Some Notes on Differential Geometry

Michael Murray

July 20, 2009

Contents

1 Co-ordinate independent calculus.

1.1 Introduction

In this section we review some elementary constructions from calculus. We will formulate them in a way that makes their dependence on co-ordinates manifest. This will make the transition to calculus on manifolds simpler.

1.2 Smooth functions

Recall that if $f\colon U \to \mathbb{R}$ is a function defined on an open subset U of $\mathbb{R}$ then we say that f is differentiable at $x \in U$ if the limit

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$$

exists. If the limit exists we call it the derivative of f at x and denote it by any of

$$df(x), \quad d_x f, \quad \text{or} \quad f'(x).$$

If f is differentiable at any x in U we just say that f is differentiable.

If $U \subset \mathbb{R}^n$ is open and $f\colon U \to \mathbb{R}$ we can define partial derivatives by varying only one of the co-ordinates. If e^i is the element of $\mathbb{R}^n$ with a 1 in the ith position and 0's elsewhere we define a curve by

$$\gamma_i(t) = x + te_i.$$

The ith partial derivative of f at x is then defined by

$$\partial_i f(x) = (f \circ \gamma_i)'(0).$$

We say that f is *smooth* if it has partial derivatives of any order. Because a differentiable function is continuous it follows that f has continuous partial derivative of any order. I am quite deliberately avoiding the notation

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

for the time being.

If $U \subset \mathbb{R}^n$ and $f\colon U \to \mathbb{R}^m$ then we say that f is smooth if the functions $f^i\colon U \to \mathbb{R}$ are smooth where $f(x) = (f^1(x), f^2(x), \dots, f^m(x))$.

1.3 Derivatives as linear operators.

Because partial derivatives are co-ordinate dependent they are not a particularly useful way of thinking about derivatives if we want to move to a co-ordinate independent setting such as differentiable manifolds. It is more useful to think of the derivative of a function $f\colon U \to \mathbb{R}$ at x, for $U \subset \mathbb{R}^n$, as a *linear* map

$$df(x)\colon \mathbb{R}^n \to \mathbb{R}$$

defined by

$$df(x)(v) = \frac{d}{dt}(t \mapsto f(x+tv))\Big|_{t=0}.$$

We think of this as the rate of change of f at x in the direction of v. For smooth functions $df(x)$ is linear. Note that $df(x)$ is akin to the notion of a directional derivative but we do not require that v is of unit length. We can recover the partial derivatives from this definition by applying the linear operator $df(x)$ to the vector e^i. The result, $df(x)(e^i)$, is just the ith partial derivative of f at x.

Similarly if $f\colon U \to \mathbb{R}^m$ then we define a linear map

$$df(x)\colon \mathbb{R}^n \to \mathbb{R}^m$$

by

$$df(x)(v) = \frac{d}{dt}(t \mapsto f(x+tv))\Big|_{t=0}.$$

As a linear map we can expand $df(x)$ in a basis and we recover the Jacobian matrix

$$df(x)(e^i) = \sum_{j=1}^{m} \partial_i f^j e^j.$$

1.4 The chain rule.

Fundamental to many of the constructions we want to consider in the following sections is the chain rule:

Theorem 1.1 (Chain Rule.). *Let $U \subset \mathbb{R}^n$ be open , $f\colon U \to \mathbb{R}^m$, $V \subset \mathbb{R}^m$ open and $g\colon V \to \mathbb{R}^k$ with $f(U) \subset V$. Let $x \in U$. If f and g are smooth so also is $g \circ f\colon U \to \mathbb{R}^k$ and*

$$d(g \circ f)(x) = dg(f(x)) \circ df(x).$$

The composition on the right hand side is the composition of linear operators. In particular if we expand both sides in terms of the standard basis of $\mathbb{R}^n$ then we have

$$\partial_j(g^i \circ f)(x) = \sum_{l=1}^{m} \partial_l g^i \circ \partial_j f^l$$

An important part of the chain rule is the fact that the composition of smooth functions is also smooth. A partial converse of this result will be important in the sequel.

Lemma 1.2. *Let U be an open subset of $\mathbb{R}^n$ and V an open subset of $\mathbb{R}^m$. A function $\phi\colon U \to V$ is smooth if and only if for every smooth function $f\colon V \to \mathbb{R}$ the composite $f \circ \phi\colon U \to \mathbb{R}$ is smooth.*

Proof. If ϕ is smooth then the result follows via the chain rule. If the result is true then take f to be the restriction to V of each of the co-ordinate functions x^i. Then x^i is smooth so $x^i \circ \phi = \phi^i$ is smooth. But if each component of ϕ^i is smooth so also is ϕ. □

1.5 Diffeomorphisms and the inverse function theorem.

A function $f\colon U \to V$ where U and V are open subsets of $\mathbb{R}^n$ is called a (smooth) *diffeomorphism* if it is smooth, invertible and has smooth inverse. If f is a diffeomorphism $f \circ f^{-1} = 1_{\mathbb{R}^n}$ so it follows from the chain rule that at any point $x \in U$

$$1_{\mathbb{R}^n} = d(1_{\mathbb{R}^n})(x) = df^{-1}(f(x)) \circ df(x)$$

so that $(df(x))^{-1} = df^{-1}(f(x))$. That is, the inverse of the linear map $df(x)$ is the linear map $df^{-1}(f(x))$. Notice that this means that a diffeomorphism necessarily goes from an open subset of $\mathbb{R}^n$ to an open subset of $\mathbb{R}^m$ where $n = m$ so we have lost nothing by putting that in the definition.

It is also useful to have the notion of a *local diffeomorphism*. We say that $f\colon U \to \mathbb{R}^n$ is a local diffeomorphism at $x \in U$ if there is an open subset V of $\mathbb{R}^n$ containing x such that $f(V)$ is open and $f\colon V \to f(V)$ is a diffeomorphism.

With this notion we have the important inverse function theorem:

Theorem 1.3 (Inverse Function Theorem). *Let U be an open subset of $\mathbb{R}^n$ and $f\colon U \to \mathbb{R}^n$ be a smooth function such that $df(x)$ is invertible. Then f is a local diffeomorphism at x and $d(f^{-1})(f(x)) = (df(x))^{-1}$.*

The Lemma proved in the previous section also gives us a characterisation of diffeomorphisms:

Lemma 1.4. *Let U and V be open subsets of $\mathbb{R}^n$. A bijection $\phi\colon U \to V$ is a diffeomorphism if and only if for every function $f\colon V \to \mathbb{R}$ we have that f is differentiable if and only if $f \circ \phi\colon U \to \mathbb{R}$ is differentiable.*

Proof. We just apply Lemma 1.2 to ϕ and ϕ^{-1}. □

2 Differentiable manifolds

2.1 Co-ordinate charts

Manifolds are sets on which we can define co-ordinates in such a way that we can do calculus. In general we don't expect to be able to define co-ordinates on all of a manifold any more than we can cover the whole of the surface of the earth with a single map. First we define:

Definition 2.1 (Co-ordinate charts)**.** A co-ordinate chart on a set M is a pair (U, ψ) where $U \subset M$ and $\psi\colon U \to \mathbb{R}^n$ is a function which is a bijection onto its image $\psi(U)$ which is open in $\mathbb{R}^n$.

If (U, ψ) is a co-ordinate chart we call U the domain of the co-ordinate chart and ψ the co-ordinates. Notice that we do not say that U is open in M because M is not a topological space yet; it is just a set.

Example 2.1. Let $1_{\mathbb{R}^n}\colon \mathbb{R}^n \to \mathbb{R}^n$ be the identity map. That is $1_{\mathbb{R}^n}(x) = (x^1, \dots, x^n)$. Then $(\mathbb{R}^n, 1_{R^n})$ is a co-ordinate chart on $\mathbb{R}^n$. We usually call these the *standard*, *usual* or *natural* co-ordinates.

Example 2.2. Let U be any open subset of $\mathbb{R}^n$ and

$$\iota\colon U \to \mathbb{R}^n$$

the inclusion map defined by $\iota(x) = x$. Then clearly $\iota(U) = U$ which is open and ι is a bijection onto $\iota(U) = U$ so that (U, ι) is a co-ordinate chart on $\mathbb{R}^n$.

Example 2.3. Let V be a finite dimensional vector space. Choose a basis $v^1, \dots, v^n$ for V and define $\psi\colon V \to \mathbb{R}^n$ by

$$u = \sum_{i=1}^{n} \psi^i(u) v^i.$$

Then ψ is a bijection, in fact a linear isomorphism. Indeed every linear isomorphism arises in this way as if $\phi\colon V \to \mathbb{R}^n$ is a linear isomorphism we can take $w^i = \phi^{-1}(e^i)$ where e^i is the vector with a 1 in the ith place and zeros everywhere else. We leave it as an exercise to show that for every $u \in V$

$$u = \sum_{i=1}^{n} \phi^i(u) w^i.$$

Example 2.4. Let

$$U = \mathbb{R}^2 - \{(x, 0) \mid x \le 0\}$$

and define polar co-ordinates

$$(r, \theta)\colon U \to (0, \infty) \times (-\pi, \pi) \subset \mathbb{R}^2$$

as follows. We define $r\colon U \to (0, \infty)$ by $r(x, y) = \sqrt{x^2 + y^2}$ and $\theta\colon U \to (-\pi, \pi)$ by the requirement that $x = r(x, y)\cos(\theta(x, y))$ and $y = r(x, y)\sin(\theta(x, y))$ and $\theta(x, 0) = 0$. Clearly (r, θ) is a bijection on the given domain and range.

Example 2.5. Let S^2 be the set of all points in $\mathbb{R}^3$ of length one. Let

$$U_0 = S^2 - \{(0, 0, 1)\} \subset \mathbb{R}^2.$$

We can define co-ordinates on U by *stereographic projection* from the point $(0, 0, 1)$ onto the X-Y plane. That is if $p = (x, y, z) \in U_0$ it has co-ordinates $\psi(p) = (\psi_0^1(p), \psi_0^2(p))$ defined uniquely by the requirement that the line through $(0,0,1)$ and p intersects the X-Y plane at $(\psi_0^1(p), \psi_0^2(p), 0)$. So we must have

$$(x-0, y-0, z-1) = (x^1 - 0, x^2 - 0, 0 - 1)$$

and hence

$$\psi_0^1(x,y,z) = \frac{x}{1-z} \qquad \text{and} \qquad \psi_0^2(x,y,z) = \frac{y}{1-z}.$$

In general a manifold will have lots of co-ordinates. We no more expect a manifold to come with a given set of co-ordinates than we expect an abstract vector space to come with a given basis. However not all co-ordinate charts will do. We want them to be able to fit together in some compatible way. The motivation for our definition comes from the desire to define differentiable (in fact smooth) functions on a manifold. Indeed we can regard co-ordinates as a device to decide which, of the many functions on M, are going to be smooth. Let (U,ψ) be a co-ordinate chart and let $f\colon U \to \mathbb{R}$ be a function. Then as U is just a set it makes no sense to ask that f be smooth. However we can ask that f be smooth with respect to the co-ordinates. That is we consider

$$f \circ \psi^{-1}\colon \psi(U) \to \mathbb{R}.$$

Now $f \circ \phi^{-1}$ is a function defined on an open subset of $\mathbb{R}^n$, namely $\psi(U)$ and we know what it means for such a function to be smooth. Consider now what happens when we change co-ordinates to some other co-ordinate chart say (V,χ) for convenience assuming that $V = U$. Then it is possible that $f \circ \psi^{-1}$ is smooth but $f \circ \chi^{-1}$ is not. To compare them we write

$$f \circ \psi^{-1} = f \circ \chi^{-1} \circ (\chi \circ \psi^{-1})$$

where

$$\chi \circ \psi^{-1}\colon \psi(U) \to \chi(V)$$

is a bijection between open subsets of $\mathbb{R}^n$. Then a sufficient condition for $f \circ \psi^{-1}$ to be smooth if $f \circ \chi^{-1}$ is is that $\chi \circ \psi^{-1}$ is smooth. As we want this to work both ways we also require that $\psi \circ \chi^{-1}$ be smooth. In other words we require that $\chi \circ \psi^{-1}$ is a diffeomorphism. If we want this to be true for any f then we have already seen in Lemma 1.2 that this becomes a necessary condition.

In practice we may not be able to find charts (U,ψ) and (V,χ) with $U = V$ so in the definition we need to allow for this.

Definition 2.2 (Compatibility of charts). A pair of charts (U,ψ) and (V,χ) are called compatible if the sets $\psi(U \cap V)$ and $\chi(U \cap V)$ are open and the map

$$\chi \circ \psi^{-1}_{|\psi(U\cap V)}\colon \psi(U \cap V) \to \chi(U \cap V)$$

is a diffeomorphism.

Note that we need to restrict the map ψ^{-1} to the set $\psi(U \cap V)$ so that it can be composed with χ. In general just writing $\chi \circ \psi^{-1}$ will not make sense.

Example 2.6. If $U \subset \mathbb{R}^2$ is the set in example 2.4 on which polar co-ordinates are defined then it has two co-ordinate charts defined on it $(U,(r,\theta))$, and (U,ι). The polar co-ordinates and the inclusion. Notice that $U \cap U = U$ so that $\iota(U \cap U)$ and $(r,\theta)(U \cap U)$ are open by assumption.

If we calculate the composition

$$\iota \circ (r,\theta)^{-1}\colon (0,\infty) \times (-\pi,\pi) \to U$$

we obtain

$$\iota \circ (r,\theta)^{-1}(s,\phi) = (s\sin(\phi), s\cos(\phi))$$

which is a diffeomorphism. Hence $(U,(r,\theta))$ and (U,ι) are compatible.

