + \sum_a^m b^a\mathbf{w}_a \tag{5.136}$$

where the sum is interpreted as a formal sum. Hence $\mathcal{V}\oplus\mathcal{W}$ is an $(n+m)$-dimensional vector space. On the other hand we can also contruct the tensor product which spanned by $\{\mathbf{v}_i\otimes\mathbf{w}_\mathbf{a}|i=1,...,n\ , a=1,...,m\}$, *i.e.*

$$\mathcal{V}\otimes\mathcal{W} = \mathrm{Span}_{i,a}\{\mathbf{v}_i\otimes\mathbf{w}_a\} \tag{5.137}$$

where $\mathbf{v}_i \otimes \mathbf{w}_a$ is a formal product. A general element is

$$\sum_i^m \sum_a^m c^{ia} \mathbf{v}_i \otimes \mathbf{w}_a \tag{5.138}$$

Therefore is $\mathcal{V} \otimes \mathcal{W}$ an nm-dimensional vector space.

Definition: An (r, s)-tensor T at a point $p \in \mathcal{M}$ is an element of

$$T_p^{(r,s)}\mathcal{M} = (\otimes^r T_p\mathcal{M}) \otimes (\otimes^s T_p^\star \mathcal{M}) \tag{5.139}$$

where $\otimes^r$ denotes the rth tensor product.

It follows that, given a local coordinate system, a local basis for (r, s)-tensors is given by

$$\frac{\partial}{\partial x^{\mu_1}}\Big|_p \otimes ... \otimes \frac{\partial}{\partial x^{\mu_r}}\Big|_p \otimes dx^{\nu_1}\Big|_p \otimes ... \otimes dx^{\nu_s}\Big|_p \tag{5.140}$$

and from the definitions earlier we have that if

$$T = \sum T^{\mu_1\mu_2...\mu_r}{}_{\nu_1\nu_2...\nu_s} \frac{\partial}{\partial x^{\mu_1}}\Big|_p \otimes ... \otimes \frac{\partial}{\partial x^{\mu_r}}\Big|_p \otimes dx^{\nu_1}\Big|_p \otimes ... \otimes dx^{\nu_s}\Big|_p \tag{5.141}$$

then the components can be computed as

$$T^{\mu_1\mu_2...\mu_r}{}_{\nu_1\nu_2...\nu_s} = T\left(dx^{\mu_1}\Big|_p, ..., dx^{\mu_r}\Big|_p, \frac{\partial}{\partial x^{\nu_1}}\Big|_p, ... \frac{\partial}{\partial x^{\nu_s}}\Big|_p\right) \tag{5.142}$$

Problem: Show that if, in a local coordinate system

$$T = \sum A^{\mu_1\mu_2...\mu_r}{}_{\nu_1\nu_2...\nu_s} \frac{\partial}{\partial x^{\mu_1}}\Big|_p \otimes ... \otimes \frac{\partial}{\partial x^{\mu_r}}\Big|_p \otimes dx^{\nu_1}\Big|_p \otimes ... \otimes dx^{\nu_s}\Big|_p \tag{5.143}$$

is an (r, s)-tensor,

$$X_i = \sum X_i^\mu \frac{\partial}{\partial x^\mu}\Big|_p, \qquad i = 1, ..., s \tag{5.144}$$

are s vectors and

$$\omega^I = \sum \omega_\nu^I dx^\nu\Big|_p, \qquad I = 1, ..., r \tag{5.145}$$

are r co-vectors, then

$$T(\omega^1, ..., \omega^r, X_1, ..., X_s) = \sum A^{\mu_1...\mu_r}{}_{\nu_1...\nu_s} \omega_{\mu_1}^1 ... \omega_{\mu_r}^r X_1^{\nu_1} ... X_1^{\nu_s} \tag{5.146}$$

Problem: Show that if T is an (r, s)-tensor, $(U_1, \phi_1 = (x^1, ..., x^n))$ and $(U_2, \phi_2 = (y^1, ..., y^n))$ are two local coordinates charts with $U_1 \cap U_2 \neq \emptyset$ such that

$$T = \sum A^{\mu_1\mu_2...\mu_r}{}_{\nu_1\nu_2...\nu_s} \frac{\partial}{\partial x^{\mu_1}}\Big|_p \otimes ... \otimes \frac{\partial}{\partial x^{\mu_r}}\Big|_p \otimes dx^{\nu_1}\Big|_p \otimes ... \otimes dx^{\nu_s}\Big|_p \tag{5.147}$$

and

$$T = \sum B^{\mu_1\mu_2...\mu_s}{}_{\nu_1\nu_2...\nu_s} \frac{\partial}{\partial y^{\mu_1}}\Big|_p \otimes ... \otimes \frac{\partial}{\partial y^{\mu_r}}\Big|_p \otimes dy^{\nu_1}\Big|_p \otimes ... \otimes dy^{\nu_s}\Big|_p \tag{5.148}$$

then

$$\begin{aligned} B^{\mu_1\mu_2...\mu_r}{}_{\nu_1\nu_2...\nu_s} &= \sum_{\lambda_1=1}^{n} ... \sum_{\lambda_r=1}^{n} \sum_{\rho_1=1}^{n} ... \sum_{\rho_s=1}^{n} A^{\rho_1\rho_2...\rho_s}{}_{\lambda_1\lambda_2...\lambda_s} \\ &\times \left(\frac{\partial}{\partial x^{\rho_1}}\Big|_p y^{\mu_1}\right) ... \left(\frac{\partial}{\partial x^{\rho_r}}\Big|_p y^{\mu_r}\right) \\ &\times \left(\frac{\partial}{\partial y^{\nu_1}}\Big|_p x^{\lambda_1}\right) ... \left(\frac{\partial}{\partial y^{\nu_s}}\Big|_p x^{\lambda_s}\right) \end{aligned} \tag{5.149}$$

where $y^\mu(x)$ is understood to be the component of the transition function $\phi_2 \circ \phi_1^{-1}$ and $x^\mu(y)$ is understood to be the components of the transition function $\phi_1 \circ \phi_2^{-1}$.

Definition: An (r, s)-tensor field is a map

$$T : \mathcal{M} \to (\otimes^r T\mathcal{M}) \otimes (\otimes^s T^\star\mathcal{M}) \quad \text{such that} \quad T(p) \in (\otimes^r T_p\mathcal{M}) \otimes (\otimes^s T_p^\star\mathcal{M}) \tag{5.150}$$

which smooth in that, for any choice of r smooth co-vector fields $\omega^1, ..., \omega^r$ and s smooth vector fields $V^1, ..., V^s$, the map $T(\omega^1, ..., \omega^r, V^1, ..., V^s)(p) : \mathcal{M} \to \mathbb{R}$ is $\mathcal{C}^\infty$.

Some tensors and tensor fields have special names:

A $(0, 0)$-tensor is a scalar. As a field it assigns a number to each point in $\mathcal{M}$.

A $(1, 0)$-tensor is a vector. As a field it assigns a tangent vector to each point in $\mathcal{M}$.

A $(0, 1)$-tensor is a 1-form. As a field it assigns a co-vector to each point in $\mathcal{M}$.

Sometimes, especially in older books, $(r, 0)$-tensors are called covariant and $(0, s)$-tensors contravariant.

An $(r, 0)$-tensor is called symmetric if

$$T\left(\omega^{P(1)}, ..., \omega^{P(r)}\right) = T\left(\omega^1, ..., \omega^r\right) \tag{5.151}$$

Similarly and $(0, s)$ tensor is called symmetric if

$$T\left(V^{P(1)}, ..., V^{P(s)}\right) = T\left(V^1, ..., V^s\right) \tag{5.152}$$

On the other hand they are called anti-symmetric if

$$T\left(\omega^{P(1)}, ..., \omega^{P(r)}\right) = \text{sgn}(P) T\left(\omega^1, ..., \omega^r\right) \tag{5.153}$$

or

$$T\left(V^{P(1)},...,V^{P(s)}\right) = \mathrm{sgn}(P)T\left(V^1,...,V^s\right) \tag{5.154}$$

Here P is a permutation and sgn(P) is its sign. Recall that a permutation P is a bijection $P : \{1,..,n\} \to \{1,...,n\}$ and can always be written in terms of either an even (sgn$(P) = 1$) or odd (sgn$(P) = -1$) number of interchanges (where two neighbouring integers are permuted).

Similarly one can talk about the symmetry properties of the $(r,0)$ and $(0,s)$ components of a mixed (r,s)-tensor seperately.

6 Differential Forms

N.B.: Conventions about form manipulations can vary from book to book (by various factors of $p!$ and minus signs). So be careful when comparing two sources.

6.1 Forms

Definition: A p-form on a manifold $\mathcal{M}$ is a smooth anti-symmetric $(0,p)$-tensor field on $\mathcal{M}$. In particular if ω is a p-form then

$$\omega\left(X_{P(1)},...,X_{P(p)}\right) = \mathrm{sgn}(P)\omega\left(X_1,...,X_p\right) \tag{6.155}$$

Definition: A 0-form on $\mathcal{M}$ is a function in $\mathcal{C}^\infty(\mathcal{M})$.

Theorem: If $\mathcal{M}$ is n-dimensional then all p-forms with $p > n$ vanish.

Proof: First note that a p-form acting on a set of vectors with the same vector appearing twice vanishes:

$$\begin{aligned} \omega\left(X_1,...,Y,X_2,...,Y,X_3,...\right) &= (-1)^n\omega\left(Y,Y,X_1,...\right) \\ &= -(-1)^n\omega\left(Y,Y,X_1,...\right) \\ &= 0 \end{aligned} \tag{6.156}$$

where in the first line we used a permutation to place the same vectors next to each other and in the penultimate line we used an interchange (which of course has no effect).

It is now easy to see that if ω is a p-form with $p > n$ then in any collection of p basis vectors at least two must be the same. Hence ω vanishes. Since it vanishes on any set of basis vectors it vanishes identically.

The space of p-forms on $\mathcal{M}$ is denoted $\Omega^p(\mathcal{M},\mathbb{R})$ and we let

$$\Omega(\mathcal{M}) = \Omega^0(\mathcal{M},\mathbb{R}) \oplus \Omega^1(\mathcal{M},\mathbb{R}) \oplus ... \oplus \Omega^n(\mathcal{M},\mathbb{R}) \tag{6.157}$$

Here we have included an explicit reference to the field $\mathbb{R}$ over which the manifold is defined.

Note that if ω and η are a p-form and q-form respectively then $\omega \otimes \eta$ will be a $(0, p+q)$ tensor field but not a $(p+q)$-form since it will not be anti-symmetric. To correct for this we consider the so-called wedge product $\wedge$

Definition: If $\omega \in \omega^p(\mathcal{M})$ and $\eta \in \omega^q(\mathcal{M})$ we define

$$(\omega \wedge \eta)(X_1, ..., X_{p+q}) = \sum_P \text{sgn}(P)(\omega \otimes \eta)(X_{P(1)}, ..., X_{P(p+q)}) \tag{6.158}$$

If should be clear that this defines a $(p+q)$-form by construction. Let us work out a few examples

Example:

$$dx^\mu\Big|_p \wedge dx^\nu\Big|_p = dx^\mu\Big|_p \otimes dx^\nu\Big|_p - dx^\nu\Big|_p \otimes dx^\mu\Big|_p \tag{6.159}$$

Example:

$$\begin{aligned} dx^\mu\Big|_p \wedge \left(dx^\nu\Big|_p \wedge dx^\lambda\Big|_p\right) \;=\;& dx^\mu\Big|_p \otimes dx^\nu\Big|_p \otimes dx^\lambda\Big|_p - dx^\mu\Big|_p \otimes dx^\lambda\Big|_p \otimes dx^\nu\Big|_p \\ & + dx^\nu\Big|_p \otimes dx^\lambda\Big|_p \otimes dx^\mu\Big|_p - dx^\nu\Big|_p \otimes dx^\mu\Big|_p \otimes dx^\lambda\Big|_p \\ & + dx^\lambda\Big|_p \otimes dx^\mu\Big|_p \otimes dx^\nu\Big|_p - dx^\lambda\Big|_p \otimes dx^\nu\Big|_p \otimes dx^\mu\Big|_p \end{aligned} \tag{6.160}$$

Problem: If

$$\begin{aligned} \omega &= \sum A_{12} dx^1 \wedge dx^2 + A_{34} dx^3 \wedge dx^4 \\ \eta &= B_{123} dx^1 \wedge dx^2 \wedge dx^3 + B_{125} dx^1 \wedge dx^2 \wedge dx^5 \end{aligned} \tag{6.161}$$

then what is $\omega \wedge \eta$?

Theorem: If $\omega \in \Omega^p(\mathcal{M})$ and $\eta \in \Omega^q(\mathcal{M})$ then $\omega \wedge \eta = (-1)^{pq} \eta \wedge \omega$

Problem: Prove this!

Theorem: A basis for $\Omega^p(\mathcal{M})$ at a point $p \in \mathcal{M}$ is given by

$$dx^{\mu_1}\Big|_p \wedge ... \wedge dx^{\mu_p}\Big|_p \tag{6.162}$$

Proof: Any $\omega \in \Omega^p(\mathcal{M})$ is defined by its action on a set of basis vectors

$$\omega\left(\frac{\partial}{\partial x^{\mu_1}}\Big|_p, ..., \frac{\partial}{\partial x^{\mu_p}}\Big|_p\right) = \omega_{\mu_1...\mu_p} \tag{6.163}$$

Since ω is antisymmetric we have that

$$\begin{aligned}
\omega_{\mu_{P(1)}\cdots\mu_{P(p)}} &= \omega\left(\frac{\partial}{\partial x^{\mu_{P(1)}}}\Big|_p, ..., \frac{\partial}{\partial x^{\mu_{P(p)}}}\Big|_p\right) \\
&= \text{sgn}(P)\omega\left(\frac{\partial}{\partial x^{\mu_1}}\Big|_p, ..., \frac{\partial}{\partial x^{\mu_p}}\Big|_p\right) \\
&= \text{sgn}(P)\omega_{\mu_1...\mu_p}
\end{aligned} \tag{6.164}$$

so

$$\begin{aligned}
\frac{1}{p!}\sum \omega_{\mu_1...\nu_p} dx^{\mu_1}\Big|_p \wedge ... \wedge dx^{\mu_p}\Big|_p &= \frac{1}{p!}\sum\sum_P \text{sgn}(P)\omega_{\mu_1...\mu_p} dx^{\mu_{P(1)}}\Big|_p \otimes ... \otimes dx^{\mu_{P(p)}}\Big|_p \\
&= \frac{1}{p!}\sum\sum_P \omega_{\mu_{P(1)}\cdots\mu_{P(p)}} dx^{\mu_{P(1)}}\Big|_p \otimes ... \otimes dx^{\mu_{P(p)}}\Big|_p \\
&= \sum \omega_{\mu_1...\nu_p} dx^{\mu_1}\Big|_p \otimes ... \otimes dx^{\mu_p}\Big|_p \\
&= \omega
\end{aligned} \tag{6.165}$$

This shows that $dx^{\mu_1}\Big|_p \wedge ... \wedge dx^{\mu_p}\Big|_p$ spans the space of p-forms. But this formula shows that they are linearly indepdent too (since we know that $dx^{\mu_1}\Big|_p \otimes ... \otimes dx^{\mu_p}\Big|_p$ are linearly independent).

6.2 Exterior Derivative

We can define a notion of a derivative on p-forms by

Definition: If $\omega \in \Omega^p(\mathcal{M})$ then we define

$$\begin{aligned}
d\omega(X_1, ..., X_{p+1}) &= \sum_i (-1)^{i+1} X_i(\omega(X_1, .., X_{i-1}, X_{i+1}, ..., X_{p+1})) \\
&\quad + \sum_{i<j} (-1)^{i+j}\omega([X_i, X_j], X_1, .., X_{i-1}, X_{i+1}, ..., , X_{j-1}, X_{j+1}, ..., X_{p+1})
\end{aligned} \tag{6.166}$$

Why on earth is this called a derivative?

Theorem: If in a local coordinate system we have that

$$\omega = \frac{1}{p!}\sum \omega_{\mu_1,...,\mu_p} dx^{\mu_1}\Big|_p \wedge ... \wedge dx^{\mu_p}\Big|_p \tag{6.167}$$

then

$$d\omega = \frac{1}{p!}\sum \partial_\nu \omega_{\mu_1,...,\mu_p} dx^\nu\Big|_p \wedge dx^{\mu_1}\Big|_p \wedge ... \wedge dx^{\mu_p}\Big|_p \tag{6.168}$$

Proof: We need only evaluate $d\omega$ at the usual choice of basis of vectors

$$d\omega\left(\frac{\partial}{\partial x^{\mu_1}}\Big|_p, ..., \frac{\partial}{\partial x^{\mu_{p+1}}}\Big|_p\right) = \sum_\nu (-1)^{\nu+1}\frac{\partial}{\partial x^\nu}\Big|_p\left(\omega\left(\frac{\partial}{\partial x^1}\Big|_p, ..., \frac{\partial}{\partial x^{\nu-1}}\Big|_p \frac{\partial}{\partial x^{\nu+1}}\Big|_p, ...\frac{\partial}{\partial x^{p+1}}\Big|_p\right)\right) \tag{6.169}$$

where we have used the fact that

$$\left[\frac{\partial}{\partial x^\mu}\Big|_p, \frac{\partial}{\partial x^\nu}\Big|_p\right] = 0 \tag{6.170}$$

so that the second term in the definition of $d\omega$ vanishes. Thus if write

$$d\omega = \frac{1}{(p+1)!}\sum (d\omega)_{\nu\mu_1...\mu_p} dx^\nu\Big|_p \wedge dx^{\mu_1}\Big|_p \wedge ... \wedge dx^{\mu_p}\Big|_p \tag{6.171}$$

we find

$$\begin{aligned} (d\omega)_{\nu\mu_1...\mu_{p+1}} &= \sum_\nu (-1)^{\nu+1}\partial_\nu \omega_{\mu_1...\nu-1,\nu+1...\mu_p} \\ &= (p+1)\partial_{[\nu}\omega_{\mu_2......\mu_{p+1}]} \end{aligned} \tag{6.172}$$

where we have used the antisymmetry of $\omega_{\mu_2......\mu_{p+1}}$ and the square brakets are defined by

$$X_{[\mu_1...\mu_p]} = \frac{1}{p!}\sum_P \mathrm{sgn}(P) X_{\mu_{P(1)}...\mu_{P(p)}} \tag{6.173}$$

However we also need to show that (we implicitly used it above):

Theorem: If $\omega \in \Omega^p(\mathcal{M})$ then $d\omega \in \Omega^{p+1}(\mathcal{M})$

Proof: It is clear that this defines something that is anti-symmetric under an interchange of any two vectors $X_\mu \leftrightarrow X_\nu$. Why does it have this funny form? Let us expand it in terms of some basis. The first term gives

$$\begin{aligned} &X_1(\omega(X_2, X_3, ..., X_p)) - X_2(\omega(X_1, X_3, ..., X_p)) + \dots \\ &= \sum_{\mu\nu...} X_1^\mu \partial_\mu(\omega_{\nu...}X_2^\nu...) - X_2^\mu\partial_\mu(\omega_{\nu...}X_1^\nu...) + \dots \\ &= \sum_{\mu\nu...}(X_1^\mu X_2^\nu... - X_2^\mu X_1^\nu...)\partial_\mu\omega_{\nu...} \\ &+ \sum_{\nu\lambda...}\omega_{\nu\lambda...}(X_1^\mu\partial_\mu X_2^\nu - X_2^\mu\partial_\mu X_1^\nu)X_3^\lambda... + \dots \end{aligned} \tag{6.174}$$

These derivative terms on X_ν need to be cancelled in order for $d\omega$ to be a form. This is the role of the second term since

$$\begin{aligned}
&\sum_{i<j}(-1)^{i+j}\omega([X_i,X_j],X_1,..,X_{i-1},X_{i+1},...,,X_{j-1},X_{j+1},...,X_{p+1})\\
&=-\sum_{\mu\nu...}\omega_{\nu\lambda...}[X_1,X_2]^\nu X_3^\lambda...+\ldots\\
&=-\sum_{\mu\nu...}\omega_{\nu\lambda...}(X_1^\mu\partial_\mu X_2^\nu-X_2^\mu\partial_\mu X_1^\nu)X_3^\lambda...+\ldots
\end{aligned} \tag{6.175}$$