Example 2.7. Let V be a vector space and $v^1, \dots, v^n$ and $w^1, \dots, w^n$ bases defining co-ordinates ψ and ϕ by

$$v = \sum_{i=1}^{n} \psi^i(v) v^i = \sum_{i=1}^{n} \phi^i(v) w^i.$$

Notice that both ϕ and ψ are onto so that $\psi(V \cap V) = \mathbb{R}^n$ is certainly open in $\mathbb{R}^n$ and likewise for ϕ. If we define a matrix X^i_j by

$$v^i = \sum_{j=1}^{n} X^i_j w^j$$

for all i then

$$\sum_{i,j=1}^{n} \psi^i(v) X^i_j w^j = \sum_{j=1}^{n} \phi^j(v) w^j$$

so that

$$\phi^j(v) = \sum_{i=1}^{n} X^i_j \psi^i(v).$$

Another way of calculating this result is to observe that

$$\phi \colon V \to \mathbb{R}^n$$

and

$$\psi \colon V \to \mathbb{R}^n$$

are linear isomorphisms so that

$$\phi \circ \psi^{-1} \colon \mathbb{R}^n \to \mathbb{R}^n$$

is the linear isomorphism with matrix X^i_j. Being linear $\phi \circ \psi^{-1}$ is certainly smooth so that (V, ϕ) and (V, ψ) are compatible.

Example 2.8. If we consider again the example of S^2 we had defined a co-ordinate chart (U_0, ψ_0) taking

$$U_0 = S^2 - \{(0, 0, 1)\}$$

and

$$\psi_0(x, y, z) = \left(\frac{x}{1 - z}, \frac{y}{1 - z} \right).$$

If we stereographically project from the point $(0, 0, -1)$ then we get co-ordinates

$$\psi_1(x, y, z) = \left(\frac{x}{1 + z}, \frac{y}{1 + z} \right).$$

defined on

$$U_1 = S^2 - \{(0, 0, -1)\}$$

We want to check that these are compatible. Note first that both $\psi_0(U_0 \cap U_1)$ and $\psi_1(U_0 \cap U_1)$ are equal to $\mathbb{R}^2 - \{(0, 0, 0)\}$ which is open in $\mathbb{R}^2$. Then an easy calculation shows that

$$\psi_0 \circ \psi_1^{-1}{}_{|\mathbb{R}^2 - \{(0,0,0)\}}(x^1, x^2) = \left(\frac{x^1}{(x^1)^2 + (x^2)^2}, \frac{x^2}{(x^1)^2 + (x^2)^2} \right)$$

which is a smooth map on $\mathbb{R}^2 - \{(0, 0, 0)\}$. Similarly for $\psi_1 \circ \psi_0^{-1}{}_{|\mathbb{R}^2 - \{(0,0,0)\}}$.

To make M into a manifold we need to be able to cover it with compatible co-ordinate charts.

Definition 2.3 (Atlas). An atlas for a set M is a collection $\{(U_\alpha, \psi_\alpha) \mid \alpha \in I\}$ of co-ordinate charts such that:

(i) for any α and β in I, (U_α, ψ_α) and (U_β, ψ_β) are compatible and;
(ii) $M = \cup_{\alpha \in I} U_\alpha$.

Then we have

Definition 2.4 (Manifold)**.** A *manifold* is a set M with an atlas $\mathcal{A}$. We call the choice of an atlas $\mathcal{A}$ for a set M a choice of *differentiable structure* for M.

Example 2.9. If there is a co-ordinate chart with domain all of M then this, by itself defines an atlas and makes M a manifold. For example $(\mathbb{R}^n, \mathrm{id})$ makes $\mathbb{R}^n$ a manifold and if U is open in $\mathbb{R}^n$ then (U, ι) makes U a manifold.

Example 2.10. If V is a vector space then any linear isomorphism from V to $\mathbb{R}^n$ makes V a manifold. The vector space V has other atlases such as the atlas of all linear isomorphisms

$$\{(V, \phi) \mid \phi\colon V \to \mathbb{R}^n \text{ a linear isomorphism}\}.$$

Example 2.11. The charts (U_0, ψ_0) and (U_1, ψ_1) are compatible and have domains that cover S^2 so they make it into a manifold. It is not difficult to show that we cannot make S^2 into a manifold with only one chart (S^2, χ) if we require that χ is continuous. Indeed if χ is continuous then because S^2 is compact we must have $\chi(S^2) \subset \mathbb{R}^n$ compact and hence closed but $\chi(S^2)$ is open so this is not possible unless $\chi(S^2) = \mathbb{R}^n$ but then it is not compact.

Example 2.12. Consider the set $\mathbb{R}P_n$ of all lines through the origin in $\mathbb{R}^{n+1}$. We shall show that this is a manifold. This manifold is called *real projective space* of dimension n. If $x = (x^0, \dots, x^n)$ is a *non-zero* vector in $\mathbb{R}^{n+1}$ we denote by $[x] = [x^0, \dots, x^n]$ the line through x. The numbers $x = (x^0, \dots, x^n)$ are often called the *homogeneous co-ordinates* of the line $[x]$. It is important to note that they are not uniquely determined by knowing the line. Indeed we have that $[x] = [y]$ if and only if there is a non-zero real number λ such that $x = \lambda y$. Define subsets $U_i \subset \mathbb{R}P_n$ by

$$U_i = \{[x] \mid x^i \neq 0\}$$

for each $i = 0, \dots, n$ and notice that these subsets cover all of $\mathbb{R}P_n$. Define maps

$$\begin{aligned} \psi_i\colon U_i &\to \mathbb{R}^n \\ [x^0, \dots, x^n] &\mapsto \Big(\frac{x^0}{x^i}, \dots, \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i}, \dots, \frac{x^n}{x^i}\Big). \end{aligned}$$

Notice that we need to check that these maps are well defined but that follows from the fact that $[x] = [y]$ only if x is a scalar multiple of y. It also straightforward to check that the ψ_i are bijections onto $\mathbb{R}^n$ and hence define co-ordinates. Lastly it is straightforward to check that these co-ordinate charts are all compatible and hence make $\mathbb{R}P_n$ into a manifold.

We need to now deal with a technical problem raised by the definition of atlas. We often want to work with co-ordinate charts that are not in the atlas $\mathcal{A}$ used to define the differentiable structure. For example if $M = \mathbb{R}^2$ we might take $\mathcal{A} = \{1_{\mathbb{R}^n}\}$. Then in a particular problem we might want to work with polar co-ordinates. But are they somehow compatible with the differentiable structure already imposed by $\mathcal{A}$? The definition of what compatibility is in this sense is easy. We could say that another co-ordinate chart is compatible with the given atlas if when we add it to the atlas we still have an atlas. In other words it is compatible with all the charts already in the atlas. We will take a different, but equivalent, approach via the notion of a *maximal* atlas containing $\mathcal{A}$ to explain these notions. We define;

Definition 2.5 (Maximal atlas.)**.** An atlas $\bar{\mathcal{A}}$ for a set M is a maximal atlas for an atlas $\mathcal{A}$ if $\mathcal{A} \subset \bar{\mathcal{A}}$ and for any other atlas $\mathcal{B}$ with $\mathcal{A} \subset \mathcal{B}$ we have $\mathcal{B} \subset \bar{\mathcal{A}}$.

We then have

Proposition 2.6. *For any atlas $\mathcal{A}$ on a set M there is a unique maximal atlas $\bar{\mathcal{A}}$ containing $\mathcal{A}$. The maximal atlas consists of every chart compatible with all the charts in $\mathcal{A}$.*

Proof. Define the set $\bar{\mathcal{A}}$ to the set of all charts which are compatible with every chart in $\mathcal{A}$. Then clearly if $\mathcal{B}$ is another atlas for M with $\mathcal{A} \subset \mathcal{B}$ then we must have $\mathcal{B} \subset \bar{\mathcal{A}}$. What is not immediate is that $\bar{\mathcal{A}}$ is an atlas. The problem is that we do not know that the charts we have added to $\mathcal{A}$ to form $\bar{\mathcal{A}}$ are compatible with each other. So let (U, ψ) and (V, χ) be charts in

$\bar{\mathcal{A}}$. We need to show that (U, ψ) is compatible with (V, χ). Recall from the definition that this is true if the sets $\psi(U \cap V)$ and $\chi(U \cap V$ are open and

$$\chi \circ \psi^{-1}_{|\psi(U\cap V)} \colon \psi(U \cap V) \to \chi(U \cap V$$

is a diffeomorphism. To prove this it suffices to show that for every x in $U \cap V$ we can find a W with $x \in W \subset U \cap V$ such that $\psi(W)$ and $\chi(W)$ are open and such that

$$\chi \circ \psi^{-1}_{|\psi(W)} \colon \psi(W) \to \chi(W)$$

is a diffeomorphism.

To find W choose a co-ordinate chart (Z, ϕ) in $\mathcal{A}$ with $x \in Z$. This is possible as the domains of the charts in an atlas cover M. Then let $W = U \cap V \cap Z$. Now (U, ψ) is compatible with (Z, ϕ) so that $\phi(U \cap Z)$ is open. Similarly $\phi(V \cap Z)$ is open so that

$$\phi(W) = \phi(U \cap Z) \cap \phi(V \cap Z)$$

is open. Using compatibility again we that

$$\psi \circ \phi^{-1}_{|\phi(U\cap Z)} \colon \phi(U \cap Z) \to \psi(U \cap Z)$$

is a diffeomorphism and hence a homeomorphism so that

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

is open as required. A similar argument shows that $\chi(W)$ is open. Then the chain rule shows that

$$\chi \circ \psi^{-1}_{|\psi(W)} = (\chi \circ \phi^{-1}_{|\phi(W)}) \circ (\phi \circ \psi^{-1}_{|\psi(W)})$$

is a diffeomorphism. □

Finally we have

Definition 2.7. If M is a manifold with atlas $\mathcal{A}$ we define a co-ordinate chart on the manifold $(M, \mathcal{A})$ to be a co-ordinate chart on the set M which is in the maximal atlas $\bar{\mathcal{A}}$.

It should be noted that having defined an atlas we tend not to refer to it very much. We usually say (U, ψ) is a co-ordinate chart on a manifold M rather than (U, ψ) is a member of the atlas $\mathcal{A}$ for a manifold $(M, \mathcal{A})$. The situation is similar to that for a topological space X with topology $\mathcal{T}$. We rarely refer to the topology $\mathcal{T}$ by name. We say U is an open subset of X rather than U is in the topology defining the open sets of X.

2.2 Linear manifolds.

There are many similarities between manifolds and vector spaces. Choosing co-ordinates is much like choosing a basis. It is useful to develop this idea further.

Definition 2.8. Define linear co-ordinates ψ on a set V to be a bijection $\psi \colon V \to \mathbb{R}^n$.

Definition 2.9. Define two sets of linear co-ordinates ψ and χ to be linearly equivalent if $\psi \circ \chi^{-1}$ is a linear isomorphism.

It is straightforward to prove that linear equivalence is an equivalence relation. We define

Definition 2.10. A linear atlas on a set V is an equivalence class of linear co-ordinates.

Definition 2.11. We define a linear manifold to be a set V with a choice of linear atlas.

We can define an addition and scalar multiplication on V by choosing some linear co-ordinates ψ from the linear atlas and defining

$$av + bw = \psi^{-1}(a\psi(v) + b\psi(w))$$

where a and b are real numbers and v and w are elements of V. We have to check that this is *well-defined* that is it is independent of the choice of ψ from the equivalence class. If χ is another choice then we have

$$\begin{aligned} av + bw &= \psi^{-1}(a\psi(v) + b\psi(w)) \\ &= \psi^{-1}(a\psi(\chi^{-1} \circ \chi(v)) + b\psi(\chi^{-1} \circ \chi(w))) \\ &= \psi^{-1}(a(\psi \circ \chi^{-1})(\chi(v)) + b(\psi \circ \chi^{-1})(\chi(w))) \\ &= \psi^{-1}(\psi \circ \chi^{-1})(a\chi(v) + b\chi(w)) \\ &= \chi^{-1}(a\chi(v) + b\chi(w)) \end{aligned}$$

where in moving from the third to the fourth lines we use the fact that $\psi \circ \chi^{-1}$ is linear. We have proved.

Proposition 2.12. *A linear manifold has a natural vector space structure which makes all of the linear co-ordinates linear isomorphisms.*

Because of Proposition 2.12 the theory of linear manifolds is really the theory of vector spaces. However it is an amusing exercise to translate everything in the theory of vector spaces into the linear manifold setting. For example a function $f\colon V \to \mathbb{R}$ is linear if $f \circ \psi^{-1}$ is linear some choice of linear co-ordinates ψ. It is then easy to prove that $f \circ \chi^{-1}\colon V \to \mathbb{R}$ for any choice of linear co-ordinates χ. Indeed we just note that

$$f \circ \chi^{-1} = (f \circ \psi^{-1}) \circ (\psi \circ \chi^{-1}).$$

2.3 Topology of a manifold

Often a manifold is defined as a topological space and the domains of the charts are required to be open sets and the co-ordinates homeomorphisms. This is really superfluous as the topology is forced once we have chosen the atlas. Given a manifold M we define a subset $W \subset M$ to be open if for every $x \in W$ there is a chart with domain U such that $x \in U \subset W$. We need to show that such a definition of open sets defines a topology on M. The only problem is showing that the intersection of two open sets is open. This follows from the following Lemma whose proof we leave as an exercise.

Lemma 2.13. *Let (U, ψ) be a co-ordinate chart on a manifold M and let $W \subset U$ be such that $\psi(W) \subset \psi(U)$ is open. Then $(W, \psi_{|W})$ is a co-ordinate chart.*

We also leave as an exercise showing that with this topology if (U, ψ) is a co-ordinate chart then $\psi\colon U \to \psi(U)$ is a homeomorphism.

Readers familiar with the notion of a basis for a topology will realise that we are claiming here that the set of all domains of coordinate charts in a maximal atlas is a basis for a topology on M. It is this topology we have described above.

We will in general require a manifold to be Hausdorff and paracompact in the topology defined by its atlas.

Now that we have defined the topology of a manifold we can discuss its dimension. Each co-ordinate function has as range some $\mathbb{R}^d$. From the definition of compatibility it is clear that d is constant on the connected components of M. We shall go further and assume that our manifolds are such that this number d is constant on all of M. We call this number the dimension of the manifold.

2.4 Exercises

Exercise 2.1. Consider the n sphere

$$S^n = \{x \in \mathbb{R}^{n+1} \mid \|x\|^2 = 1\}.$$

Let

$$U_i = \{x \in S^n \mid x^i \neq 1\} = S^n - \{e_i\}$$

where e_i has all components 0 except the ith which is equal to one. If x is a point in U_i show that there is a unique line through x and the vector e_i. Show that this line intersects the plane

$$\{x \mid x^i = 0\}$$

in exactly one point. Writing this point as

$$(\psi_i^1(x), \psi_i^2(x), \dots, \psi_i^{i-1}(x), 0, \psi_i^i(x), \dots, \psi_i^n(x))$$

defines a function

$$\psi_i = (\psi_i^1, \dots, \psi_i^n) \colon U_i \to \mathbb{R}^n.$$

Show that (U_i, ψ_i) is a co-ordinate chart on S^n and that

$$\{(U_i, \psi_i) \mid i = 1, \dots, n\}.$$

is an atlas for S^n.

The functions ψ_i are are said to arise by *stereographic projection* from e_i onto the plane $\{x \mid x^i = 0\}$.

Exercise 2.2. Consider the sphere S^n again. Define

$$U_i^+ = \{x \in S^n \mid x^i > 0\}$$

and define $\psi_i^+ \colon U_i^+ \to \mathbb{R}^n$ by

$$\psi_i^+(x^1, \dots, x^{n+1}) = (x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^n).$$

Show that (U_i^+, ψ_i^+) is a co-ordinate chart for S^n. Similarly define

$$U_i^- = \{x \in S^n \mid x^i < 0\}$$

and define $\psi_i^- \colon U_i^+ \to \mathbb{R}^n$ by

$$\psi_i^-(x^1, \dots, x^{n+1}) = (x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^n).$$

Again show that (U_i^-, ψ_i^-) is a co-ordinate chart for S^n.

Show that

$$\{(U_i^+, \psi_i^+), (U_i^-, \psi_i^-) \mid i = 1, \dots, n\}$$

is an atlas for S^n.

Exercise 2.3. Show that the atlases in Exercises 2.1 and 2.2 define the same maximal atlas on S^n.

Exercise 2.4. Let $\mathbb{R}P_n$ be the set of all lines (through the origin) in $\mathbb{R}^{n+1}$. This space is called real, projective space of dimension n. If x is a non-zero vector in $\mathbb{R}^{n+1}$ denote by $[x]$ the line through x. Show that $[x] = [y]$ if and only if there is a non-zero real number λ such that $x = \lambda y$.

Define subsets U_i of $\mathbb{R}P_n$ by

$$U_i = \{[x^0, \dots, x^n] | x^i \neq 0\}$$

and maps $\varphi_i : U_i \to \mathbb{R}^n$ by

$$\varphi_i([x^0, \dots, x^n]) = (\frac{x^0}{x^i}, \dots, \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i}, \dots, \frac{x^n}{x_i})$$

for every $i = 0, \dots, n$. Show that φ_i is well defined and that (U_i, φ_i) is a co-ordinate chart on $\mathbb{R}P_n$. Show that

$$\{(U_i, \varphi_i) | i = 0, \dots, n\}$$

is an atlas for $\mathbb{R}P_n$.

Exercise 2.5. Show that if M_1 and M_2 are manifolds then there is a natural way of making $M_1 \times M_2$ into a manifold so that $\dim(M_1 \times M_2) = \dim(M_1) + \dim(M_2)$.

Exercise 2.6. Repeat exercise (2.4) for $\mathbb{C}^n$ to define $2n$ dimensional complex projective space $\mathbb{C}P_n$ as the space of complex lines through zero in $\mathbb{C}^{n+1}$.

3 Smooth functions on a manifold.

We motivated the definition of the compatibility of charts by the problem of defining smooth functions on a manifold. Let us return to that problem now.

Definition 3.1. A function $f \colon M \to \mathbb{R}$ on a manifold M is *smooth* if we can cover the manifold with co-ordinate charts (U, ψ) such that $f \circ \psi^{-1} \colon \psi(U) \to \mathbb{R}$ is smooth.

Notice that we do not know that $f \circ \psi^{-1} \colon \psi(U) \to \mathbb{R}$ is smooth for any chart (U, ψ) but only that we can cover M with charts for which this is so. To get this stronger result we need the following Lemma.

Lemma 3.2. *If $f \colon M \to \mathbb{R}$ is a smooth function and (V, χ) is a co-ordinate chart then $f \circ \chi^{-1} \to \mathbb{R}$ is smooth.*

Proof. It suffices to show that for every $x \in V$ there is a $W \subset V$ containing x such that $f \circ \chi^{-1}_{|\chi(W)}$ is smooth. Pick any such x. Then by definition there is a chart (U, ψ) with $x \in U$ and $f \circ \psi^{-1} \colon \psi(U) \to \mathbb{R}$ smooth. Let $W = U \cap V$. Then

$$f \circ \chi^{-1}_{|\chi(W)} = (f \circ \psi^{-1} | \psi(W)) \circ (\psi \circ \chi^{-1} | \chi(W))$$

which is smooth by the chain rule and compatibility of charts. □

We will also be interested in smooth functions of a single real variable into a manifold or paths. We have

Definition 3.3. If x is a point of a manifold and $\gamma \colon (-\epsilon, \epsilon) \to M$ we say that γ is a (smooth) *path* through x if $\gamma(0) = x$ and there is a chart (U, ψ) with $\gamma((-\epsilon, \epsilon)) \subset U$ and such that $\psi \circ \gamma$ is smooth.

Example 3.1. If x is a point in $\mathbb{R}^n$ and v is a vector in $\mathbb{R}^n$ then the function

$$t \mapsto x + tv$$

is a path through x.

Example 3.2. If $x \in S^2$ and $v \in \mathbb{R}^3$ with $\langle x, v \rangle = 0$ then

$$t \mapsto \frac{x + tv}{\|x + tv\|}$$

is a path in S^2 through x.

We have a similar type of lemma as before.

Lemma 3.4. *If $\gamma \colon (-\epsilon, \epsilon) \to M$ is a path in M and (V, χ) is a chart with $\gamma((-\epsilon, \epsilon)) \subset V$ then $\chi \circ \gamma$ is a path in M.*

Proof. Chain rule and compatibility. □

3.1 Exercises

Exercise 3.1. Define $h\colon \mathbb{R} \to \mathbb{R}$ by

$$h(x) = \begin{cases} \exp(\frac{-1}{1-x^2}) & \text{if } -1 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

and show that h is smooth. By integrating h find a smooth function $g\colon \mathbb{R} \to \mathbb{R}$ with the property that $g(x)$ is zero for $x < -1$ and $g(x)$ is one for $x > 1$. Show that for any $\epsilon > \delta > 0$ there is a smooth function $\phi\colon \mathbb{R}^n \to \mathbb{R}$ with $\phi(x)$ equal to zero if $\|x\| > \epsilon$ and ϕ equal to one if $\|x\| < \delta$. Now consider a manifold M and a point x. By using co-ordinates show that if U is any open subset of M containing x then there are open subsets U_1 and U_2 with $x \in U_1 \subset U_2 \subset U$ and a smooth function $f\colon M \to \mathbb{R}$ with f equal to 1 on all of U_1 and equal to zero outside of U_2.

Exercise 3.2. Let x be point in a manifold M. Let X_x be the set of all pairs (U, f) where U is a open set containing x and $f\colon U \to \mathbb{R}$ is a smooth function. Define a relation on X_x by saying that $(U, f) \simeq (V, g)$ if there is an open set W with $x \in W \subset U \cap V$ and $f_{|W} = g_{|V}$. Show that this an equivalence relation. Equivalence classes are called *germs* at x and the set of them we will denote by G_x. Show that G_x is an algebra under pointwise addition, scalar multiplication and multiplication. If $f \in C^\infty(M, \mathbb{R})$ the algebra of all smooth functions on M it defines the germ containing (M, f). Show that the map this induces $C^\infty(M, \mathbb{R}) \to G_x$ is onto. [Hint: Use 3.1.]

Exercise 3.3. Consider the map $F\colon \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$F(x, y, z) = (x^2 + y^2 + z^2 - 9, x + y + z - 3).$$

If we identity the tangent spaces to $\mathbb{R}^3$ and $\mathbb{R}^2$ with $\mathbb{R}^3$ and $\mathbb{R}^2$ respectively calculate the tangent map

$$T_{(x,y,z)}F\colon \mathbb{R}^3 \to \mathbb{R}^2.$$

Exercise 3.4. Define a map $F\colon S^2 \to \mathbb{C}P_1$ by

$$F(x, y, z) = [x + iy, 1 - z].$$

By using the co-ordinates defined in Exercises (2.1) and (2.6) show that this map is well defined as $z \to 1$ and that it is, in fact, a diffeomorphism.

4 The tangent space.

Most of the theory of calculus on manifolds needs the idea of tangent vectors and tangent spaces. The name 'tangent vector' comes of course from examples like $S^2 \subset \mathbb{R}^3$ where a tangent vector at $x \in S^2$ is a vector in $\mathbb{R}^3$ tangent to the sphere which in that particular case means orthogonal to x. However in the case of a general manifold M it does not come to us sitting inside some $\mathbb{R}^N$ and we have to work a little harder to develop a notion of tangent vector.

Although we do not have a notion of tangent vector yet we do have the notion of a smooth path in a manifold. Let us we what this does for us in $\mathbb{R}^n$. In that case if $f\colon \mathbb{R}^n \to \mathbb{R}$ and a path $\gamma\colon \mathbb{R} \to \mathbb{R}^n$ with $\gamma(0) = x$ then we can consider

$$(f \circ \gamma)'(0)$$

the rate of change of f along γ as we go through 0. By the chain rule we can write this as

$$(f \circ \gamma)'(0) = df(x)(\gamma'(0))$$

where $\gamma'(0)$ is the tangent vector to γ at $t = 0$. Notice that this equation tells us the $(f \circ \gamma)'(0)$ depends on γ only through $\gamma'(0)$, that is if we replace γ by another path ρ with $\rho(0) = x$ and $\rho'(0) = \gamma'(0)$ then

$$(f \circ \gamma)'(0) = (f \circ \rho)'(0).$$

On a manifold we do not have the vector space structure of $\mathbb{R}^n$ so we cannot, immediately, differentiate a path. However we can compose a smooth path γ and a smooth function f to obtain a function

$$f \circ \gamma \colon (-\epsilon, \epsilon) \to \mathbb{R}.$$

Moreover if we choose co-ordinates (U, ψ) with $\gamma((-\epsilon, \epsilon)) \subset U$ then we have that

$$f \circ \gamma = (f \circ \psi^{-1}) \circ (\psi \circ \gamma)$$

so that $f \circ \gamma$, being the composition of two smooth functions, is smooth and it makes sense to consider

$$(f \circ \gamma)'(0).$$

If we insert the co-ordinates again and apply the chain rule this is

$$(f \circ \gamma)'(0) = d(f \circ \psi^{-1})(\psi(x))(\psi \circ \gamma)'(0).$$

We would like $(f \circ \gamma)'(0)$ to be the rate of change of f in the direction $\gamma'(0)$ but because we are on a manifold we do not know what $\gamma'(0)$ is. To avoid this problem we just define $\gamma'(0)$ to be the set of all paths which should have the same tangent vector $\gamma'(0)$. We do this as follows.

Definition 4.1 (Tangency). Let $\gamma \colon (-\epsilon, \epsilon) \to M$ and $\rho \colon (-\delta, \delta) \to M$ be paths through a point x. We say that γ and ρ are tangent at $t = 0$ if for every $f \in C^\infty(M, \mathbb{R})$ we have

$$(f \circ \gamma)'(0) = (f \circ \rho)'(0)$$

Let $U \subset M$ is an open set containing m and $g \colon U \to \mathbb{R}$ be smooth. Then if $\gamma \colon (-\epsilon, \epsilon) \to M$ is a path through m we have that $\gamma^{-1}(U)$ is an open neighbourhood of $0 \in (-\epsilon, \epsilon)$ so there exists ϵ' such that $\gamma(-\epsilon', \epsilon') \subset U$ and we can define $(g \circ \gamma_{|(-\epsilon', \epsilon')})'(0)$. Clearly this is independent of the choice of ϵ' and we will just denote it by $g \circ \gamma)'(0)$.

Lemma 4.2. *Let $\gamma \colon (-\epsilon, \epsilon) \to M$ and $\rho \colon (-\delta, \delta) \to M$ be paths though a point x. Then γ and ρ are tangent at 0 if and only if there is a chart (V, χ) containing x such that*

$$\chi \circ \gamma'(0) = \chi \circ \rho'(0).$$

Proof. If γ and ρ are tangent then the result follows by applying the definition to the co-ordinate functions χ multiplied by a smooth bump function which is one in a neighbourhood of x. Conversely if the result is true then by the chain rule

$$(f \circ \gamma)'(0) = d(f \circ \psi^{-1})(\psi(x))(\psi \circ \gamma)'(0) = d(f \circ \psi^{-1})(\psi(x))(\psi \circ \rho)'(0) = (f \circ \rho)'(0).$$

□

It is easy to see that tangency is an equivalence relation on the set of all paths through the point x. The equivalence classes are called tangent vectors (although we have not yet shown that they are vectors). The equivalence class containing a path γ is denoted by $\gamma'(0)$ or $t_0(\gamma)$. If X is a tangent vector and $\gamma \in X$ then we usually say that X is tangent to γ rather than that γ is an element of X. The set of all tangent vectors at x we denote by T_xM. We want to show now that T_xM has the structure of a vector space.