Example: Consider $\mathbb{R}^3$. A 0-form is just a function and

$$df=\sum\partial_\mu f dx^\mu \tag{6.176}$$

is just the gradient of f. A 1-form $\omega=\sum\omega_\mu dx^\mu$ has

$$\begin{aligned}
d\omega &= \sum\partial_\nu\omega_\nu dx^\nu\wedge dx^\mu\\
&= (\partial_1\omega_2-\partial_2\omega_1)dx^1\wedge dx^2+(\partial_2\omega_3-\partial_3\omega_2)dx^2\wedge dx^3+(\partial_3\omega_1-\partial_1\omega_3)dx^3\wedge dx^1
\end{aligned} \tag{6.177}$$

whose components are just those of curl(ω). Lastly for a 2-form $\omega=\frac{1}{2}\sum\omega_{\mu\nu}dx^\mu\wedge dx^\nu$ we have

$$\begin{aligned}
d\omega &= \frac{1}{2}\sum\partial_\lambda\omega_{\mu\nu}dx^\lambda\wedge dx^\mu\wedge dx^\nu\\
&= \frac{1}{2}(\partial_1\omega_{23}+\partial_2\omega_{31}+\partial_3\omega_{12})dx^1\wedge dx^2\wedge dx^3
\end{aligned} \tag{6.178}$$

and these are the components of div($\tilde{\omega}$) where

$$\tilde{\omega}=\frac{1}{2}\omega_{23}dx^1+\frac{1}{2}\omega_{31}dx^2+\frac{1}{2}\omega_{12}dx^3 \tag{6.179}$$

Next we prove the most important property of the exterior derivative:

Theorem: $d^2=0$

Proof: Let us choose a coordinate system as above so that

$$d\omega=\frac{1}{p!}\sum\partial_\nu\omega_{\mu_1,...,\mu_p}dx^\nu\Big|_p\wedge dx^{\mu_1}\Big|_p\wedge...\wedge dx^{\mu_p}\Big|_p \tag{6.180}$$

Then it follows that

$$d^2\omega=\frac{1}{p!}\sum\partial_\lambda\partial_\nu\omega_{\mu_1,...,\mu_p}dx^\lambda\Big|_p\wedge dx^\nu\Big|_p\wedge dx^{\mu_1}\Big|_p\wedge...\wedge dx^{\mu_p}\Big|_p \tag{6.181}$$

but this must vanish as

$$\partial_\lambda \partial_\nu \omega_{\mu_1,\ldots,\mu_p} = \partial_\nu \partial_\lambda \omega_{\mu_1,\ldots,\mu_p} \tag{6.182}$$

Theorem: If $\omega \in \Omega^p(\mathcal{M})$ and $\eta \in \Omega^q(\mathcal{M})$ then $d(\omega \wedge \eta) = (d\omega) \wedge \eta + (-1)^p \omega \wedge d\eta$

Problem: Prove this.

Let us return to the notion of a pull-back. This can be easily extended to any $(0,p)$-tensor (not necessarily a p-form). Consider a $\mathcal{C}^\infty$ map $F : \mathcal{M} \to \mathcal{N}$ between two manifolds and let ω be a $(0,p)$-tensor on $\mathcal{N}$. Then we define

$$f^\star \omega(X_1, \ldots, X_p) = \omega\left(f_\star X_1, \ldots, f_\star X_1\right) \tag{6.183}$$

Just as with a co-vector this defines a $(0,p)$-tensor on $\mathcal{M}$. Clearly $f^\star\omega$ will be anti-symmetric if ω. Thus if $\omega \in \Omega^p(\mathcal{N}, \mathbb{R})$ then $f^\star\omega \in \Omega^p(\mathcal{M}, \mathbb{R})$. We can also define the pull back of a 0-form or function $g : \mathcal{N} \to \mathbb{R}$ by the rule

$$f^\star g = g \circ f \tag{6.184}$$

so that $f^\star g : \mathcal{M} \to \mathbb{R}$.

Theorem: Let $f : \mathcal{M} \to \mathcal{N}$, $(y^1, \ldots, y^n)$ be local coordinates on $V \subset \mathcal{N}$ and $(x^1, \ldots, x^m)$ local coordinates on $U \cap f^{-1}(V) \subset \mathcal{M}$. If

$$\omega = \sum_{\nu=1}^{n} \omega_{\nu_1,\ldots,\nu_p}(f(p)) dy^{\nu_1}\Big|_{f(p)} \wedge \ldots \wedge dy^{\nu_p}\Big|_{f(p)} \tag{6.185}$$

then

$$f^\star \omega = \sum_{\mu=1}^{m} \sum_{\nu=1}^{n} \omega_{\nu_1 \ldots \nu_p}(f(p)) \frac{\partial}{\partial x^{\mu_1}}(y^{\nu_1} \circ f) \ldots \frac{\partial}{\partial x^{\mu_p}}(y^{\nu_p} \circ f) dx^{\mu_1}\Big|_p \wedge \ldots \wedge dx^{\mu_p}\Big|_p \tag{6.186}$$

Problem: Prove this too.

Theorem: The exterior derivative and the pull back commute: $d(f^\star\omega) = f^\star d\omega$.

Proof: In a local coordinate system we saw that that

$$f^\star \omega = \frac{1}{p!} \sum \sum \omega_{\nu_1 \ldots \nu_p}(f(x)) \frac{\partial f^{\nu_1}}{\partial x^{\mu_1}} \ldots \frac{\partial f^{\nu_p}}{\partial x^{\mu_p}} dx^{\mu} \wedge \ldots \wedge dx^{\mu_p} \tag{6.187}$$

where we have used the short hand $f^\nu = y^\nu \circ f$ Therefore we have that

$$df^\star \omega \;\; = \;\; \frac{1}{p!} \sum \partial_\lambda \left(\omega_{\nu_1 \ldots \nu_p}(f(x)) \frac{\partial f^{\nu_1}}{\partial x^{\mu_1}} \ldots \frac{\partial f^{\nu_p}}{\partial x^{\mu_p}} \right) dx^\lambda \wedge dx^{\mu_1} \ldots \wedge dx^{\mu_p}$$

$$
\begin{aligned}
&= \frac{1}{p!}\sum \frac{\partial f^\rho}{\partial x^\lambda}\partial_\rho\omega_{\nu_1...\nu_p}\frac{\partial f^{\nu_1}}{\partial x^{\mu_1}}...\frac{\partial f^{\nu_p}}{\partial x^{\mu_p}}dx^\lambda\wedge dx^{\mu_1}...\wedge dx^{\mu_p} \\
&\quad +\frac{1}{p!}\sum \omega_{\nu_1...\nu_p}\frac{\partial^2 f^{\nu_1}}{\partial x^\lambda\partial x^{\mu_1}}...\frac{\partial f^{\nu_p}}{\partial x^{\mu_p}}dx^\lambda\wedge dx^{\mu_1}...\wedge dx^{\mu_p}+... \\
&= \frac{1}{p!}\sum \partial_\rho\omega_{\nu_1...\nu_p}\frac{\partial f^\rho}{\partial x^\lambda}\frac{\partial f^{\nu_1}}{\partial x^{\mu_1}}...\frac{\partial f^{\nu_p}}{\partial x^{\mu_p}}dx^\lambda\wedge dx^{\mu_1}...\wedge dx^{\mu_p} \\
&= f^\star d\omega
\end{aligned}
\tag{6.188}
$$

Note that in the third line we used the fact that second partial derivatives are symmetric.

6.3 Integration on Manifolds

Let us first recall how we would integrate $\omega = p(x,y)dx + q(x,y)dy$ along a curve $C:[0,1]\to\mathbb{R}^2$ in $\mathbb{R}^2$. A natural presciption is

$$
\int_C p(x,y)dx + q(x,y)dy = \int_0^1 \left(p(C(t))\frac{dC^x}{dt} + q(C(t))\frac{dC^y}{dt}\right)dt \tag{6.189}
$$

Here we are thinking of dx and dy as the infinitessimal change in x and y along the curve C; $(dx,dy) = (dC^x/dtdt, dC^y/dtdt)$. We can rewrite this as

$$
\int_C \omega = \int_0^1 C^\star\omega \tag{6.190}
$$

Problem: Show that

$$
C^\star\omega = \left(p(C(t))\frac{dx\circ C}{dt} + q(C(t))\frac{dy\circ C}{dt}\right)dt \tag{6.191}
$$

This definition clearly extends to the integral of a 1-form along a curve in an arbitrary manifold. To define the integral of a general p-form over a manifold we need to generalise a curve to a p-dimensional surface.

Definition: Let $I_p = [0,1]^p = \{(x^1,...,x^n)\in\mathbb{R}^n|0\le x^\mu\le 1\}$ be a p-cube in $\mathbb{R}^n$.

i) A p-symplex on $\mathcal{M}$ is a $\mathcal{C}^\infty$ map $C: J\to\mathcal{M}$ where J is an open set in $\mathbb{R}^n$ that contains I_p.

ii) A 0-symplex is a map from $\{0\}\to\mathcal{M}$, *i.e.* it is just a point in $\mathcal{M}$.

iii) The support $|C|$ of a p-symplex is the set $C(I_p)\subset\mathcal{M}$

Next we can consider "sums" of such surfaces:

Definition: A p-chain on $\mathcal{M}$ is a finite formal linear combination of p-symplexes on $\mathcal{M}$ with real coefficients, *i.e.* a general p-chain is

$$
\sigma_p = r_1C_1 + ... + r_kC_k \tag{6.192}
$$

where $r_i \in \mathbb{R}$ and C_i are p-symplexs. The support of a p-chain $\sigma_p = r_1 C_1 + ... + r_k C_k$ is $|\sigma_p| = \cup |C_i|$

Definition: We define the maps $\pi_i^{(1)} : I_{p-1} \to I_p$ and $\pi_i^{(0)} : I_{p-1} \to I_p$ by

$$\begin{aligned} \pi_i^{(1)}(t_1, ..., t_{p-1}) &= (t_1, ..., t_{i-1}, 1, t_{i+1}, ..., t_p) \\ \pi_i^{(0)}(t_1, ..., t_{p-1}) &= (t_1, ..., t_{i-1}, 0, t_{i+1}, ..., t_p) \end{aligned} \tag{6.193}$$

i.e. these project a $(p-1)$-cube in $\mathbb{R}^n$ onto a side of a p-cube in $\mathbb{R}^n$. Thus we have a vector space structure on the space of p-chains.

The boundary of a p-cube can then be constructed as the sum of all sides, weighted with a plus sign for front sides and a minus sign for the back sides, in other words a $(p-1)$-chain.