Let γ be a path and (U, ψ) a choice of co-ordinates with U containing the image of γ. Then

$$\psi \circ \gamma \colon (-\epsilon, \epsilon) \to \mathbb{R}^n$$

is a smooth path in $\mathbb{R}^n$. This has a tangent vector at zero which is the vector

$$(\psi \circ \gamma)'(0)$$

in $\mathbb{R}^n$ at $\psi(x)$. Notice that from the lemma this depends only on $\gamma'(0)$. We define a map

$$d\psi(x) \colon T_xM \to \mathbb{R}^n$$

by

$$d\psi(x)(\gamma'(0)) = (\psi \circ \gamma)'(0).$$

By definition of tangency this map is injective we want to prove

Proposition 4.3. *The map $d\psi(x)$ is a bijection.*

Proof. As we have already noted it suffices to show that this map is onto. Let (U, ψ) be a chart about x. If v is a vector in $\mathbb{R}^n$ then $t \mapsto \psi(x) + tv$ is a path in $\mathbb{R}^n$ with tangent vector v. Because $\psi(U)$ is open we can find an $\epsilon > 0$ such that if $|t| < \epsilon$ then $\psi(x) + tv \in \psi(U)$. Then we can define $\gamma\colon (-\epsilon, \epsilon) \to M$ by

$$\gamma(t) = \psi^{-1}(x + tv).$$

Then we have $\psi \circ \gamma(t) = \psi(x) + tv$ so that $(\psi \circ \gamma)'(0) = v$. $\square$

Lemma 4.4. *If χ and ψ are co-ordinates on M and γ is a path through x then*

$$(\chi \circ \gamma)'(0) = d(\chi \circ \psi^{-1})(\psi(x))(\psi \circ \gamma)'(0).$$

or

$$d\chi(x) = d(\chi \circ \psi^{-1})(\psi(x)) \circ d\psi(x)$$

Proof. The lemma follows immediately from the chain rule applied to the composition of maps

$$\chi \circ \gamma = (\chi \circ \psi^{-1}) \circ (\psi \circ \gamma).$$

Notice that all the maps here are defined on open subsets of $\mathbb{R}^n$ so that we can apply the standard chain rule. $\square$

From the discussion in the previous section the maps $d\chi(x)$ define linear co-ordinates on T_xM and hence by Proposition 2.12 T_xM has a unique vector space structure which makes all the maps $d\psi(x)$ linear isomorphisms.

Example 4.1. As always the first example is $M = \mathbb{R}^n$. In that case we have a preferred set of co-ordinates. These are just the identity. So two paths γ and ρ are tangent if and only if $\gamma'(0) = \rho'(0)$. In other words two paths are tangent if they have the same tangent vector at x. Notice also that if v is any vector there is a preferred path whose tangent vector is v. That is the straight line $t \mapsto x + tv$. So in the case of $\mathbb{R}^n$ there is no reason to introduce all the extra machinery of equivalence classes of paths.

Example 4.2. The second example is $M = V$ a finite dimensional vector space. Notice that if γ is a path through v in V then we can make sense of of the derivative of γ at 0 directly by

$$\gamma'(0) = \lim_{t \to 0} \frac{\gamma(t) - \gamma(0)}{t}.$$

Of course $\gamma'(0)$ defined in this way is a vector in V whereas above we have defined $\gamma'(0)$ as an equivalence class of paths. This correspondence defines an isomorphism

$$\begin{aligned} T_v(V) &\to V \\ t_0(\gamma) &\mapsto \gamma'(0) \end{aligned}$$

Notice that the inverse to this map sends a vector w to the tangency class of the straight line $w \mapsto v + tw$ and that each tangency class $t_0(\gamma)$ contains a unique straight line $t \mapsto \gamma(0) + t\gamma'(0)$. Again in this case the extra machinery of equivalence classes of paths adds nothing to what we already know.

In the introduction to this section we remarked on the case of the two-sphere which is a submanifold of $\mathbb{R}^3$. We will return to submanifolds shortly but first we need to consider the notion of the derivative of a function.

4.1 The derivative of a function.

Recall that a function $f\colon M \to \mathbb{R}$ is smooth if we can cover M with co-ordinates (U, ψ) such that $f\colon \psi^{-1}\colon \psi(M) \to \mathbb{R}$ is smooth . If γ is a smooth path through $x \in M$ then if follows from the chain rule that

$$f \circ \gamma = (f \circ \psi^{-1}) \circ (\psi \circ \gamma)$$

is smooth. Hence we can differentiate the function $f \circ \gamma$ at $t = 0$. By the chain rule we have that

$$(f \circ \gamma)'(0) = d(f \circ \psi^{-1})(\psi(x))((\psi^i \circ \gamma)'(0)).$$

It follows that $(f \circ \gamma)'(0) = (f \circ \rho)'(0)$ if ρ and γ are in the same tangency class. Hence if $X = t_0(\gamma)$ is a tangent vector in T_xM we can define

$$df(x)(X) = (f \circ \gamma)'(0).$$

We call this the rate of change of f in the direction X. Notice that we can calculate $df(x)(X)$ without explicit reference to the path γ by the formula

$$df(x)(X) = d(f \circ \psi^{-1})(\psi(x))d\psi(x)(X).$$

As we vary the tangent vector X we define a map

$$df(x)\colon T_xM \to \mathbb{R}$$

called the differential of f at x. This map satisfies the formula

$$df(x) = d(f \circ \psi^{-1})(\psi(x)) \circ d\psi(x)$$

and hence, being a composition of linear maps, is linear.

We call the set of linear maps from T_xM to $\mathbb{R}$ the *cotangent* space to M at x and denote it by T_x^*M. So we have

$$df(x) \in T_x^*M.$$

Elements of T_x^*M are also called *one-forms*.

If $f\colon M \to V$ is a smooth function into a vector space we can also define $df(x)$ by

$$df(x) = f \circ \gamma'(0)$$

which is a vector in V. The one-form df is then an example of a vector space valued one-form.

4.2 Co-ordinate tangent vectors and one-forms.

Let (U, ψ) be a set of co-ordinates on M where $\psi = (\psi^1, \dots, \psi^n)$. Then each of the component functions ψ^i is a real function so we can define n one-forms $d\psi^i(x) \in T_x^*M$ called the co-ordinate one-forms.

We have seen that

$$d\psi^{-1}\colon T_xM \to \mathbb{R}^n$$

is a linear isomorphism. We denote by

$$\frac{\partial}{\partial \psi^i}(x) = d\psi^{-1}(\psi(x))(e^i)$$

the image under this map of the standard basis vector e^i in $\mathbb{R}^n$. We call the set of these the basis of co-ordinate tangent vectors. Consider what happens when we apply $d\psi^i(x)$ to $\partial/\partial\psi^j(x)$. We have

$$d\psi^i(x)\left(\frac{\partial}{\partial \psi^j}(x)\right) = d\psi^i(x)(d\psi^{-1}(x)(e^j)) = d(\psi^i \circ \psi^{-1})(\psi(x))(e^j)).$$

Notice that if $y = (y^1, \dots, y^n)$ is a point in $\mathbb{R}^n$ then $\psi^i \circ \psi^{-1}(y) = y^i$ is a linear map so equal to its own derivative. Hence $d(\psi^i \circ \psi^{-1})(\psi(x))(e^j)$ is the ith component of the vector e^j or just δ_{ij}. It follows from linear algebra that $d\psi^1(x), \dots, d\psi^n(x)$ is a basis of T_xM and, in fact, the dual basis to the basis $\partial/\partial\psi^i$.

4.3 How to calculate.

It is useful for calculations to know how to expand various quantities in these co-ordinate bases. First let f be a smooth function on M then we must have

$$df(x) = \sum_{i=1}^{n} a_i d\psi^i(x)$$

for some real numbers a^i. This is just linear algebra as is the fact that if we apply both sides of this equation to $\partial/\partial\psi^j(x)$ and use the dual basis relation we deduce that

$$a^i = df(x)\left(\frac{\partial}{\partial\psi^i}(x)\right)$$

we *define*

$$\frac{\partial f}{\partial\psi^i}(x) = df(x)\left(\frac{\partial}{\partial\psi^i}(x)\right)$$

and hence have the formula:

$$df(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial\psi^i}(x) d\psi^i(x).$$

If γ is a path through x then its tangent at 0, $\gamma'(0)$ can be expanded as

$$\gamma'(0) = \sum_{i=1}^{n} b^i \frac{\partial}{\partial\psi^i}(x).$$

Applying $d\psi^j(x)$ to both sides and using the chain rule we deduce that

$$\gamma'(0) = \sum_{i=1}^{n} (\psi^i \circ \gamma)'(0) \frac{\partial}{\partial\psi^i}(x).$$

4.4 Tangent vectors as differential operators

There is a useful equivalent definition of tangent vectors implicit in the notation $\partial/\partial\psi^i$. Note that the set $C^\infty(M, \mathbb{R})$ of all smooth functions is an algebra over $\mathbb{R}$. If $m \in M$ we define:

Definition 4.5. An m-derivation of $C^\infty(M, \mathbb{R})$ is a linear map $\delta \colon C^\infty(M, \mathbb{R}) \to \mathbb{R}$ satisfying

$$\delta(fg) = \delta(f)g(m) + f(m)\delta(g)$$

for all $f, g \in C^\infty(M, \mathbb{R})$.

If $X \in T_m M$ then it is easy to check that $\delta_X(f) = X(f)$ is an m-derivation. In fact we have

Proposition 4.6. *If $\delta \colon C^\infty(M, \mathbb{R}) \to \mathbb{R}$ is an m-derivation then there exists $X \in T_m M$ with $\delta = \delta_X$.*

Proof. First assume that f is zero in an open neighbourhood U of m and that χ is a smooth bump function with support in U and 1 in a smaller open set containing m. Then $f\chi = 0$ so that $0 = \delta(f\chi) = f(m)\delta(\chi) + \delta(f)\chi(m) = \delta(f)$. Also if f is any function and we choose a similar χ then $0 = \delta(f(1-\chi))$ so that $\delta(f) = \delta(f\chi)$. Assume then that there is some $X \in T_x M$ such that $\delta(f) = \delta_X(f)$ whenever f has support in the domain of a co-ordinate chart U. Then choosing χ as above but with support in the same domain we have $\delta(f) = \delta(f\chi) = \delta_X(f\chi) = \delta_X(f)$. It suffices then to prove the result for the manifold U or equivalently for an open ball in $\mathbb{R}^n$ with $x = 0$. In such a case, using the fundamental theorem of calculus we have

$$f(x) = f(0) + \sum_{i=1}^{n} x^i \int_0^1 \frac{\partial f}{\partial x^i}(tx)dt \tag{4.1}$$
$$= f(0) + \sum_{i=1}^{n} x^i g_i(x) \tag{4.2}$$
$$\tag{4.3}$$

where

$$g_i(x) = \int_0^1 \frac{\partial f}{\partial x^i}(tx)dt$$

so that

$$g_i(0) = \frac{\partial f}{\partial x^i}(0).$$

It follows that we have

$$\chi f = \chi f(0) + \sum_{i=1}^{n} \chi x^i g_i$$

and thus

$$\delta(\chi f) = 0 + \sum_{i=1}^{n} \delta(\chi x^i)\frac{\partial f}{\partial x^i}(0)$$

If we let

$$X = \sum_{i=1}^{n} \delta(\chi x^i)\frac{\partial}{\partial x^i}$$

then

$$\delta_X(\chi f) = \delta_X(f) = \sum_{i=1}^{n} \delta(\chi x^i)\frac{\partial f}{\partial x^i}(0) = \delta(f).$$

□

Notice that the set of all m-derivations is naturally a vector space which is isomorphic to the tangent space at m.

5 Submanifolds

Historically the theory of differential geometry arose from the study of surfaces in $\mathbb{R}^3$. We want to consider the more general case of submanifolds in $\mathbb{R}^n$. Recall that we have

Definition 5.1. A set $Z \subset \mathbb{R}^n$ is a submanifold of dimension d if there is a smooth map $f \colon \mathbb{R}^n \to \mathbb{R}^{n-d}$ for some d such that $Z = f^{-1}(0)$ and $df(z)$ is onto for all $z \in Z$.

We will show that any submanifold of $\mathbb{R}^n$ is a manifold by constructing co-ordinate charts. The idea is simple. At any point $z \in Z$ we define the subspace tangent to Z. This is the kernel K_z of the map $df(z)$. Then we consider orthogonal projection of K_z onto Z. The inverse of this map defines co-ordinates on an open neighbourhood of z. In fact these are a very special kind of co-ordinate. To prove this we first prove

Proposition 5.2. *Let Z be a submanifold of $\mathbb{R}^n$ of dimension d. Then at any $z \in Z$ we can find a co-ordinate chart (U, ψ) on $\mathbb{R}^n$ such that if $\psi = (\psi^1, \dots, \psi^n)$ then*

$$U \cap Z = \{x \in U \mid \psi^{d+1}(x) = \dots = \psi^n(x) = 0\}.$$

Proof. Let K_z be the kernel of $df(z)$ and let $\pi\colon \mathbb{R}^n \to K_z$ be the orthogonal projection onto K_z. Choose a basis $v^1, \dots, v^d$ and write π_z with respect to this basis as

$$\pi(x) = \sum_{i=1}^{d} \pi^i(v)v^i.$$

Define a map $\chi\colon \mathbb{R}^n \to \mathbb{R}^n$ by

$$\psi(x) = (\pi^1(x), \dots, \pi^d(x), f^1(x), \dots, f^{n-d}(x)).$$

Because π is a linear map it is its own derivative so we have

$$d\psi(z) = (\pi^1(z), \dots, \pi^d(z), df^1(z), \dots, df^{n-d}(z)).$$

Consider v in the kernel of $d\psi(z)$. Then $d\pi(z)(v) = 0$ and $df(z)(v) = 0$. Hence v is both orthogonal to K_z and in K_z so it must be zero. So $d\psi(z)$ is injective and hence a dimension count surjective so a bijection. Now we can apply the inverse function theorem so there is an open set U in $\mathbb{R}^n$ such that $\psi(U)$ is open and

$$\psi_{|U}\colon U \to \psi(U)$$

is a diffeomorphism. But this just means that $(U, \psi_{|U})$ is a co-ordinate chart on $\mathbb{R}^n$. Notice that $\psi^{d+1}(x) = \cdots = \psi^n(x) = 0$ if and only if $f(x) = 0$ if and only if x is in $U \cap Z$. □

Let (U, ψ) be a set of co-ordinates on $\mathbb{R}^n$ such that

$$U \cap Z = \{x \in U\colon \mid \psi^{d+1}(x) = \cdots = \psi^n(x) = 0\}.$$

Then consider $(U \cap Z, \bar{\psi})$ where $\bar{\psi} = (\psi^1_{|U\cap Z}, \dots, \psi^{n-d}_{|U\cap Z})$. This is a co-ordinate chart. The only thing to check is that that because $\psi(U)$ is open then

$$\bar{\psi}(U \cap Z) = \{x \in \mathbb{R}^d \mid (x^1, \dots, x^d, 0, \dots, 0) \in \psi(U)\}$$

is open. This is an elementary fact about the topology of $\mathbb{R}^n$. Consider two such co-ordinate charts (U, ψ) and (V, χ). We will prove that $(U \cap Z, \bar{\psi})$ and $(V \cap Z, \bar{\chi})$ are compatible. This follows essentially from the fact that (U, ψ) and (V, χ) are compatible. First we note that

$$\bar{\chi}(U \cap V \cap Z) = \{x \in \mathbb{R}^d \mid (x^1, \dots, x^d, 0, \dots, 0) \in \chi(U \cap V)\}$$

and, again, this is open. For the smoothness of the co-ordinate change map we note that

$$\bar{\psi}^i \circ \bar{\chi}^{-1}_{|\chi(U\cap V\cap Z)}(x) = \psi^i \circ \chi^{-1}_{|\chi(U\cap V)}(x, 0)$$

where $(x, 0) = (x^1, \dots, x^d, 0, \dots, 0)$. Hence the result follows.

We have now proved

Theorem 5.3. *If Z is a submanifold the set of charts above is an atlas.*

This result gives us many examples of submanifolds:

Example 5.1 (Spheres). Consider the sphere $S^n \subset \mathbb{R}^{n+1}$ defined as the set of points x whose length $\|x\|$ is equal to one. Here

$$\|x\|^2 = \sum_{i=1}^{n} (x^i)^2.$$

We can prove it is a submanifold of $\mathbb{R}^n$ and hence a manifold by considering the map $f\colon \mathbb{R}^{n+1} \to 1$ defined by

$$f(x) = \|x\|^2 - 1.$$

Clearly this is smooth and has zero set Z equal to the sphere S^n. To check that the derivative is smooth note that

$$df(x)(v) = 2\langle x, v\rangle$$

and is a linear map onto a one-dimensional space so to show it is onto we just need to show that it is not equal to the zero linear map.

Example 5.2 (The orthogonal group). The orthogonal group is the group of all linear transformations of $\mathbb{R}^n$ that preserve the usual inner product on $\mathbb{R}^n$. We shall think of it as a group of n by n matrices:

$$O(n) = \{X \mid XX^t = 1\}.$$

We can identify the set of all n by n matrices with $\mathbb{R}^{n^2}$. There are various ways of doing this. So as to be concrete let us assume we have done it by writing down the rows one after the other. With this identification in mind define a smooth map $f\colon \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$ by $f(X) = XX^t - 1$. It is clear we have $f^{-1}(0) = O(n)$. Define the linear subspace $S \subset \mathbb{R}^{n^2}$ to be the set of all symmetric matrices. This can be identified with $\mathbb{R}^d$ where $d = n(n+1)/2$. It is easy to check that f takes its values in S so we will think of f as a smooth map $f\colon \mathbb{R}^{n^2} \to \mathbb{R}^d$.

We want to calculate $df(X)$ the derivative of f at a matrix X. By differentiating the path $t \mapsto X + tY$ we obtain

$$df(X)(Y) = YX^t + XY^t.$$

If B is any symmetric matrix then it is easy to check that $df(X)((1/2)BX) = B$ if we use that fact that $XX^t = 1$ and $B = B^t$. We have therefore shown that $df(X)$ is onto for any $X \in O(n)$ so that $O(n)$ is a submanifold of dimension $n - d = n(n-1)/2$.

5.1 Tangent space to a submanifold of $\mathbb{R}^n$

If Z is a submanifold of $\mathbb{R}^n$ then there is a natural notion of the plane tangent to Z at any point x independent of abstract notions such as equivalence classes of paths and co-ordinates. It is just the subspace of $\mathbb{R}^n$ tangential to Z at x. More precisely if $Z = f^{-1}(0)$ it is the kernel of $df(x)$ which we denoted by K_x. To relate this to the abstract notion of tangent vector consider a smooth path

$$\gamma\colon (-\epsilon, \epsilon) \to Z.$$

Because $Z \subset \mathbb{R}^n$ this is naturally a path in $\mathbb{R}^n$. We check first that this is smooth. To do this choose co-ordinates (U, ψ) for $\mathbb{R}^n$ about x satisfying

$$U \cap Z = \{x \in Z\colon \mid \psi^{d+1}(x) = \cdots = \psi^n(x) = 0\}.$$

and denote by $\bar{\psi}$ the corresponding co-ordinates on $U \cap Z$. Smoothness of γ means that the functions $\hat{\psi}^i \circ \gamma = \psi^i \circ \gamma$ are smooth for each $i = 1, \dots, d$. Because γ has image inside Z we also have that $\psi^i \circ \gamma = 0$ for each $i = d+1, \dots, n$ and hence these are also smooth. So γ is a smooth path in $\mathbb{R}^n$. Consider the vector $\gamma'(0)$ in $\mathbb{R}^n$. We have that $f \circ \gamma(t) = 0$ for all t so by the chain rule $df(x)(\circ'(0)) = 0$ so that $\gamma'(0) \in K_x$.

We can now define a map $T_x Z \to K_x$. If $X \in T_x Z$ then we choose a path γ whose tangent vector at 0 is X and map X to $\gamma'(0) \in K_x$. We have to check first that this is well-defined. Let ρ be another such path and consider the co-ordinates $\bar{\psi}$. By definition we have

$$(\bar{\psi}^i \circ \gamma)'(0) = (\bar{\psi}^i \circ \rho)'(0)$$

for every $i = 1, \dots, d$. Hence we also have

$$(\psi^i \circ \gamma)'(0) = (\psi^i \circ \rho)'(0)$$

for every $i = 1, \dots, d$. But

$$(\psi^i \circ \gamma)'(0) = (\psi^i \circ \rho)'(0) = 0$$

for $i = d+1, \dots, n$ so we have

$$(\psi^i \circ \gamma)'(0) = (\psi^i \circ \rho)'(0)$$

for $i = 1, \dots, n$. Hence X maps to the same element of K_x whether we use γ or ρ. To show that this map is injective we use a similar argument. It is easy to see that this map is linear. Hence, counting, dimensions we see that this is a linear isomorphism.

We conclude that if $Z \subset \mathbb{R}^n$ is a submanifold and we consider the tangents to all the paths through $x \in Z$, thought of as maps into $\mathbb{R}^n$ then they span the space K_x.

5.2 Exercise

Exercise 5.1. Show that the set defined by the equation

$$r^2 - a^2 = (\sqrt{x^2 + a^2} - a)^{1/2}$$

is a smooth submanifold of $\mathbb{R}^3$ if a and r are real numbers with $r < a$.

Exercise 5.2. Show that the following subset of $\mathbb{R}^3$ is a submanifold:

$$Q = \{(x, y, z) \mid x^2 + y^2 + z^2 = 9 \quad \text{and} \quad x + y + z = 3\}.$$

5.3 Smooth functions between manifolds

The definition of a smooth function on a manifold and a smooth path can all be subsumed in the following definition.

Definition 5.4. Let $f\colon M \to N$ be a map between manifolds. Then f is called smooth if for every point $x \in M$ there is a co-ordinate chart (U, ψ) on M and (V, χ) on N such that $x \in U$, $f(U) \subset V$ and

$$\chi \circ f \circ \psi^{-1}\colon \psi(U) \to \chi(V)$$

is smooth.

Again we have the usual lemma:

Lemma 5.5. *Let $f\colon M \to N$ be a smooth map between manifolds. Assume that there are co-ordinate charts (U, ψ) on M and (V, χ) on N such that $f(U) \subset V$. Then*

$$\chi \circ f \circ \psi^{-1}\colon \psi(U) \to \chi(V)$$

is smooth.

5.4 The tangent to a smooth map.

If $f\colon M \to N$ is a smooth map and $\gamma\colon (-\epsilon, \epsilon) \to M$ is a smooth path through x then $f \circ \gamma$ is a smooth path in N through $f(x)$. Moreover if we consider another path ρ which is tangent to γ then $f \circ \gamma$ and $f \circ \rho$ are tangent. To see this choose co-ordinates (U, ψ) and (V, χ) with $f(U) \subset V$. Assume without loss of generality that $\gamma(-\epsilon, \epsilon)$ and $\rho(-\epsilon, \epsilon)$ are in U. Then we have

$$\chi \circ (f \circ \gamma)'(0) = d(\chi \circ f \circ \psi^{-1})(\psi(x))(\psi \circ \gamma)'(0)$$

and

$$\chi \circ (f \circ \rho)'(0) = d(\chi \circ f \circ \psi^{-1})(\psi(x))(\psi \circ \rho)'(0)$$

so that $(\psi \circ \rho)'(0) = (\psi \circ \gamma)'(0)$ implies that $\chi \circ (f \circ \rho)'(0) = \chi \circ (f \circ \gamma)'(0)$ and hence $f \circ \gamma$ and $f \circ \rho$ are tangent. So associated with f there is a well-defined map from $T_x M$ to $T_{f(x)} N$ that sends $\gamma'(0)$ to $(f \circ \gamma)'(0)$. This map is denoted $T_x f$ and called the tangent to f at x. So we have that

$$T_x(f)(\gamma'(0)) = (f \circ \gamma)'(0).$$

Notice that the tangent map satisfies

$$T_x(f) = d\chi(f(x))^{-1} \circ T_x(f) \circ d\psi(x).$$

so that, being a composition of three linear maps it is itself linear. Moreover this formula also shows that with respect to the bases of $T_x M$ and $T_{f(x)} N$ given by the co-ordinate vector fields we have

$$T_x(M)(\frac{\partial}{\partial \psi}(x)) = \sum_{i=1}^{n} \frac{\partial \chi^j \circ f}{\partial \psi^i}(x) \frac{\partial}{\partial \chi^j}(f(x)).$$

In other words it is given by the action of a matrix whose entries are the partial derivatives of the co-ordinate expression for f.

Example 5.3. The tangent space to $\mathbb{R}^n$ at $\psi(x)$ is just $\mathbb{R}^n$ again. The map $d\psi(x)\colon T_xM \to \mathbb{R}^n$ is just the map $T_x\psi\colon T_xM \to T_{\psi(x)}\mathbb{R}^n$.

Example 5.4. If $F\colon \mathbb{R}^n \to \mathbb{R}^m$ is a smooth map then, after identifying $T_x\mathbb{R}^n$ with $\mathbb{R}^n$ and $T_{F(x)}\mathbb{R}^m$ with $\mathbb{R}^m$ we see that the tangent map $T_x(F)$ is just the matrix of partial derivatives $dF(x)$.

The chain rule for smooth functions in $\mathbb{R}^n$ generalises to manifolds as follows.

Proposition 5.6 (Chain Rule). *Let $f\colon M \to N$ and $g\colon N \to P$ be smooth functions. Then the map $g \circ f\colon M \to P$ is smooth and $T_x(g \circ f) = T_{f(x)}(g) \circ T_x(f)$.*

If $f\colon M \to V$ is a smooth map from a manifold to a vector space then we can define, as for a real valued function, a derivative by

$$df(x)(t_0(\gamma)) = f \circ \gamma'(0)$$

which is a vector on V. On the other hand we have just defined

$$T_x(f)\colon T_x(M) \to T_{f(x)}(V).$$

To understand the relation between these two notions of derivative recall from Example 4.2 that we have seen that the tangent spaces to a vector space are naturally identified with the vector space itself by differentiating each path. If we compose $T_x(f)$ with this identification $T_{f(x)}(V) \simeq V$ then we obtain $df(x)$.

5.5 Submanifolds again.

If M is a manifold then we can define a submanifold of M by using the principal property of submanifolds in $\mathbb{R}^n$.

Definition 5.7 (Submanifolds.). We say that a subset $N \subset M$ is a submanifold of dimension d of a manifold M of dimension m if for every $x \in N$ we can find a co-ordinate chart (U, ψ) for M with $x \in N$ and such that

$$U \cap N = \{y \in N\colon \mid \psi^{d+1}(y) = \cdots = \psi^n(y) = 0\}.$$

Just as before we can define co-ordinates $(U \cap N, \bar{\psi})$ on N by letting

$$\bar{\psi}^i(y) = \psi^i(y)$$

for each $i = 1, \ldots, d$. Similarly we have

Proposition 5.8. *The set consisting of all the charts $(U \cap N, \bar{\psi})$ constructed in this manner is an atlas. Moreover it makes N a manifold in such a way that the inclusion map $\iota_N\colon N \to M$ defined by $\iota_N(n) = n$ is smooth.*

Because the condition for being a submanifold is local we can use the inverse function theorem as in Proposition 5.2 to prove

Proposition 5.9. *Let $f\colon M \to N$ be a smooth map between manifolds of dimension m and n respectively. Let $n \in N$ and $Z = f^{-1}(n)$. Then if T_zf is onto for all $z \in M$ the set Z is submanifold of M. Moreover the image of ι_Z in T_zM is precisely the kernel of T_zF.*

6 Vector fields.

We have seen how to define tangent vectors at a point of a manifold. In many problems we are interested in vector fields X, that is a choice of vector $X(x) \in T_xM$ at every point of a manifold. We need to make sense of the notion of such a vector $X(x)$ depending smoothly on x. We do this as follows. Choose a chart (U, ψ). Then at every point $x \in U$ we have a basis

$$\frac{\partial}{\partial\psi^1}(x), \ldots, \frac{\partial}{\partial\psi^n}(x)$$

of $T_x M$ and we can expand $X(x)$ as a linear combination of these tangent vectors:

$$X(x) = \sum_{i=1}^{n} X^i(x) \frac{\partial}{\partial \psi^i}(x).$$

We call the functions $X^i \colon U \to \mathbb{R}$ the *components* of the vector field with respect to the co-ordinate chart. We have

Definition 6.1. A vector field X on a manifold M is smooth if its components with respect to a collection of co-ordinate charts whose domains cover M are all smooth.

We have the usual lemma.