Example: If we look at I_2 then

$$\begin{aligned} \pi_1^{(0)}(t) &= (0, t) \qquad & \pi_2^{(0)}(t) &= (t, 0) \\ \pi_1^{(1)}(t) &= (1, t) \qquad & \pi_2^{(1)}(t) &= (t, 1) \end{aligned} \tag{6.194}$$

As a point set the boundary is

$$\{(1, t) | t \in I_1\} \cup \{(t, 1) | t \in I_1\} \cup \{(0, t) | t \in I_1\} \cup \{(t, 0) | t \in I_1\} \tag{6.195}$$

but this doesn't take into account the fact that some sides are oriented differently to others. This is achieved by considering the 1-chain

$$\pi_1^{(1)} - \pi_1^{(0)} - \pi_2^{(1)} + \pi_2^{(0)} \tag{6.196}$$

whose support is the point set consisting of the boundary.

This allows us to define the boundary of a p-symplex in $\mathcal{M}$

Definition: If C is a p-symplex in $\mathcal{M}$ then the boundary of C is denoted by ∂C and defined to be the $(p-1)$-chain

$$\partial C = \sum_{i=1}^{p} (-1)^{i+1} \left(C \circ \pi_i^{(1)} - C \circ \pi_i^{(0)} \right) \tag{6.197}$$

For a p-chain $\sigma = r_1 C_1 + ... + r_k C_k$ we define

$$\partial \sigma = \sum r_i (\partial C_i) \tag{6.198}$$

The next theorem summarises the notion that boundaries have no boundaries.

Theorem: $\partial^2 = 0$

Proof: Clearly it suffices to show this for a p-symplex C in $\mathcal{M}$.

$$\begin{aligned}
\partial C &= \sum_{i=1}^{p}(-1)^{i+1}\left(C\circ\pi_i^{(1)} - C\circ\pi_i^{(0)}\right)\\
\partial^2 C &= \sum_{i=1}^{p}(-1)^{i+1}\left(\partial(C\circ\pi_i^{(1)}) - \partial(C\circ\pi_i^{(0)})\right)\\
&= \sum_{j=1}^{p-1}\sum_{i=1}^{p}(-1)^{i+j+1}\left(C\circ\pi_i^{(1)}\circ\pi_j^{(1)} - C\circ\pi_i^{(1)}\circ\pi_j^{(0)} - C\circ\pi_i^{(0)}\circ\pi_j^{(1)} + C\circ\pi_i^{(0)}\circ\pi_j^{(0)}\right)
\end{aligned} \tag{6.199}$$

We note that, if $i \le j$

$$\begin{aligned}
\pi_i^{(\alpha)}\circ\pi_j^{(\beta)}(t_1,...,t_p) &= \pi_i^{(\alpha)}(t_1,...,\beta,...,t_p)\\
&= (t_1,...,\alpha,...,\beta,...,t_p)\\
\pi_{j+1}^{(\beta)}\circ\pi_i^{(\alpha)}(t_1,...,t_p) &= \pi_{j+1}^{(\beta)}(t_1,...,\alpha,...,t_p)\\
&= (t_1,...,\alpha,...,\beta...,t_p)
\end{aligned} \tag{6.200}$$

and the α and β appear in the same place. Thus, if $i \le j$,

$$\pi_i^{(\alpha)}\circ\pi_j^{(\beta)} = \pi_{j+1}^{(\beta)}\circ\pi_i^{(\alpha)} \tag{6.201}$$

This shift of $j \to j+1$ introduces a minus sign into the sum due to the $(-1)^{i+j+1}$ factor. Hence we see that the first term and last term in $\partial^2 C$ each sum to zero and the middle two terms will together sum to zero.

Finally we can define the integral of a p-form over a p-chain.

Definition: Let C be a p-symplex in $\mathcal{M}$ and ω a p-form then

$$\int_C \omega = \int_{I_p} C^\star\omega \tag{6.202}$$

where if $C^\star\omega = \frac{1}{p!}f(t_1,..,t_p)dt_1\wedge....\wedge dt_p$ the right hand side is understood to mean the usual integral expression

$$\int_{I_p} C^\star\omega = \int_0^1 ... \int_0^1 f(t_1,..,t_p)dt_1...dt_p \tag{6.203}$$

If $\sigma = \sum r_i C_k$ is a p-chain then

$$\int_\sigma \omega = \sum_i r_i \int_{C_i}\omega \tag{6.204}$$

Example: Consider the manifold $\mathbb{R}^2 - \{(0,0)\}$, the 1-form

$$\omega = \frac{ydx}{x^2+y^2} - \frac{xdy}{x^2+y^2} \tag{6.205}$$

and the curve $C(t) = (\cos(2\pi t), \sin(2\pi t))$. Then

$$\begin{aligned} \int_C \omega &= \int_0^1 C^\star \omega dt \\ &= \int_0^1 \left(\sin(2\pi t)\frac{d\cos(2\pi t)}{dt} - \cos(2\pi t)\frac{d\sin(2\pi t)}{dt} \right) dt \\ &= -2\pi \int_0^1 \left(\sin^2(2\pi t) + \cos^2(2\pi t) \right) dt \\ &= -2\pi \end{aligned} \tag{6.206}$$

Problem: Consider the manifold $\mathbb{R}^2 - \{(0,0)\}$ and the 1-form

$$\omega = \frac{ydx}{x^2+y^2} - \frac{xdy}{x^2+y^2} \tag{6.207}$$

What is

$$\int_C \omega \tag{6.208}$$

along the curve $C(t) = (2+\cos(2\pi t), 2+\sin(2\pi t))$. Next consider the 2-form

$$\omega = \frac{dx \wedge dy}{x^2+y^2} \tag{6.209}$$

What is

$$\int_C \omega \tag{6.210}$$

where $C : I_2 \to \mathbb{R}^2 - \{(0,0)\}$ is given by $C(t_1,t_2) = (t_1+1)(\cos(2\pi t_2), \sin(2\pi t_2))$.

Finally we arrive at a central theorem in differential geometry.

Theorem: (Stokes) If $\omega \in \Omega^{p-1}(\mathcal{M}, \mathbb{R})$ and σ is a p-chain then

$$\int_\sigma d\omega = \int_{\partial\sigma} \omega \tag{6.211}$$

Proof: By linearity we need only show that this is true for p-symplexes C. Recalling that

$$\int_C d\omega = \int_{I_p} C^\star d\omega = \int_{I_p} dC^\star \omega \tag{6.212}$$

it is sufficient to show that

$$\int_{I_p} d\psi = \int_{\partial I_p} \psi \tag{6.213}$$

for any $(p-1)$-form ψ on $\mathbb{R}^n$. Since this condition is linear in ψ it is sufficient to consider

$$\psi = f(x)dx^1 \wedge ... \wedge dx^{p-1} \tag{6.214}$$

so that

$$d\psi = \sum \partial_\lambda f dx^\lambda \wedge dx^1 \wedge ... \wedge dx^{p-1} \tag{6.215}$$

and again by linearity we can consider

$$d\psi = (-1)^{p-1}\partial_p f dx^1 \wedge ... \wedge dx^p \tag{6.216}$$

And now we just compute

$$\begin{aligned}
\int_{I_p} d\psi &= (-1)^{p-1}\int_{I_p}\partial_p f dx^1...dx^p \\
&= (-1)^{p-1}\int dx^1...dx^{p-1}\int_0^1 \partial_p f dx^p \\
&= (-1)^{p-1}\int dx^1...dx^{p-1}\Big(f(x^1,...,x^{p-1},1) - f(x^1,...,x^{p-1},0)\Big)
\end{aligned} \tag{6.217}$$

On the other hand the boundary of I_p is

$$\begin{aligned}
\partial I_p &= \sum_{i=1}^{p}(-1)^{i+1}\left(\pi_i^{(1)} - \pi_i^{(0)}\right) \\
&= \sum_{i=1}^{p}(-1)^{i+1}\Big(\{(x^1,...,1,...,x^p)|x^j \in I_1\} - \{(x^1,...,0,...,x^p)|x^j \in I_1\}\Big)
\end{aligned} \tag{6.218}$$

Now $\psi = f dx^1 \wedge ... \wedge dx^{p-1}$ will only have a non-vanishing contribution to the total integral on those boundary components with x^p constant and $x^1, ..., x^{p-1}$ varying. Hence

$$\begin{aligned}
\int_{\partial I_p}\psi &= (-1)^{p+1}\int_{\{(x^1,...,x^{p-1},1)|x^i\in I_1\}}\psi - (-1)^{p+1}\int_{\{(x^1,...,x^{p-1},0)|x^i\in I_1\}}\psi \\
&= (-1)^{p+1}\int dx^1...dx^{p-1}f(x^1,...,x^{p-1},1) - (-1)^{p+1}\int dx^1...dx^{p-1}f(x^1,...,x^{p-1},0)
\end{aligned} \tag{6.219}$$

and thus we prove the theorem.

This is beautiful generalisation of the following well known result for 1-forms on $\mathbb{R}$

$$\int_a^b df = f(b) - f(a) \tag{6.220}$$

6.4 de Rahm Cohomology and Homology

The exterior deriviative has one very important property:

$$d^2 = 0 \tag{6.221}$$

Thus of $\omega = d\eta$ then it follows that $d\omega = 0$. This motivates two definitions:

Definition:A p-form ω is closed if $d\omega = 0$. We denote the set of closed p-forms on $\mathcal{M}$ by $Z^p(\mathcal{M}, \mathbb{R})$.

Definition:A p-form ω is exact if $\omega = d\eta$ for some $(p-1)$-form on $\mathcal{M}$. We denote the set of exact p-forms on $\mathcal{M}$ by $B^p(\mathcal{M}, \mathbb{R})$.

Theorem: $B^p(\mathcal{M}, \mathbb{R}) \subset Z^p(\mathcal{M}, \mathbb{R})$

Proof: This is obvious since if $\omega = d\eta \in B^p(\mathcal{M}, \mathbb{R})$ then $d\omega = d^2\eta = 0$ so that $\omega \in Z^p(\mathcal{M}, \mathbb{R})$

However the important point is that the converse is not true, not all closed forms are exact. One way to think about this is that d can be viewed as acting from $\Omega^p(\mathcal{M}, \mathbb{R}) \to \Omega^{p+1}(\mathcal{M}, \mathbb{R})$ in sucession:

$$0 \to \Omega^0(\mathcal{M}, \mathbb{R}) \to \Omega^1(\mathcal{M}, \mathbb{R}) \to \ldots \to \Omega^{n-1}(\mathcal{M}, \mathbb{R}) \to \Omega^n(\mathcal{M}, \mathbb{R}) \to 0 \tag{6.222}$$

where we take $d0 = 0$ to be the zero-function in $\Omega^0(\mathcal{M}, \mathbb{R})$. The fact that $d^2 = 0$ means that

$$\text{Image}(d : \Omega^p(\mathcal{M}, \mathbb{R}) \to \Omega^{p+1}(\mathcal{M}, \mathbb{R})) \subset \text{Kernal}(d : \Omega^{p+1}(\mathcal{M}, \mathbb{R}) \to \Omega^{p+2}(\mathcal{M}, \mathbb{R})) \tag{6.223}$$

Such a group of maps is called a differential complex.