Lemma 6.2. *If X is a smooth vector field then its components with respect to any co-ordinate chart are smooth.*

Proof. Let (U, ψ) be a co-ordinate chart and let $x \in U$. Choose a co-ordinate chart (V, χ) with $x \in V$ and such that the components of X are smooth with respect to (V, χ). Then write

$$X = \sum_{i=1}^{n} X^i(x) \frac{\partial}{\partial \psi^i}(x) = \sum_{a=1}^{n} X^i(x) \frac{\partial}{\partial \chi^a}(x).$$

From the results of section 4.3 we have

$$d\chi^a = \sum_{i=1}^{n} \frac{\partial \chi^a}{\partial \psi^i} d\psi^i$$

so using the property of dual bases we have.

$$\frac{\partial}{\partial \psi^i} = \sum_{a=1}^{n} \frac{\partial \chi^a}{\partial \psi^i} \frac{\partial}{\partial \chi^a}.$$

Hence

$$X^i(y) = \sum_{a=1}^{n} \frac{\partial \psi^i}{\partial \chi^a}(y) X^a(y)$$

for all $y \in U \cap V$ so that the X^i are smooth on $U \cap V$ and hence smooth on all of U. □

Classical texts on differential geometry, in particular those on tensor calculus, downplay the co-ordinates and charts and concentrate on the components of vector fields and similar tensors. Assume that M is covered by the domains of co-ordinate charts (U_α, ψ_α). For each chart (U_α, ψ_α) we write

$$X_{|U_\alpha} = \sum_{i=1}^{n} X^i_\alpha \frac{\partial}{\partial \psi^i} \tag{6.1}$$

and then as in the proof above we have that for x in the intersection of U_α and U_β we have

$$X^i_\alpha(x) = \sum_{j=1}^{n} \frac{\partial \psi^i_\alpha}{\partial \psi^j_\beta}(x) X^j_\beta(x). \tag{6.2}$$

The converse is also true. If we have a collection of maps $X^i_\alpha \colon U_\alpha \to \mathbb{R}$ satisfying 6.2 then we can define a vector field using 6.1 and check that it is well-defined. Classical and physics texts generally suppress the α index and also the sum by applying the Einstein summation convention. This convention is that any index that occurs in an expression in both a raised and lowered position is summed over. So a typical writing of 6.2 would be to say that we have co-ordinates x^i and co-ordinates $x^{i'}$ and that the vector field transforms as

$$X^i = \frac{\partial x^i}{\partial x^{j'}} X^{j'}.$$

Notice that even if we do not exploit the Einstein summation convention it is a useful guide to memorising expressions like. To apply it correctly we need to remember that the index on a co-ordinate is a superscript.

6.1 The Lie bracket.

One use of this discussion is the definition of the Lie bracket of two vector fields. Let X and Y be two vector fields and write them locally as

$$X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial \psi^i}$$

and

$$Y = \sum_{i=1}^{n} Y^i \frac{\partial}{\partial \psi^i}.$$

so that

Then define

$$[X, Y] = \sum_{i,j=1}^{n} (X^j \frac{\partial Y^i}{\partial \psi^j} - Y^j \frac{\partial X^i}{\partial \psi^j}) \frac{\partial}{\partial \psi^i}.$$

We leave it as an exercise to show that this transforms as a vector field. We call $[X, Y]$ the Lie bracket of the vector fields X and Y. Lie is named after Sophus Lie and pronounced "lee".

6.2 Exercise

Exercise 6.1. Let X and Y be vector fields on a manifold M. Define a new vector field $[X, Y]$ by defining it in local co-ordinates (U, ϕ) by

$$[X, Y]_{|U} = \sum_{i,j} (X_i \frac{\partial}{Y}^j \partial \phi^i - Y_i \frac{\partial X^j}{\partial \phi^i}) \frac{\partial}{\partial \phi^j}.$$

Show that this makes sense. That is it doesn't really depend on the choice of co-ordinates. The vector field $[X, Y]$ is called the Lie bracket of X and Y.

6.3 Vector fields and the tangent bundle.

There is a more sophisticated way to define vector fields which we now consider, partly for its own interest and partly because it is the motivating example of the notion of a vector bundle.

Let M be a manifold and take the union of all the tangent spaces, denote it by

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

and call it the *tangent bundle* to M. There is an important map $\pi: TM \to M$ called the *projection* that sends a vector $X \in T_x M$ to the point $\pi(X) = x$ at which it is located. A vector field is a map $X: M \to TM$ with the special property that $X(x) \in T_x M$. This property can be also written as $\pi \circ X = \mathrm{id}_M$, that is $\pi(X(x)) = x$. Such a map $X: M \to TM$ is called a *section* of the projection map π. We want to consider smooth vector fields and as we already have a notion of smooth function between manifolds the simplest way to define smooth vector fields is to make TM a manifold. To do this involves a construction that we will use again later so we ill state it in more general form than is immediately necessary.

Let E be a set with a surjection $\pi: E \to M$ where M is a manifold. Denote by E_x the fibre of E over x, that is the set $\pi^{-1}(x)$. Let V be a finite dimensional vector space. Assume that we can cover M by co-ordinate charts $\{(U_\alpha, \psi_\alpha\}_{\alpha \in I}$ such that for every $\alpha \in I$ and for every $x \in U_\alpha$ there is a bijection

$$\phi_\alpha(x): E_x \to V$$

such that the map

$$\begin{array}{ccc} U_\alpha \cap U_\beta & \to & GL(V) \\ x & \mapsto & \phi_\alpha(x) \circ \phi_\beta(x)^{-1} \end{array}$$

is smooth where $GL(V)$ is the group of all linear isomorphisms of V. Then it is possible to make E a manifold as follows. We define bijections

$$\begin{array}{rcll} \chi_a & : & \pi^{-1}(U) & \to & U \times V \\ & & x & \mapsto & (\pi(x), \phi_\alpha(\pi(x))v) \end{array}$$

To make these into charts we should really identify V with some $\mathbb{R}^k$ but we will not bother to do that. To check compatibility we note that $\chi_\alpha(U_\alpha \cap U_\beta) = U_\alpha \cap U_\beta \times V$ which is open in $\mathbb{R}^n \times V$. Likewise for $\chi_\beta(U_\alpha \cap U_\beta)$. Then the map we want to check is smooth is the map

$$U_\alpha \cap U_\beta \times V \to U_\alpha \cap U_\beta \times V$$

which sends (x, v) to $(x, \phi_\alpha(x) \circ \phi_\beta(x)^{-1}v)$ and this is smooth and invertible. By interchanging α and β we deduce that this map is a diffeomorphism. Hence we have made E into a manifold. Notice that with this manifold structure the map χ_α is a diffeomorphism, as the co-ordinate charts of a manifold are diffeomorphisms. Notice also that each E_x is a vector space from Lemma 2.12. Moreover it easy to check that the addition and scalar multiplication are smooth. Define a section of $\pi\colon E \to M$ to be a smooth map $s\colon M \to E$ which satisfies $s(x) \in E_x$ for all $x \in M$. If s is such a section then on restriction to U_α we can define a map $s_a\colon U \to V$ by $s_\alpha(x) = \phi_\alpha(x)(s(x))$. The s_a are clearly smooth. The converse is also true if s is any map and the s_a defined in this way are smooth then s is smooth.

Consider now the case of the tangent bundle. Let (U_α, ψ_α) be a co-ordinate chart on M. Then $V = \mathbb{R}^n$ and $\phi_\alpha(x) = d\psi_a(x)$. The condition we require to hold is that the map

$$x \mapsto d\phi_\beta(\psi^{-1}(x)) \circ d\psi_\alpha^{-1}(x) = d(\psi_\beta \circ \psi_a^{-1})(x) = d_i(\psi_\beta^j \circ \psi_\alpha^{-1})(x)$$

is smooth. But this is just the Jacobian matrix of partial derivatives which depends smoothly on x.

It is straightforward to now check the following Proposition.

Proposition 6.3. *A vector field on a manifold M is smooth if and only if it is a smooth section of the tangent bundle.*

To understand what it means to be smooth in terms of co-ordinates recall the definition of $d\psi(x)$. We have the co-ordinate vector fields $(\partial/\partial\psi^i)(x)$ for $i = 1, \ldots, n$. Then

$$d\psi(x)(\frac{\partial}{\partial\psi^i})(x) = e^i$$

where e^i is the standard basis vector of $\mathbb{R}^n$. So clearly this is a smooth map so that the co-ordinate vector fields are smooth.

More generally if X is a vector field we can write it as

$$X(x) = \sum_{i=1}^n X^i(x)\frac{\partial}{\partial\psi^i}(x)$$

for any $x \in U$, and functions $X^i\colon U \to \mathbb{R}$. Then

$$d\psi(x)(X(x)) = (X^1(x), \ldots, X^n(x)).$$

This proves

Proposition 6.4. *Let X be a vector field on a manifold M. Then if X is smooth and (U, ψ) is a co-ordinate chart then if we let*

$$X(x) = \sum_{i=1}^n X^i(x)\frac{\partial}{\partial\psi^i}(x)$$

the functions $X^i\colon U \to \mathbb{R}$ are smooth. Conversely if X is a vector field and we can cover M with co-ordinate charts (U, ψ) such that the corresponding $X^i\colon U \to \mathbb{R}$ are smooth then X is smooth.

6.4 Vector fields and derivations.

Recall from Section 4.4 that the tangent vectors at $m \in M$ can be identified with the m-derivations of the algebra $C^\infty(M, \mathbb{R})$. To understand what this means for vector fields we first have:

Definition 6.5. A derivation of $C^\infty(M)$ is a linear map

$$D\colon C^\infty(M, \mathbb{R}) \to C^\infty(M, \mathbb{R})$$

such that

$$D(fg) = D(f)g + fD(g).$$

A vector field X gives rise to a derivation $f \mapsto X(f)$ and using the previous Lemma we have

Proposition 6.6. *Every derivation arises from a vector field.*

Proof. Let D be a derivation. Then note that for any x $f \mapsto D(f)(x)$ is a derivation at x. Hence there is a tangent vector $X(x)$ such that $D(f)(x) = X(x)(f)$ for all x. We have to check that $X(x)$ depends smoothly on x. But if we choose local co-ordinates ψ as in the proof for m-derivations and extend them to global functions ψ then we have

$$X(x) = \sum_{i=1}^{n} D(\psi^i)(x)\frac{\partial}{\partial \psi^i}(x)$$

but $D(\psi)$ is a smooth function so X is smooth. □

The advantage of thinking of a vector field as a derivation is that derivations have a natural bracket operation. If D and D' are two derivations then a simple calculation shows that $[D, D']$ defined by

$$[D, D'](f) = D(D'(f)) - D'(D(f)).$$

is also a derivation. So we can define the bracket of two vector fields X and Y and called the Lie bracket $[X, Y]$. To calculate $[X, Y]$ we apply it to ψ^i then we have

$$[X, Y](\psi^i) = X(Y(\psi^i)) - Y(X(\psi^i))$$

so that if

$$X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial \psi^i}$$

and

$$Y = \sum_{i=1}^{n} Y^i \frac{\partial}{\partial \psi^i}$$

so that

$$[X, Y](\psi^i) = \sum_{j=1}^{n} X^j \frac{\partial Y^i}{\partial \psi^j} - Y^j \frac{\partial X^i}{\partial \psi^j}.$$

Hence

$$[X, Y] = \sum_{i,j=1}^{n} (X^j \frac{\partial Y^i}{\partial \psi^j} - Y^j \frac{\partial X^i}{\partial \psi^j})\frac{\partial}{\partial \psi^i}.$$

7 Tensor products and tensor fields.

If V and W are finite dimensional vector spaces then the Cartesian product $V \times W$ is naturally a vector space called the direct sum of V and W and denoted $V \oplus W$. The tensor product is a more complicated object. To define it we start by defining for any *set* X the free vector space over X, $F(X)$. This is the set of all maps from X to $\mathbb{R}$ which are zero except at a finite number of points. We define the vector space structure by adding and scalar multiplying maps. Each x gives rise to a function $\delta(x)$ which is one at x and zero elsewhere. We therefore have a map $\delta\colon X \to F(X)$. By construction the span of the image of δ is all of $F(X)$.

The special property of the free vector space over X is the following.

Proposition 7.1. *Let $f\colon X \to U$ be any map from X into a vector space U then there is a unique linear map $\hat{f}F(X) \to U$ such that $\hat{f} \circ \delta = f$.*

Proof. The general element of $F(X)$ is

$$\sum_{i=1}^{n} a_i \delta(x_i)$$

for $a_i \in \mathbb{R}$. We define

$$\hat{f}(\sum_{i=1}^{n} a_i \delta(x_i)) = \sum_{i=1}^{n} a_i f(x^i).$$

□

Given two vector spaces V and W we can define $F(V \times W)$. This is an infinite dimensional vector space. We shall denote $\delta((v, w))$ by $\delta(v, w)$. Consider the subspace Z defined as the span of all elements of the form

$$\delta(\lambda v + \mu v', w) - \lambda\delta(v, w) - \mu\delta(v', w)$$

and

$$\delta(v, \lambda w + \mu w') - \lambda\delta(v, w) - \mu\delta(v, w')$$

for any real numbers λ and μ and vectors $v, v' \in V$ and $w, w' \in W$. Let us denote

$$V \otimes W = F(V \times W)/Z$$

and define a map

$$\otimes\colon V \times W \to V \otimes W$$

by

$$v \otimes w = \delta(v, w) + Z.$$

We have

Proposition 7.2. *The map $\otimes\colon V \times W \to V \otimes W$ is bilinear.*

Proof. We check the first factor only

$$\begin{aligned}
(\lambda v + \mu v') \otimes w =& \delta(\lambda v + \mu v', w) + Z \\
=& \delta(\lambda v + \mu v', w) - \lambda\delta(v, w) \\
& - \mu\delta(v', w) + \lambda\delta(v, w) + \mu\delta(v', w) + Z \\
=& \lambda\delta(v, w) + \mu\delta(v', w) + Z \\
=& \lambda(\delta(v, w) + Z) + \mu(\delta(v', w) + Z) \\
=& \lambda v \otimes w + \mu v' \otimes w
\end{aligned}$$

□

From Proposition 7.1 we know that any map $f\colon V \times W \to U$, where U is a vector space extends to a map $\hat{f}\colon F(V \times W) \to U$. Standard linear algebra tells us that we can take the quotient to get a map $\tilde{f}\colon V \otimes W \to U$ if $\hat{f}(Z) = 0$. The map is defined by $v \otimes w \to f(v, w)$. For example if $v^* \in V^*$ and $w^* \in W^*$ then $v \otimes w \to v^*(v)w^*(w)$ defines a linear map from $V \otimes W \to \mathbb{R}$.