Since the space of p-forms is a vector space we can define the following

Definition: The pth de Rahm cohomology group $H^p(\mathcal{M}, \mathbb{R})$ is the quotient space

$$H^p(\mathcal{M}, \mathbb{R}) = \frac{Z^p(\mathcal{M}, \mathbb{R})}{B^p(\mathcal{M}, \mathbb{R})} \tag{6.224}$$

where two p-forms are viewed to be equivalent iff their difference is an exact form. The dimension of $H^p(\mathcal{M}, \mathbb{R})$ is called the pth Betti number b_p.

Example: Consider $H^0(\mathcal{M}, \mathbb{R})$. It is clear that a closed 0-form is just a constant function. Although if $\mathcal{M}$ is disconnected a closed 0-form can take on a different constant value for each connected component of $\mathcal{M}$. Since there are no -1-forms we simply define $Z^0(\mathcal{M}, \mathbb{R})$ to be empty. Thus

$$H^0(\mathcal{M}, \mathbb{R}) \equiv \mathbb{R}^n \tag{6.225}$$

where n is the number of connected components of $\mathcal{M}$.

Problem: Consider the manifold $\mathbb{R}^2 - \{(0,0)\}$ and the 1-form

$$\omega = \frac{ydx}{x^2+y^2} - \frac{xdy}{x^2+y^2} \tag{6.226}$$

Show that this is closed. Is it exact?

Theorem: If $\mathcal{M}$ and $\mathcal{N}$ are two manifolds and $f : \mathcal{M} \to \mathcal{N}$ is a diffeomorphsim then $H^p(\mathcal{M}, \mathbb{R}) \cong H^p(N, \mathbb{R})$.

Proof: Recall that we proved that the pull back and exterior derivative commute. Therefore if ω is a closed p-form on $\mathcal{N}$ then $f^\star\omega$ is a closed form on $\mathcal{M}$;

$$df^\star\omega = f^\star(d\omega) = 0 \tag{6.227}$$

Furthermore if $\omega = d\eta$ is a exact p-form on $\mathcal{N}$ then

$$f^\star\omega = f^\star(d\eta) = d(f^\star\eta) \tag{6.228}$$

is an exact p-form on $\mathcal{M}$. Similarly closed forms on $\mathcal{M}$ are pulled back using f^{-1} to closed forms on $\mathcal{N}$ and exact forms on $\mathcal{M}$ are pulled back to exact forms on $\mathcal{N}$.

Thus the de Rahm cohomology groups are capable of distingishing between two distinct manifolds. By distinct here we mean that two manifolds are equivalent if there is a diffeomorphism between them. Note though that the converse is not true. There are plenty of examples of inequivalent manifolds that have the same de Rham cohomology groups $H^k(\mathcal{M}, \mathbb{R})$. The general idea of cohomology can be applied to any operator which is nilpotent *i.e.* whose action squares to zero, and is a central element of modern algebraic and geometric topology.

So why is it called cohomology? Well we also saw that the boundary operator acting on p-chains satisfies $\partial^2 = 0$. By definition p-chains formed a real vector space. We can therefore construct analogues of closed and exact p-form.

Definition: A p-cycle is a p-chain C such that $\partial C = 0$. We denote the space of p-cycles by $Z_p(\mathcal{M}, \mathbb{R})$.

Definition: If a p-cycle C can be written as $C = \partial D$ where D is a $(p-1)$-chain then it is called a p-boundary. We denote the space of p-boundaries by $B_p(\mathcal{M}, \mathbb{R})$.

In a sense these are more fundamental notions since they don't require a differential structure, *i.e.* we can define them for any topological space $\mathcal{M}$ not necessarily a manifold. They give rise to the so-called homology groups

Definition: The pth homology group $H_p(\mathcal{M}, \mathbb{R})$ is the quotient space

$$H_p(\mathcal{M}, \mathbb{R}) = \frac{Z_p(\mathcal{M}, \mathbb{R})}{B_p(\mathcal{M}, \mathbb{R})} \tag{6.229}$$

where two p-cycles are equivalent iff their difference is a p-boundary.

Thus the homology groups measure which cycles are not the boundaries of chains. This is perhaps a more geometrical object to think about than the de Rahm cohomology groups which are related to differential properties of the manifold. However there is a deep theorem which we won't have time to prove:

Theorem: (de Rham) If $\mathcal{M}$ is compact then $H_p(\mathcal{M}, \mathbb{R})$ and $H^p(\mathcal{M}, \mathbb{R})$ are finite dimensional vector spaces and each others dual. Hence

$$H_p(\mathcal{M}, \mathbb{R}) \cong H^p(\mathcal{M}, \mathbb{R}) \tag{6.230}$$

To see that they are duals of each other one considers the map $\Lambda : H_p(\mathcal{M}, \mathbb{R}) \times H^p(\mathcal{M}, \mathbb{R}) \to \mathbb{R}$

$$\Lambda(\sigma, \omega) = \int_\sigma \omega \tag{6.231}$$

Thus for each p-cycle σ one constructs a linear map from $H^p(\mathcal{M}, \mathbb{R}) \to \mathbb{R}$. Similarly for each p-form ω one constructs a linear map from $H_p(\mathcal{M}, \mathbb{R}) \to \mathbb{R}$. Note in particular that Λ is independent of the choice of representative of the (co)homology group

$$\begin{aligned} \Lambda(\sigma + \partial\rho, \omega + d\eta) &= \int_\sigma \omega + \int_{\partial\rho} \omega + \int_\sigma d\eta + \int_{\partial\rho} d\eta \\ &= \int_\sigma \omega + \int_\rho d\omega + \int_{\partial\sigma} \eta + \int_\rho d^2\eta \\ &= \int_\sigma \omega \end{aligned} \tag{6.232}$$

The important point about exact forms is that they can be written as $\omega = d\eta$ with η well defined everywhere. Stokes theorem can give us a check as to whether or not a form can be exact or not. If $\omega = d\eta$ is exact then

$$\int_\sigma \omega = \int_\sigma d\eta = \int_{\partial\sigma} \eta = 0 \tag{6.233}$$

for any cycle σ since $\partial\sigma = 0$.

So what can fail if we have a closed form and we want make it exact? In a local coordinate system the condition $\omega = d\eta$ is just a set of first order differential equations. A key point is that locally, that is in a sufficently small open set, any closed form is exact. This is summed up by the following theorem, which we don't have time to prove.

Theorem: (Poincare lemma) In a coordinate chart U of $\mathcal{M}$ is contractable to a point then any closed p-form on U is exact. Here constractable to a point means that there exists a smooth map $F : U \times I_1 \to U$ such that $F(p, 0) = p$ and $F(p, 1) = p_0$ for all $p \in U$ where p_0 is some fixed point in U.

What fails then is that the solution to the condition $\omega = d\eta$ can fail to exist globally. This is why the de Rahm Cohomology groups contain topological data.

Example: Consider S^1 with coordinate patches defined by

$$\begin{aligned} U_1 &= S^1 - \{(1,0)\}\ , \qquad \phi_1(\cos\theta, \sin\theta) = \theta \in (0, 2\pi) \\ U_2 &= S^1 - \{(-1,0)\}\ , \qquad \phi_2(\cos\theta, \sin\theta) = \theta \in (-\pi, \pi) \end{aligned} \tag{6.234}$$

It should be easy to check now that this defines a differential structure for S^1. We can consider the curve

$$C(t) = (\cos 2\pi t, \sin 2\pi t) \tag{6.235}$$

This defines a vector field on S^1, namely the tangent to C,

$$T(f) = \frac{d}{dt} f \circ C(t) \tag{6.236}$$

at each point $(\cos 2\pi t, \sin 2\pi t) \in S^1$. In either coordinate system this vector field can be written as

$$\begin{aligned} T(f) &= \frac{d}{dt}\Big(f \circ \phi_i^{-1} \circ \phi_i \circ C(t)\Big) \\ &= \frac{d\phi_i \circ C}{dt}\frac{d}{d\theta}\Big(f \circ \phi_i^{-1}\Big) \\ &= 2\pi \frac{d}{d\theta} f(\theta) \end{aligned} \tag{6.237}$$

i.e.

$$T^\theta = 2\pi \tag{6.238}$$

This in turn defines a dual 1-form ω which is locally

$$\omega = (2\pi)^{-1} d\theta \tag{6.239}$$

i.e.

$$\omega_\theta = (2\pi)^{-1} \tag{6.240}$$

By construction ω is closed, for example $d\omega$ would be a two form on a one-dimensional manifold. But is it exact? Now the curve C is 1-cycle, *i.e.* it has no boundary since

$$\partial C = C \circ \pi^{(1)} - C \circ \pi^{(0)} = C(1) - C(0) = 0 \tag{6.241}$$

We can therefore evaluate

$$\begin{aligned} \int_C \omega &= \int_0^1 C^\star \omega \\ &= \int_0^1 \frac{d\theta \circ C}{dt} \omega_\theta \\ &= (2\pi)^{-1} \int_0^1 2\pi dt \\ &= 1 \end{aligned} \tag{6.242}$$

Since this is non-zero ω cannot be exact.

Thus we see that $H^1(S^1, \mathbb{R})$ is non trivial. In fact $H^1(S^1, \mathbb{R}) \equiv \mathbb{R}$.

We should compare this with another connected 1-dimensional manifold $\mathcal{M} = \mathbb{R}$. If $\omega = \omega_x dx$ is a closed form then this just means that ω_x is a $\mathcal{C}^\infty$ function. However we can then simply define

$$\eta(x) = \int_0^x \omega_x(y) dy \tag{6.243}$$

so that $d\eta = \omega_x dx$. Hence ω is exact and we see that

$$H^1(\mathbb{R}^1, \mathbb{R}) = \mathbf{1} \tag{6.244}$$

is trivial. In this very elementary example we see that that $H^1(\mathcal{M}, \mathbb{R})$ is counting the number of noncontractable 'holes' on the manifold. This is the same as the number of 1-cycles which are not boundaries of 2-cycles. In particular we see that $H^1(\mathcal{M}, \mathbb{R})$ can destinguish between two inequivalent manifolds.

7 Connections, Curvature and Metrics

So far everything that we have discussed about manifolds has been intrinsic to the manifold, defined as a topological space with a differentiable structure, and has not required introducing any additional structures. However it is very common and intuitive to introduce some additional structures.

7.1 Connections, Curvature and Torsion

The first additional structure that we can introduce is that of a connection. We have been emphasising that we can't just differentiate a generic object such as a tensor on a manifold because we don't know how to construct something like $x + \epsilon$ where $x \in \mathcal{M}$ and ϵ is a small parameter.

Problem: Show that if

$$X = \sum_\mu X^\mu \frac{\partial}{\partial x^\mu}\Big|_p \tag{7.245}$$

is a vector field expanded in some local coordinate chart $\phi_i(p) = (x^1(p), ..., x^n(p))$ and we define

$$dX = \sum_\mu \partial_\nu X^\mu dx^\nu\Big|_p \otimes \frac{\partial}{\partial x^\mu}\Big|_p \tag{7.246}$$

then dX will not be a tensor. However check that if $\omega = \sum \omega_\mu dx^\mu|_p$ is a co-vector (1-form) then

$$d\omega = \sum_\mu (\partial_\nu \omega_\mu - \partial_\mu \omega_\nu) dx^\nu|_p \otimes dx^\mu|_p \tag{7.247}$$

is a tensor. Hint: consider how things look in a different coordinate chart $\psi(p) = (y^1, ..., y^n)$.

However we can simply declare that there is a suitable derivative.

Definition: A connection (or more acurately an affine connection) on a manifold $\mathcal{M}$ is an operator D which assigns to each vector field X on $\mathcal{M}$ a mapping $D_X : T\mathcal{M} \to T\mathcal{M}$ such that, for all $Y, Z \in T\mathcal{M}$ and $f \in \mathcal{C}(\mathcal{M})$
i) $D_X(Y+Z) = D_X(Y) + D_X(Z)$
ii) $D_{X+Y}Z = D_X Z + D_Y Z$
iii) $D_{fX}Y = fD_X Y$
iv) $D_X(fY) = X(f)Y + fD_X(Y)$
$D_X Y$ is called the covariant derivative of Y along X.

N.B.: The commutator (or as we saw Lie derivative) obeys all but condition (iii).

In other words D_X acts as a directional derivative along the direction determined by X. However a the existance of connection does not follow from the definition of a manifold but requires us to add it in. In particular a typical manifold can be endowed with infinitely many different connections.

It is conventient to introduce a new notation. If $x^1, ..., x^\mu$ are local coordinates in some chart of $\mathcal{M}$ we let

$$D_\mu = D_{\frac{\partial}{\partial x^\mu}|_p} \tag{7.248}$$

It is easy to convince yourself that D_X is entirely determined by its action on a set of basis vectors hence we introduce

$$D_\mu \frac{\partial}{\partial x^\nu}\Big|_p = \sum_\lambda \Gamma^\lambda_{\mu\nu} \frac{\partial}{\partial x^\lambda}\Big|_p \tag{7.249}$$

and the $\Gamma^\lambda_{\mu\nu}$ are known as the connection coefficients. Thus if

$$X = \sum_\mu X^\mu \frac{\partial}{\partial x^\mu}\Big|_p \qquad \text{and} \qquad Y = \sum_\mu Y^\mu \frac{\partial}{\partial x^\mu}\Big|_p \tag{7.250}$$

then it follows from the definition of D_X that

$$\begin{aligned} D_X Y &= D_{\sum X^\mu \frac{\partial}{\partial x^\mu}|_p} \left(\sum Y^\nu \frac{\partial}{\partial x^\nu}\Big|_p \right) \\ &= \sum X^\mu D_\mu \left(\sum Y^\nu \frac{\partial}{\partial x^\nu}\Big|_p \right) \\ &= \sum X^\mu \left(\frac{\partial}{\partial x^\mu}\Big|_p (Y^\nu) \frac{\partial}{\partial x^\nu}\Big|_p + Y^\nu D_\mu \frac{\partial}{\partial x^\nu}\Big|_p \right) \\ &= \sum \left(X^\mu \partial_\mu Y^\lambda + \Gamma^\lambda_{\mu\nu} X^\mu Y^\nu \right) \frac{\partial}{\partial x^\lambda}\Big|_p \end{aligned} \tag{7.251}$$

Theorem: Let $\Gamma^\lambda_{\mu\nu}$ be the components of a tensor in a coordinate system $(x^1,...,x^n)$ then in another overlapping coordinate system $(y^1,...,y^n)$ we have

$$\tilde{\Gamma}^\lambda_{\mu\nu} = \sum \frac{\partial x^\rho}{\partial y^\mu}\frac{\partial x^\tau}{\partial y^\nu}\frac{\partial y^\lambda}{\partial x^\sigma}\Gamma^\sigma_{\rho\tau} + \sum \frac{\partial y^\mu}{\partial x^\sigma}\frac{\partial^2 x^\sigma}{\partial y^\mu \partial y^\nu} \tag{7.252}$$

where once again we think of $x^\mu(y^\nu)$ as the transition function.

Proof: We have seen that the natural basis of vector field in one coordinate system as compared to the other is

$$\frac{\partial}{\partial y^\mu}\Big|_p = \sum M_\mu{}^\nu \frac{\partial}{\partial x^\nu}\Big|_p , \qquad M_\mu{}^\nu = \frac{\partial x^\nu(y)}{\partial y^\mu} \tag{7.253}$$

By definition we have that

$$D_{\frac{\partial}{\partial y^\mu}|_p}\left(\frac{\partial}{\partial y^\nu}\Big|_p\right) = \tilde{\Gamma}^\lambda_{\mu\nu}\frac{\partial}{\partial y^\lambda}\Big|_p \tag{7.254}$$

so we compute

$$\begin{aligned} D_{\frac{\partial}{\partial y^\mu}|_p}\left(\frac{\partial}{\partial y^\nu}\Big|_p\right) &= D_{\sum M_\mu{}^\rho \frac{\partial}{\partial x^\rho}|_p}\left(\sum M_\nu{}^\lambda \frac{\partial}{\partial x^\lambda}\Big|_p\right) \\ &= \sum M_\mu{}^\rho D_\rho \left(\sum M_\nu{}^\lambda \frac{\partial}{\partial x^\lambda}\Big|_p\right) \\ &= \sum \left(M_\mu{}^\rho \partial_\rho M_\nu{}^\lambda \frac{\partial}{\partial x^\lambda}\Big|_p + M_\mu{}^\rho M_\nu{}^\lambda \Gamma^\sigma_{\rho\lambda}\frac{\partial}{\partial x^\sigma}\Big|_p\right) \end{aligned} \tag{7.255}$$

Next we need to compute this expression on

$$dy^\tau\Big|_p = \sum (M^{-1})^\tau{}_\pi dx^\pi\Big|_p \tag{7.256}$$

Noting that

$$\frac{\partial}{\partial x^\sigma}\Big|_p\left(dy^\tau\Big|_p\right) = \sum (M^{-1})^\tau{}_\sigma \tag{7.257}$$

we find

$$\begin{aligned} D_{\frac{\partial}{\partial y^\mu}|_p}\left(\frac{\partial}{\partial y^\nu}\Big|_p\right)\left(dy^\tau\Big|_p\right) &= \sum \left(M_\mu{}^\rho \partial_\rho M_\nu{}^\lambda (M^{-1})^\tau{}_\lambda + M_\mu{}^\rho M_\nu{}^\lambda \Gamma^\sigma_{\rho\lambda}(M^{-1})^\tau{}_\sigma\right) \\ &= \tilde{\Gamma}^\tau_{\mu\nu} \end{aligned} \tag{7.258}$$

To see that this agrees with the theorem we just note that

$$\sum M_\mu{}^\rho \partial_\rho M_\nu{}^\lambda = \sum \frac{\partial x^\rho}{\partial y^\mu}\partial_\rho M_\nu{}^\lambda = \frac{\partial}{\partial y^\mu} M_\nu{}^\lambda \tag{7.259}$$

Thus the connection coeficients cannot be thought of as the components of a tensor. However from the connection we can construct two associated tensors.

The first is the so-called torsion tensor. This is a $(1,2)$-tensor defined by

$$T(X,Y,\omega) = (D_X Y - D_Y X - [X,Y])(\omega) \tag{7.260}$$

Check that this makes sense: $D_X Y$, $D_Y X$ and $[X,Y]$ are all vector fields and hence they can be evaluated on a co-vector.

Theorem: In a local coordinate system the torsion tensor is

$$\begin{aligned} T &= \sum T^{\lambda}_{\mu\nu}\, dx^{\mu}\Big|_p \otimes dx^{\nu}\Big|_p \otimes \frac{\partial}{\partial x^{\lambda}}\Big|_p \\ &= \sum \left(\Gamma^{\lambda}_{\mu\nu} - \Gamma^{\lambda}_{\nu\mu}\right) dx^{\mu}\Big|_p \otimes dx^{\nu}\Big|_p \otimes \frac{\partial}{\partial x^{\lambda}}\Big|_p \end{aligned} \tag{7.261}$$

Proof: We can simply evaluate T on the basis elements

$$\begin{aligned} T\left(\frac{\partial}{\partial x^{\mu}}, \frac{\partial}{\partial x^{\nu}}, dx^{\lambda}\right) &= \left(D_{\mu}\left(\frac{\partial}{\partial x^{\nu}}\right) - D_{\nu}\left(\frac{\partial}{\partial x^{\mu}}\right)\right)(dx^{\lambda}) \\ &= \sum_{\lambda} \left(\Gamma^{\rho}_{\mu\nu} - \Gamma^{\rho}_{\nu\mu}\right) \frac{\partial}{\partial x^{\rho}}(dx^{\lambda}) \\ &= \sum_{\lambda} \left(\Gamma^{\lambda}_{\mu\nu} - \Gamma^{\lambda}_{\nu\mu}\right) \end{aligned} \tag{7.262}$$

Secondly we have the curvature $(1,3)$-tensor

$$R(X,Y,Z,\omega) = \Big(-D_X(D_Y Z) + D_Y(D_X Z) + D_{[X,Y]}Z\Big)(\omega) \tag{7.263}$$

again this makes sense as all the terms are vector fields acting on a co-vector ω.