Let $\{v^1, \dots, v^n\}$ be a basis of V and $\{w^1, \dots, w^m\}$ be a basis of W. Consider the set of mn vectors $v^i \otimes w^j$ in $V \otimes W$. We wish to show that they form a basis. First we check that they span the space $V \otimes W$. As the elements of $V \otimes W$ are finite linear combinations of elements of the form $v \otimes w$ it suffices to show that these are all in the span of the vectors $v^i \otimes w^j$. But this follows from the bilinearity. If $v = \sum_{i=1}^n a_i v^i$ and $w = \sum_{j=1}^m b_j w^j$ then

$$v \otimes w = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j v^i \otimes w^j.$$

To show that they are linearly independent assume that

$$0 = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} v^i \otimes w^j.$$

Let v_i^* and w_j^* be the dual bases of $V^*$and W^*. That is $v_i^*(v^j) = \delta_i^j$ and $w_i^*(w^j) = \delta_i^j$. Then apply the map $V \otimes W \to \mathbb{R}$ defined by v_i^* and w_j^* to this equation to obtain $a_{ij} = 0$. So we have proved.

Proposition 7.3. *If V and W are finite dimensional vector spaces then*

$$\dim(V \otimes W) = \dim(V)\dim(W).$$

We can iterate tensor products. If V and W and U are vector spaces we can form $(V \otimes W) \otimes U$ and $V \otimes (W \otimes U)$. These different vector spaces are in fact isomorphic via the map

$$(v \otimes w) \otimes u \mapsto v \otimes (w \otimes u).$$

We use this map to identify these two spaces and ignore the brackets. We write $V \otimes U \otimes W$ for the triple tensor product. More generally we can form finitely many tensor products.

We also need to know about tensor products of maps. If $X\colon V \to V'$ is linear and $Y\colon W \to W'$ is linear then we can define a map

$$V \times W \to V' \otimes W'$$

by $(v, w) \mapsto X(v) \otimes Y(v)$. This is a bilinear map so factors to a map $V \otimes W \to V' \otimes W'$ which we denote by $X \otimes Y$. It is defined by $(X \otimes Y)(v \otimes w) = X(v) \otimes Y(w)$.

We have seen that any bilinear map $V \times W \to \mathbb{R}$ gives rise to a linear map $V \otimes W \to \mathbb{R}$. It is easy to show that this is an isomorphism. More generally if for any collection of vector spaces $V_1, \dots, V_k$ we denote by $\mathrm{Mult}(V_1 \times \cdots \times V_k, \mathbb{R})$ the space of all multilinear maps from $V_1 \times \dots V_k \to \mathbb{R}$ we have

Proposition 7.4. *If $V_1, \dots, V_k$ are vectors spaces then there is a natural isomorphism*

$$Mult(V_1 \times \cdots \times V_k, \mathbb{R}) \to (V_1 \otimes V_2 \otimes \cdots \otimes V_k)^*$$

defined by

$$f \mapsto (v_1 \otimes \cdots \otimes v_k \mapsto f(v_1, \dots, v_k)).$$

7.1 Tensor fields

Consider now a manifold M. For any $r, s \geq 0$ we can define

$$T_m^{(r,s)}(M) = T_m(M) \otimes \cdots \otimes T_m(M) \otimes T_m^*(M) \otimes \cdots \otimes T_m^*(M)$$

where there are r copies of $T_m(M)$ and s copies of $T_m^*(M)$ so that $T_m^{(1,0)}(M) = T_m(M)$ and $T_m^{(0,1)}M = T_m^*(M)$. $T_m^{(r,s)}(M)$ has a basis given by the tensor products

$$\frac{\partial}{\partial \psi^{j_1}} \otimes \cdots \otimes \frac{\partial}{\partial \psi^{j_r}} \otimes d_m \psi^{i_1} \otimes \cdots \otimes d_m \psi^{i_s}$$

and we can use these to make the (r,s) tensor bundle

$$T^{(r,s)}(M) = \bigcup_{m \in M} T_m^{(r,s)}(M)$$

into a manifold following the approach for vectors. An (r,s) tensor field X is a section of this tensor bundle and locally looks like

$$X(m) = \sum_{1 \le j_1,\dots,j_r,i_1\dots i_s \le n} X_{j_1,\dots,j_r,i_1\dots i_s}(m) \frac{\partial}{\partial \psi^{j_1}} \otimes \cdots \otimes \frac{\partial}{\partial \psi^{j_r}} \otimes d_m \psi^{i_1} \otimes \cdots \otimes d_m \psi^{i_s}$$

The functions $X_{j_1,\dots,j_r,i_1\dots i_s}$ recover the classical idea of a tensor field as a collection of functions transforming in a certain way.

8 Differential forms.

In vector calculus in $\mathbb{R}^3$ extensive use is made of the idea of vector fields and the differential operators grad, div and curl. Differential forms and their associated exterior derivative are the generalisations to higher dimensions, and manifolds of these ideas.

8.1 The exterior algebra of a vector space.

If V is a vector space we define a k-linear map to be a map

$$\omega \colon V \times \cdots \times V \to \mathbb{R},$$

where there are k copies of V, which is linear in each factor. That is

$$\begin{aligned}\omega(v_1,\dots,v_{i-1},\alpha v+\beta w,v_{i+1},v_k) &= \alpha\omega(v_1,\dots,v_{i-1},v,v_{i+1},v_k)\\ &\quad +\beta\omega(v_1,\dots,v_{i-1},w,v_{i+1},v_k).\end{aligned}$$

We define a k-linear map ω to be totally antisymmetric if

$$\omega(v_1,\dots,v_i,v_{i+1},\dots,v_k) = -\omega(v_1,\dots,v_{i+1},v_i,\dots,v_k)$$

for all vectors $v_1,\dots,v_k$ and all i. Note that it follows that

$$\omega(v_1,\dots,v,v,\dots,v_k) = 0$$

and if $\pi \in S_k$ is a permutation of k letters then

$$\omega(v_1,v_2,\dots,v_k) = \operatorname{sgn}(\pi)\omega(v_{\pi(1)},v_{\pi(2)},\dots,v_{\pi(k)})$$

where $\operatorname{sgn}(\pi)$ is the sign of the permutation π. We denote the vector space of all k-linear, totally antisymmetric maps by $\Lambda^k(V^*)$. and call them k-forms. If $k=1$ the $\Lambda^1(V^*)$ is just V^* the space of all linear functions on V and if $k=0$ we make the convention that $\Lambda^0(V^*) = \mathbb{R}$. We need to collect some results on the linear algebra of these spaces.

Assume that V has dimension n and that $v_1,\dots,v_n$ is a basis of V. Let ω be a k form. Then if $w_1,\dots,w_k$ are arbitrary vectors and we expand them in the basis as

$$w_i = \sum_{j=1}^n w_{ij} v_j.$$

then we have

$$\omega(w_1, \ldots, w_k) = \sum_{j_1,\ldots,j_k=1}^{n} w_{1j_1} w_{2j_2} \ldots w_{kj_k} \omega(v_{j_1}, \ldots, v_{j_k})$$

so that it follows that ω is completely determined by its values on basis vectors. In particular if $k > n$ then $\Lambda^k(V^*) = 0$.

If α^1 and α^2 are two linear maps in V^* then we define an element $\alpha^1 \wedge \alpha^2$, called the wedge product of α^1 and α^2, in $\Lambda^2(V^*)$ by

$$\alpha^1 \wedge \alpha^2(v_1, v_2) = \alpha^1(v_1)\alpha^2(v_2) - \alpha^1(v_2)\alpha^2(v_1).$$

More generally if $\omega \in \Lambda^p(V^*)$ and $\rho \in \Lambda^q(V^*)$ we define $\omega \wedge \rho \in \Lambda^{p+q}(V^*)$ by

$$\begin{aligned}(\omega \wedge \rho)&(w_1, \ldots, w_{p+q}) \\ &= \frac{1}{p!q!} \sum_{\pi \in S_{p+q}} \operatorname{sgn}(\pi)\omega(w_{\pi(1)}, \ldots, w_{\pi(p)})\rho(w_{\pi(p+1)}, \ldots, w_{\pi(p+q)}).\end{aligned}$$

Assume that $\dim(V) = n$. Then we leave as an exercise the following proposition.

Proposition 8.1. *The direct sum*

$$\Lambda(V^*) = \bigoplus_{k=1}^{n} \Lambda^k(V^*)$$

with the wedge product is an associative algebra.

We call $\Lambda(V^*)$ the exterior algebra of V^*. We call an element $\omega \in \Lambda^k(V^*)$ an element of degree k. Because of associativity we can repeatedly wedge and disregard brackets. In particular we can define the wedge product of m elements in V^* and we leave it as an exercise to show that

$$\begin{aligned}\alpha^1 \wedge \alpha^2 \wedge \cdots \wedge \alpha^m(v_1, \ldots, v_m) &= \sum_{\pi \in S_m} \operatorname{sgn}(\pi)\alpha^1(v_{\pi(1)})\alpha^2(v_{\pi(2)}) \ldots \alpha^m(v_{\pi(m)}) \\ &= \det(\alpha^i(v_j))\end{aligned}$$

Notice that

$$\alpha^1 \wedge \cdots \wedge \alpha^i \wedge \alpha^{i+1} \wedge \cdots \wedge \alpha^m = -\alpha^1 \wedge \cdots \wedge \alpha^{i+1} \wedge \alpha^i \wedge \cdots \wedge \alpha^m$$

and that

$$\alpha^1 \wedge \cdots \wedge \alpha \wedge \alpha \wedge \cdots \wedge \alpha^m = 0.$$

Still assuming that V is n dimensional choose a basis $v_1, \ldots, v_n$ of V. Define the dual basis of V^*, $\alpha^1, \ldots, \alpha^n$, by

$$\alpha^i(v_j) = \delta^i_j$$

for all i and j. We want to define a basis of $\Lambda^k(V^*)$. Define elements of $\Lambda^k(V)$ by choosing k numbers $i_1, \ldots, i_k$ between 1 and n and considering

$$\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}.$$

As we are trying to form a basis we may as well keep the i_j distinct and ordered $1 \le i_1 < \cdots < i_k \le n$. We show first that these elements span $\Lambda^k(V^*)$. Let ω be an element of $\Lambda^k(V^*)$. Notice that

$$\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}(v_{j_1}, \ldots, v_{j_k})$$

equals zero unless there is a permutation π such that $j_l = i_{\pi(l)}$ for all l and equals $\operatorname{sgn}(\pi)$ if there is such a permutation. Consider vectors $w_1, \ldots, w_k$ and expand them in the basis as

$$w_i = \sum_j w_{ij} v_j.$$

Then we have

$$\omega(w_1,\ldots,w_k) = \sum_{j_1,\ldots,j_k} w_{1j_1}w_{2j_2}\ldots w_{kj_k}\omega(v_{j_1},\ldots,v_{j_k})$$

so that it follows that ω is completely determined by its values on basis vectors. For any ordered k-tuple $1 \le i_1 < \cdots < i_k \le n$ define

$$\omega_{i_1\ldots i_k} = \omega(v_{i_1},\ldots,v_{i_k})$$

and consider

$$\tilde{\omega} = \sum_{1\le i_1<\cdots<i_k\le n} \omega_{i_1\ldots i_k}\alpha^{i_1}\wedge\cdots\wedge\alpha^{i_k}.$$

We show that $\omega = \tilde{\omega}$. It suffices to apply both sides to vectors $(v_{i_1},\ldots,v_{i_k})$ for any $1 \le i_1 < \cdots < i_k \le n$ and show that they are equal but that is clear from previous discussions. So $\Lambda^k(V^*)$ is spanned by the basis vectors $\alpha^{i_1}\wedge\cdots\wedge\alpha^{i_k}$. We have

Proposition 8.2. *The vectors $\alpha^{i_1}\wedge\cdots\wedge\alpha^{i_k}$ where $1 \le i_1 < \cdots < i_k \le n$ are a basis for $\Lambda^k(V^*)$.*

Proof. We have already seen that these k-forms span. It suffices to show that they are linearly independent.

We show by induction on m that if we fix any subset B_m consisting of exactly m of the α^j then the k-forms $\alpha^{i_1}\wedge\cdots\wedge\alpha^{i_k}$ where $1 \le i_1 < \cdots < i_k \le n$, and all the α^j are chosen only from B_m, are linearly independent. If $m = 1$ the result is clear. More generally assume the result for m and that we have a linear relation amongst some of the $\alpha^{i_1}\wedge\cdots\wedge\alpha^{i_k}$ where each of the α^j are chosen from some set B_{m+1} of size $m+1$. There must be some $\alpha^i \in B_{m+1}$ which occurs in some but not all of the k-forms $\alpha^{i_1}\wedge\cdots\wedge\alpha^{i_k}$. Otherwise we must have $k = m+1$ and there is only one term in the linear expression which is not possible. So now wedge the whole expression with α^i. The terms containing α^i disappear and we obtain a relation between the $k+1$ forms constructed from vectors in $B_{m+1} - \{\alpha^i\}$ which is a contradiction. □

It is sometimes useful to sum over all k-tuples $i_1,\ldots,i_k$ not just ordered ones. We can do this — an keep the uniqueness of the coefficients $\omega_{i_1\ldots i_k}$ — if we demand that they be antisymmetric. That is

$$\omega_{j_1\ldots j_i j_{i+1}\ldots j_k} = -\omega_{j_1\ldots j_{i+1}j_i\ldots j_k}.$$

Then we have

$$\begin{aligned}\omega &= \textstyle\sum_{1\le i_1<\cdots<i_k\le n}\omega_{i_1\ldots i_k}\alpha^{i_1}\wedge\cdots\wedge\alpha^{i_k}\\ &= \textstyle\sum_{1\le i_1,\ldots,i_k\le n}\frac{1}{k!}\omega_{i_1\ldots i_k}\alpha^{i_1}\wedge\cdots\wedge\alpha^{i_k}.\end{aligned}$$

We will need one last piece of linear algebra called *contraction*. Let $\omega \in \Lambda^k(V)$ and $v \in V$. Then we define a $k-1$ form $\iota_v\omega$, the contraction of ω and v by

$$\iota_v(\omega)(v_1,\ldots,v_{k-1}) = \omega(v_1,\ldots,v_{k-1},v)$$

where $v_1,\ldots,v_{k-1}$ are any $k-1$ elements of V.