Theorem: In a local coordinate system the curvature tensor is

$$\begin{aligned} R &= \sum R_{\mu\nu\lambda}{}^{\rho} dx^{\mu}\Big|_p \otimes dx^{\nu}\Big|_p \otimes dx^{\lambda}\Big|_p \otimes \frac{\partial}{\partial x^{\rho}}\Big|_p \\ &= \sum \left(-\partial_{\mu}\Gamma^{\rho}_{\nu\lambda} + \partial_{\nu}\Gamma^{\rho}_{\mu\lambda} + \Gamma^{\sigma}_{\mu\lambda}\Gamma^{\rho}_{\nu\sigma} + \Gamma^{\sigma}_{\nu\lambda}\Gamma^{\rho}_{\mu\sigma}\right) dx^{\mu}\Big|_p \otimes dx^{\nu}\Big|_p \otimes dx^{\lambda}\Big|_p \otimes \frac{\partial}{\partial x^{\rho}}\Big|_p \end{aligned} \tag{7.264}$$

Problem: Show this.

The important things about these tensors is that they contain coordinate independent information. In particular if a tensor, such as the torsion or curvature, vanishes in one coordinate system then it vanishes in all. This cannot be said of things like the connection coefficients or other quantities that you might encounter while working in a particular coordinate system.

Definition: A vector field X is said to be parallel transported along a curve C if

$$D_{T_C}X = 0 \tag{7.265}$$

at each point on C, where T_C is the tangent to C. This means that we think of X as transported is such a way that it points in the same direction along the curve. This is possible because we have a connection which tells us how to compare vectors in the tangent spaces above different points.

We also can give a geometrical meaning of the torsion and curvature tensors. Consider an infinitessimal displacement of the coordinate x^μ by a vector X

$$\delta_X x^\nu = \epsilon X^\nu \tag{7.266}$$

We can then parallel transport this displacement along a direction given by Y. Parallel transport means that

$$0 = Y^\mu(\partial_\mu X^\nu + \sum \Gamma^\nu_{\mu\lambda}X^\lambda) \tag{7.267}$$

so that

$$\delta_Y X^\nu = -\epsilon \sum Y^\mu \Gamma^\nu_{\mu\lambda}X^\lambda \tag{7.268}$$

Thus

$$\delta_Y \delta_X x^\nu = -\epsilon^2 \sum Y^\mu \Gamma^\nu_{\mu\lambda}X^\lambda \tag{7.269}$$

On the other hand we could first consider the displacement along Y

$$\delta_Y x^\nu = \epsilon Y^\nu \tag{7.270}$$

and then parallel transport this along X

$$\delta_X Y^\nu = -\epsilon^2 X^\mu \sum \Gamma^\nu_{\mu\lambda}Y^\lambda \tag{7.271}$$

This yeilds

$$\delta_X \delta_Y x^\nu = -\epsilon^2 \sum X^\mu \Gamma^\nu_{\mu\lambda}Y^\lambda \tag{7.272}$$

Therefore the difference between these two is measured by the torsion

$$[\delta_X, \delta_Y]x^\nu = -\epsilon^2 \sum X^\mu Y^\lambda T^\nu_{\mu\lambda} \tag{7.273}$$

To understand the curvature we first parallel transport a vector Z around a curve with tangent X by an infinitessimal amount. You can use the formula above to get

$$Z^\rho \to Z^\rho - \epsilon \sum X^\mu \Gamma^\rho_{\mu\lambda}Z^\lambda \tag{7.274}$$

Let us now parallel transport this around X

$$\begin{aligned}
Z^\rho &\rightarrow Z^\rho - \epsilon \sum X^\mu \Gamma^\rho_{\mu\lambda} Z^\lambda - \epsilon \sum Y^\pi \Gamma^\rho_{\pi\sigma}(x+\epsilon X)(Z^\sigma - \epsilon X^\mu \Gamma^\sigma_{\mu\lambda} Z^\lambda) \\
&= Z^\rho - \epsilon \sum X^\mu \Gamma^\rho_{\mu\lambda} Z^\lambda - \epsilon \sum Y^\pi (\Gamma^\rho_{\pi\sigma} + \epsilon X^\tau \partial_\tau \Gamma^\rho_{\pi\sigma})(Z^\sigma - \epsilon X^\mu \Gamma^\sigma_{\mu\lambda} Z^\lambda) \\
&= Z^\rho - \epsilon \sum X^\mu \Gamma^\rho_{\mu\lambda} Z^\lambda - \epsilon \sum Y^\mu \Gamma^\rho_{\mu\sigma} Z^\sigma \\
&\qquad -\epsilon^2 \sum Y^\mu X^\tau \partial_\tau \Gamma^\rho_{\mu\sigma} Z^\sigma + \epsilon^2 \sum \Gamma^\sigma_{\mu\lambda} \Gamma^\rho_{\pi\lambda} Y^\pi X^\mu Z^\lambda + \ldots
\end{aligned} \tag{7.275}$$

If we first transport about Y and then X we find

$$\begin{aligned}
Z^\rho \rightarrow\ & Z^\rho - \epsilon \sum Y^\mu \Gamma^\rho_{\mu\lambda} Z^\lambda - \epsilon \sum X^\mu \Gamma^\rho_{\mu\sigma} Z^\sigma \\
&-\epsilon^2 \sum X^\mu Y^\tau \partial_\tau \Gamma^\rho_{\mu\sigma} Z^\sigma + \epsilon^2 \sum \Gamma^\sigma_{\mu\lambda} \Gamma^\rho_{\pi\lambda} X^\pi Y^\mu Z^\lambda + \ldots
\end{aligned} \tag{7.276}$$

Thus, with a little work, one sees that the parallel transport along Y then X minus parallel transport first along Y and then X is precisely the curvature

$$[\delta_X, \delta_Y] Z^\rho = \epsilon^2 \sum R_{\mu\nu\lambda}{}^\rho X^\mu Y^\nu Z^\lambda \tag{7.277}$$

The connection can be extended to define a covariant dervative on any tensor field. We start by defining it on a co-vector ω by

$$D_X \omega(Y) = X(\omega(Y)) - \omega(D_X Y) \tag{7.278}$$

for any vector field Y. In coordinates this is

$$D_\mu \omega_\nu = \partial_\mu \omega_\nu - \sum \omega_\lambda \Gamma^\lambda_{\mu\nu} \tag{7.279}$$

where we have taken $X = \frac{\partial}{\partial x^\mu}|_p$ and $Y = \frac{\partial}{\partial x^\nu}|_p$ so that

$$\omega(Y) = \omega_\nu \qquad \text{and} \qquad (D_X Y)^\lambda = \Gamma^\lambda_{\mu\nu} \tag{7.280}$$

This implies that

$$D_\mu(dx^\nu) = -\sum \Gamma^\nu_{\mu\lambda} dx^\lambda \tag{7.281}$$

The extension to an (r,s)-tensor is

$$\begin{aligned}
D_X T(\omega^1, ..., \omega^r, Y_1, ..., Y_s) &= X(T(\omega^1, ..., \omega^r, Y_1, ..., Y_s)) \\
&\quad - \sum_i T(\omega^1, ..., D_X\omega^i, ..., \omega^r, Y_1, ..., Y_s) \\
&\quad - \sum_i T(\omega^1, ..., \omega^r, Y_1, ..., D_X Y_i, ..., Y_s)
\end{aligned} \tag{7.282}$$

Problem: Convince yourself that in a local coordinate system

$$\begin{aligned}
D_\lambda T^{\mu_1 \ldots \mu_r}{}_{\nu_1 \ldots \nu_r} &= \partial_\mu T^{\mu_1 \ldots \mu_r}{}_{\nu_1 \ldots \nu_r} \\
&\quad + \sum \Gamma^{\mu_i}_{\lambda\rho} T^{\mu_1 \ldots \rho \ldots \mu_r}{}_{\nu_1 \ldots \nu_r} \\
&\quad - \sum \Gamma^\rho_{\lambda\nu_i} T^{\mu_1 \ldots \mu_r}{}_{\nu_1 \ldots \rho \ldots \nu_r}
\end{aligned} \tag{7.283}$$

7.2 Riemannian Manifolds

Another object that is frequently discussed is a metric. This allows one to measure distances and angles on the manifolds. Again this is not implicit to a manifold and typically there are infinitely many possible metrics for a given manifold. For example General Relativity is a theory of gravity which postulates that spacetime is a manifold. The dynamical equations of General Relativity (Einstein's equation) then determine the metric.

Definition: A metric g on a manifold $\mathcal{M}$ is a non-degenerate inner product on $T_p\mathcal{M}$ for each point $p \in \mathcal{M}$. As such this defines a symmetric $(0,2)$ tensor, that is a map $g_p : T_p\mathcal{M} \otimes T_p\mathcal{M} \to \mathbb{R}$, since inner products are linear and symmetric in each of its entries, *i.e.*

$$g(X,Y) = g(Y,X) \ , \qquad g(X, Y + fZ) = g(X) + fg(X,Z) \tag{7.284}$$

We will of course assume that g is a smooth $(0,2)$ tensor field on $\mathcal{M}$

Definition: A Riemannian manifold is a manifold with a positive definite metric tensor field (positive definite means that $g(X,X) \geq 0$ with equality iff $X = 0$). If the metric is not positive definite it is called a pseudo Riemannian manifold.

As you know from elementary linear algebra an inner product allows us to define the lenths and angles of vectors. Thus with a metric we can define the length and angles of tangent vectors. For example we can now now define the angle between to curves as they intersect to be

$$\arccos\left(\frac{g(T_1,T_2)}{\sqrt{g(T_1,T_1)g(T_2,T_2)}}\right) \tag{7.285}$$

where T_i is the tagent vector to the ith curve at the point where they intersect. We can also define the length of a curve to be

$$\int_C \sqrt{g(T,T)}d\tau \tag{7.286}$$

where T is the tangent to the curve C. So we simply integrate the length of the tangent vector at each point along the curve.

Thus we can give a metric structure to the manifold by defining

$$d(p,q) = \inf_C \int_C \sqrt{g(T,T)}d\tau \tag{7.287}$$

where C is a curve on $\mathcal{M}$ such that $C(0) = p$ and $C(1) = q$.

Example: The Euclidean metric on $\mathbb{R}^n$ is simply

$$g\left(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right) = \delta_{\mu\nu} \tag{7.288}$$

This is in Cartesian coordinates. The length of a curve is therefore just

$$\int_C \sqrt{\sum_\mu \frac{dC^\mu}{d\tau}\frac{dC^\mu}{d\tau}}d\tau \tag{7.289}$$

so in particular for a straight line $C^\mu = p^\mu + (q^\mu - p^\mu)\tau$ one has

$$\int_0^1 \sqrt{\sum_\mu (p^\mu - q^\mu)(p^\mu - q^\mu)}d\tau = \sqrt{\sum_\mu (p^\mu - q^\mu)(p^\mu - q^\mu)} \tag{7.290}$$

which is the Pythagorian distance. But in general the metric coefficients $g_{\mu\nu}$ can be arbitrary functions of the coordinates. Indeed even $\mathbb{R}^n$ with a different coordinate system will have non-trival $g_{\mu\nu}$.

Problem: Consider $\mathbb{R}^2$ and cover it (or more precisely all but the origin) with polar coordinates show that the Euclidean metric is

$$g = dr \otimes dr + r^2 d\theta \otimes d\theta \tag{7.291}$$

Example: We can think of $\mathcal{M} = \mathbb{R}^2 - \{0, 0\}$. This is clearly a manifold as it is an open subset of $\mathbb{R}^2$. By putting different metrics on it though we can think of it in a variety of ways.

With the flat metric

$$g = dr \otimes dr + r^2 d\theta \otimes d\theta \tag{7.292}$$

then this is just want we naturally think of as $\mathcal{M} = \mathbb{R}^2 - \{0, 0\}$ as a subset of the plane.

But we could also consider

$$g' = dr \otimes dr + d\theta \otimes d\theta \tag{7.293}$$

This turns the manifold into a cylinder $S^1 \times \mathbb{R}$, although since $r > 0$ it is really only half a cylinder.

There are also other more exotic possibilities such as

$$g'' = dr \otimes dr + \cosh^2 r d\theta \otimes d\theta \tag{7.294}$$

This looks like a funnel where radius of the circle starts at one and then and grows exponentially with r.

However all of these are diffeomorphic to each other as manifolds. As such the de Rahm cohomology groups will be the same. They just have a different metric put on them.

As is well known an inner product induces an isomorphism between a vector space and its dual. Therefore it follows that a metric tensor induces an isomorphism between

$T_p\mathcal{M}$ and $T_p^\star\mathcal{M}$ at each point $p \in \mathcal{M}$. To be precise, given a vector field X, we can construct a co-vector ω_X by

$$\omega_X(Y) = g(X, Y) \qquad (7.295)$$

clearly this defines a linear map, *i.e.* $\omega_X \in T_p^\star\mathcal{M}$. If in a local coordinate system we have

$$X = \sum X^\mu \frac{\partial}{\partial x^\mu} \qquad \text{and} \qquad g = \sum g_{\mu\nu} dx^\mu \otimes dx^\nu \qquad (7.296)$$

then

$$\begin{aligned} \omega_X(Y) &= \sum (\omega_X)_\nu Y^\nu \\ &= \sum g_{\mu\nu} X^\mu Y^\nu \end{aligned} \qquad (7.297)$$

therefore we see that

$$(\omega_X)_\nu = \sum g_{\mu\nu} X^\mu \qquad (7.298)$$

To see that all co-vectors arise in this way suppose that ω is a co-vector, *i.e.* a linear map from $T_p\mathcal{M}$ to $\mathbb{R}$, then it is defined by its action on a basis

$$\omega\left(\frac{\partial}{\partial x^\mu}\right) = \omega_\mu \qquad (7.299)$$

where

$$\omega = \sum \omega_\mu dx^\mu \qquad \text{and} \qquad g = \sum g_{\mu\nu} dx^\mu \otimes dx^\nu \qquad (7.300)$$

We can therefore consider the vector

$$X_\omega = \sum X_\omega^\nu \frac{\partial}{\partial x^\nu} = \sum g^{\mu\nu} \omega_\nu \frac{\partial}{\partial x^\mu} \qquad (7.301)$$

Here $g^{\mu\nu}$ is the matrix inverse to $g_{\mu\nu}$ which exists since g is non-degenerate. It now follows that

$$\begin{aligned} g(X_\omega, \frac{\partial}{\partial x^\mu}) &= \sum g_{\mu\nu} X_\omega^\nu \\ &= \sum g_{\mu\nu} g^{\nu\lambda} \omega_\lambda \\ &= \omega_\mu \end{aligned} \qquad (7.302)$$

which agrees with the action of ω.

A metric tensor gives rise to an inverse metric $(2, 0)$ tensor by

$$g^{-1}(\omega_X, \omega_Y) = g(X, Y) \qquad (7.303)$$

where we have used the fact that each covector can be identified with a unique vector. Since this identification is linear we see that

$$g^{-1} : T^\star\mathcal{M} \otimes T^\star\mathcal{M} \to \mathbb{R} \qquad (7.304)$$

is linear an symmetric in each entry and hence is a symmetric $(2,0)$-tensor. We have already used this tensor. Using the above form for ω_X and ω_Y shows that, in a particular coordinate system, the left hand side of (7.303) is

$$\begin{aligned} g^{-1}(\omega_X, \omega_Y) &= \sum (g^{-1})^{\mu\nu} (\omega_X)_\mu (\omega_X)_\nu \\ &= \sum (g^{-1})^{\mu\nu} g_{\mu\lambda} X^\lambda g_{\nu\rho} Y^\rho \end{aligned} \tag{7.305}$$

whereas the right hand side of (7.303) is

$$g(X,Y) = \sum g_{\lambda\rho} X^\lambda Y^\rho \tag{7.306}$$

since these must agree we see that

$$(g^{-1})^{\mu\nu} = g^{\mu\nu} \tag{7.307}$$

i.e. the inverse metric. In other words a metric tensor allows us to raise and lower the indices on tensors.

Once a metric is supplied there is a natural choice of connection, known as the Levi-Civita connection.

Theorem: On a (pseudo) Riemannian manifold there is a unique connection D such that
i) $D_X g = 0$ for any vector X
ii) The torsion of D vanishes.

Proof: Let us start by assuming that such a connection exists. From the definition of a covariant derivative on a $(0,2)$-tensor we have

$$0 = D_X g(Y,Z) = X(g(Y,Z)) - g(D_X Y, Z) - g(Y, D_X Z) \tag{7.308}$$

for three vector fields X, Y and Z This implies, along with its cyclic permutations,

$$\begin{aligned} X(g(Y,Z)) &= g(D_X Y, Z) + g(Y, D_X Z) \\ Y(g(Z,X)) &= g(D_Y Z, X) + g(Z, D_Y X) \\ Z(g(X,Y)) &= g(D_Z X, Y) + g(X, D_Z Y) \end{aligned} \tag{7.309}$$

Next we assume that D has no torsion, so that $D_X Y - D_Y X = [X,Y]$. These conditions become

$$\begin{aligned} X(g(Y,Z)) &= g(D_Y X, Z) + g(D_X Z, Y) + g([X,Y], Z) \\ Y(g(Z,X)) &= g(D_Z Y, X) + g(D_Y X, Z) + g([Y,Z], X) \\ Z(g(X,Z)) &= g(D_X Z, Y) + g(D_Z Y, X) + g([Z,X], Y) \end{aligned} \tag{7.310}$$

Thus if we consider the second plus the third minus the first of these equations we find

$$\begin{aligned} Z(g(X,Y)) + Y(g(Z,X)) - X(g(Y,Z)) &= 2g(D_Z Y, X) \\ &\quad + g([Y,Z],X) + g([Z,X],Y) - g([X,Y],Z) \end{aligned} \tag{7.311}$$

Rearranging gives

$$\begin{aligned} 2g(D_Z Y, X) &= Z(g(X,Y)) + Y(g(Z,X)) - X(g(Y,Z)) \\ &\quad - g([Y,Z],X) - g([Z,X],Y) + g([X,Y],Z) \end{aligned} \tag{7.312}$$

Because g is non degenerate and X is arbitrary this will uniquely determine $D_Y Z$. Thus we need only show that this does indeed define a connection.

Properties (i) and (ii) of the definition of the connection follow trivially. To test property (iii) we substitute $Z \to fZ$ in the above equation

$$\begin{aligned} 2g(D_{fZ} Y, X) &= fZ(g(X,Y)) + Y(g(fZ,X)) - X(g(Y,fZ)) \\ &\quad - g([Y,fZ],X) - g([fZ,X],Y) + g([X,Y],fZ) \\ &= 2fg(D_Z Y, Y)) + Y(f)g(Z,X) - X(f)g(Y,Z) \\ &\quad - Y(f)g(Z,X) + X(f)g(Z,Y) \\ &= 2g(fD_Z Y, X) \end{aligned} \tag{7.313}$$

To test property (iv) we substitute $Y \to fY$

$$\begin{aligned} 2g(D_Z fY, X) &= Z(g(X,fY)) + fY(g(Z,X)) - X(g(fY,Z)) \\ &\quad - g([fY,Z],X) - g([Z,X],fY) + g([X,fY],Z) \\ &= 2fg(D_Z Y, X) + Z(f)g(X,Y) - X(f)g(Y,Z) \\ &\quad + Z(f)g(Y,X) + X(f)g(Y,Z) \\ &= 2g(fD_Z Y + Z(f)Y, X) \end{aligned} \tag{7.314}$$

And we prove the theorem.

Therefore given a metric tensor we also find a natural curvature tensor, namely the one corresponding to the Levi-Civita connection. This is called the Riemann curvature

Theorem: In a local coordinate system $(x^1, ..., x^n)$ the coefficients of the Levi-Civita connection are

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}(\partial_\mu g_{\rho\nu} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}) \tag{7.315}$$

Problem: Show this.

References

[1] C. Isham, *Modern Differential Geometry for Physicists*, World Scientific, 1989.

[2] M. Nakahara, *Geometry, Topology and Physics*, IOP, 1990.

[3] I. Madsen and J. Tornehave, *From Calculus to Cohomology*, CUP, 1997.

[4] M. Göckler and T. Schüker, *Differential Geometry, Gauge Thoeries and Gravity*, CUP, 1987.

[5] S. Kobayahi and K. Nomizu, *Foundations of Differential Geometry, vol. I*, Wiley, 1963.