Example 8.1. Consider the vector space $\mathbb{R}^3$. Then we know that zero forms and one forms are just real numbers and linear maps respectively. Notice that in the case of $\mathbb{R}^3$ we can identify any linear map v with the vector $v = (v^1, v^2, v^3)$ where

$$v(x) = \sum_{i=1}^{3} v^i x^i.$$

Let α^i be the basis of linear functions defined by $\alpha^i(x) = x^i$. We have seen that every two form ω on $\mathbb{R}^3$ has the form

$$\omega = \omega_1\alpha^2\wedge\alpha^3 + \omega_2\alpha^3\wedge\alpha^1 + \omega_3\alpha^1\wedge\alpha^2.$$

Every three-form μ takes the form

$$\mu = a\alpha^1 \wedge \alpha^2 \wedge \alpha^3.$$

It follows that in $\mathbb{R}^3$ we can identify three-forms with real numbers by identifying μ with a and we can identify two-forms with vectors by identifying ω with $(\omega_1, \omega_2, \omega_3)$.

It is easy to check that with these identifications the wedge product of two vectors v and w is identified with the vector $v \times w$. In other words wedge product corresponds to cross product.

8.2 Differential forms and the exterior derivative.

We can now apply the constructions of the previous section to the tangent space to a manifold. We define a k-form on the tangent space at $x \in M$ to be an element of

$$\Lambda^k T_x^* M.$$

We want to define k-form 'fields' in the same way we define vector fields except that we do not call them k-form fields we call the differentiable k-forms or sometimes just k-forms. Choose co-ordinates (U, ψ) on M. Then $\omega(x)$ in $\Lambda^k(T_x^* M)$ can be be written as

$$\omega(x) = \sum_{i_1,\dots,i_k} \frac{1}{k!}\omega_{i_1,\dots,i_k} d\psi^{i_1} \wedge \cdots \wedge d\psi^{i_k}$$

at each $x \in U$. Hence we have defined a function

$$\omega_{i_1,\dots,i_k} \colon U \to \mathbb{R}$$

for each set of k indices. We call these functions the *components* of ω with respect to the co-ordinate chart. The components satisfy the anti-symmetry conditions in the previous section. We can also define the $\omega_{i_1,\dots,i_k}$ as

$$\omega_{i_1,\dots,i_k} = \omega(\frac{\partial}{\partial\psi^{i_1}}, \dots, \frac{\partial}{\partial\psi^{i_k}}).$$

We define a smooth differential form by

Definition 8.3 (Differential form.)**.** A differential form ω is smooth if its components with respect to a collection of co-ordinate charts whose domains cover M are smooth.

We have the usual Lemma

Lemma 8.4. *If a differential form is smooth then its components with respect to any co-ordinate chart are smooth.*

We denote by $\Omega^k(M)$ the set of all smooth differentiable k forms on M. Notice that $\Omega^0(M)$ is just $C^\infty(M)$ the space of all smooth functions on M.

Using the equation

$$\omega_{i_1,\dots,i_k} = \omega(\frac{\partial}{\partial\psi^{i_1}}, \dots, \frac{\partial}{\partial\psi^{i_k}}).$$

for the components of the differential form we can calculate the way the components change if we use another co-ordinate chart (V, χ). We have

$$\frac{\partial}{\partial\psi^i} = \sum_{a=1}^{n} \frac{\partial\chi^a}{\partial\psi^i}\frac{\partial}{\partial\chi^a}$$

and substituting this into the formula gives

$$\omega_{i_1,\dots,i_k} = \sum_{a_1,\dots,a_k=1}^{n} (\frac{\partial\chi^{a_1}}{\partial\psi^{i_1}} \cdots \frac{\partial\chi^{a_n}}{\partial\psi^{i_n}})\omega_{a_1,\dots,a_k}.$$

The usual derivative on functions defines a linear differential operator

$$d\colon \Omega^0(M) \to \Omega^1(M).$$

As well as being linear d satisfies the Leibniz rule:

$$d(fg) = fdg + (df)g.$$

We want to prove

Proposition 8.5. *If the dimension of M is n then there are unique linear maps*

$$d\colon \Omega^p(M) \to \Omega^{p+1}(M)$$

for all $p = 0, \ldots, n-1$ satisfying:

1. *If $p = 0$ d is the usual derivative,*
2. *$d^2 = 0$, and*
3. *$d(\omega \wedge \rho) = (d\omega) \wedge \rho + (-1)^p \omega \wedge (d\rho)$ where $\omega \in \Omega^p(M)$ and $\rho \in \Omega^q(M)$.*

Proof. We define d recursively. We have the ordinary definition of d if $p = 0$. We assume that we have it defined for all $p < k$ and that the conditions (i), (ii) and (iii) hold when ever they make sense. Consider a k form ω. Let (U, ψ) be a co-ordinate chart and let

$$\omega = \sum_{i_1 \ldots i_k} \frac{1}{k!} \omega_{i_1 \ldots i_k} d\psi^{i_1} \wedge \cdots \wedge d\psi^{i_k}.$$

Then define

$$\omega_i = \sum_{i_2 \ldots i_k} \frac{1}{k!} \omega_{i i_2 \ldots i_k} d\psi^{i_2} \wedge \cdots \wedge d\psi^{i_k}$$

so that

$$\omega = \sum_{i=1}^{k} \omega_i \wedge d\psi^i.$$

Notice that ω_i is uniquely defined by this equation because

$$\omega_i = \iota_{\frac{\partial}{\partial \psi^i}} \omega.$$

Consider now another choice of co-ordinates (V, χ). We have

$$\omega = \sum_{a=1}^{k} \omega_a \wedge d\chi^a$$

where

$$\omega_a = \iota_{\frac{\partial}{\partial \chi^a}} \omega.$$

It is easy to check that on $U \cap V$ we have

$$\omega_a = \sum_{a=1}^{n} \frac{\partial \psi^i}{\partial \chi^a} \omega_i.$$

Then if the proposition is to be true we must have

$$\begin{aligned} d\omega &= d(\textstyle\sum_{i=1}^{k} \omega_i \wedge d\psi^i) \\ &= \textstyle\sum_{i=1}^{k} d\omega_i \wedge d\psi^i \end{aligned}$$

This defines a differential $k+1$ form on the open set U. On the open set V it is defined by

$$\sum_{a=1}^{k} d\omega_a \wedge d\chi^a$$

and we need to check that these two agree. We have

$$\begin{aligned} d\omega_a =& d(\sum_{i=1}^{n} \frac{\partial \psi^i}{\partial \chi^a} \omega_i) \\ =& \sum_{b,i=1}^{n} \frac{\partial^2 \psi^i}{\partial \chi^b \partial \chi^a} d\chi^b \wedge \omega_i + \sum_{i=1}^{n} \frac{\partial \psi^i}{\partial \chi^a} d\omega_i. \end{aligned}$$

Hence

$$\sum_{a=1}^{n} d\omega_a \wedge d\chi^a = \sum_{i,a,b=1}^{n} \frac{\partial^2 \psi^i}{\partial \chi^b \partial \chi^a} d\chi^b \wedge \omega_i \wedge d\chi^a + \sum_{i,a=1}^{n} \frac{\partial \psi^i}{\partial \chi^a} d\omega_i \wedge d\chi^a.$$

The first term vanishes because the partial derivative is symmetric in a and b and the wedge product is anti-symmetric. Hence we have

$$\sum_{a=1}^{n} d\omega_a \wedge d\chi^a = \sum_{i=1}^{n} d\omega_i \wedge d\chi^i$$

as required. Clearly condition (i) is still true. For (ii) note that $dd\omega = d(\sum_{i=1}^{n} d\beta_i \wedge d\psi^i) = \sum_{i=1}^{n} dd\beta_i \wedge d\psi^i = 0$. For the final condition let $\rho = \sum_{i=1}^{n} \rho_i \wedge d\psi^i$. Assume that ρ has degree q. Then

$$\omega \wedge \rho = \sum_{i}^{n} \left(\sum_{j=1}^{n} (-1)^q \omega_i \wedge \rho_j \wedge \psi^j \right) \wedge \psi^i$$

so that

$$d(\omega \wedge \rho) = \sum_{i}^{n} \left(\sum_{j=1}^{n} (-1)^q d(\omega_i \wedge \rho_j \wedge d\psi^j) \right) \wedge d\psi^i.$$

Then applying the result for degrees lower than k we have

$$d(\omega_i \wedge \rho_j \wedge d\psi^j) = d(\omega_i) \wedge \rho_j \wedge d\psi^j + (-1)^{p-1} \omega_i \wedge d\rho_j \wedge d\psi^j.$$

Putting this altogether we have

$$\begin{aligned} d(\omega \wedge \rho) =& \sum_{j=1}^{n} \left[(-1)^q d(\omega_i) \wedge \rho_j \wedge d\psi^j + (-1)^{p+q-1} \omega_i \wedge d\rho_j \wedge d\psi^j \right] \wedge d\psi^i \\ =& \sum_{j=1}^{n} d(\omega_i) \wedge d\psi^i \wedge \rho_j \wedge d\psi^j + (-1)^q \omega_i \wedge d\psi^i \wedge d\rho_j \wedge d\psi^j \\ =& d(\omega) \wedge \rho + (-1)^q \omega \wedge d\rho. \end{aligned}$$

as required. □

Note that this proposition implies that if

$$\omega = \sum_{i_1,\ldots,i_k} \frac{1}{k!} \omega_{i_1,\ldots,i_k} d\psi^{i_1} \wedge \cdots \wedge d\psi^{i_k}.$$

then we have

$$d\omega = \sum_{i_1,\ldots,i_k} \frac{1}{k!} d\omega_{i_1,\ldots,i_k} \wedge d\psi^{i_1} \wedge \cdots \wedge d\psi^{i_k}$$

or

$$d\omega = \sum_{i_0,i_1,\ldots,i_k} \frac{1}{k!} \frac{\omega_{i_1,\ldots,i_k}}{\partial \psi^{i_0}} d\psi^{i_0} \wedge d\psi^{i_1} \wedge \cdots \wedge d\psi^{i_k}.$$

It is a useful exercise to reprove this lemma using this definition of the exterior derivative.

Example 8.2. Recall from 8.1 the way in which we identified one-forms and two-forms on $\mathbb{R}^3$ with vectors. It follows that differentiable one and two forms on $\mathbb{R}^3$ can be identified with vector-fields. Similarly zero and three forms are functions. With these identifications it is straightforward to check that the exterior derivative of zero, one and two forms corresponds to the classical differential operators grad, curl and div.

8.3 Pulling back differential forms

We have seen that if $f\colon M \to N$ is a smooth map then it has a derivative or tangent map $T_x(f)$ that acts on tangent vectors in T_xM by sending them to $T_{f(x)}N$. Moreover $T_x(f)$ is linear. Recall that if $X\colon V \to W$ is a linear map between vector spaces then it has an adjoint or dual $X^*\colon W^* \to V^*$ defined by

$$X^*(\xi)(v) = \xi(X(v))$$

where $\xi \in W^*$ and $v \in V$. Notice that X^* goes in the opposite direction to X. So we have a linear map called the cotangent map

$$T_x^*(f)\colon T_{f(x)}^*N \to T_x^*M$$

which is just the adjoint of the tangent map. It is defined by

$$T_x^*(f)(\omega)(X) = \omega(T_x(f)(X)).$$

This action defines a map on differential forms called the pull-back by f and denoted f^*. if $\omega \in \Omega^k(N)$ then we define $f^*(\omega) \in \Omega^k(M)$ by

$$f^*(\omega)(x)(X_1, \dots, X_k) = \omega(f(x))(T_x(f)(X_1), \dots, T_x(f)(X_k))$$

for any $X_1, \dots, X_k$ in T_xM.

Notice that if ϕ is a zero form or function on N then $f^{-1}(\phi) = \phi \circ f$. The pull back map

$$f^*\colon \Omega^q(N) \to \Omega^q(M).$$

satisfies the following proposition.

Proposition 8.6. *If $f\colon M \to N$ is a smooth map and ω and μ is a differential form on N then:*

1. *$df^*(\omega) = f^*(d\omega)$, and*
2. *$f^*(\omega \wedge \mu) = f^*(\omega) \wedge f^*(\mu)$.*

Proof. Exercise. □

8.4 Integration of differential forms

If M is a manifold a partition of unity is a collection of smooth non-negative functions $\{\rho_\alpha\}_{\alpha\in I}$ such that every $x \in M$ has neighbourhood on which only a finite number of the ρ are non-vanishing and such that $\sum_{\alpha\in I} \rho_\alpha = 1$.

Recall that if $f\colon M \to \mathbb{R}$ is smooth function then we define $\operatorname{supp}(f)$ to be the closure of the set on which f is non-zero. There are two basic existence results on a paracompact, Hausdorff manifold.

1. If $\{U_\alpha\}_{\alpha\in I}$ is an open cover of M there is a partition of unity $\{\rho_\alpha\}_{\alpha\in I}$ with $\operatorname{supp}(\rho_\alpha) \in U_\alpha$. Such a partition of unity is called subordinate to the cover.
2. If $\{U_\alpha\}_{\alpha\in I}$ is an open cover of M there is a partition of unity $\{\rho_\alpha\}_{\beta\in J}$, with a possibly different indexing set J with each $\operatorname{supp}(\rho_\beta)$ in some U_α.

Let $U \subset \mathbb{R}^n$ and $\psi\colon \mathbb{R}^n \to \mathbb{R}^n$ be a diffeomorphism. Then it is well known that if $f\colon \psi(U) \to \mathbb{R}$ is a function then

$$\int_U f \circ \psi |\det\left(\frac{\partial \psi^i}{\partial x^j}\right)| dx^1 \dots dx^n = \int_{\psi(U)} f dx^1 \dots dx^n.$$

In this formula we regard $dx^1 \dots dx^n$ as the symbol for Lebesgue measure. However it is very suggestive of the notation for differential forms developed in the previous section.

If ω is a differential n form on $V = \psi(U)$ then we can write it as

$$\omega(x) = f(x) dx^1 \wedge \cdots \wedge dx^n.$$

If we pull it back with the diffeomorphism ϕ then, as we seen before,

$$\psi^*(\omega) = f(x) \det\left(\frac{\partial \psi^i}{\partial x^j}\right) dx^1 \wedge \cdots \wedge dx^n.$$

So differential n forms transform by the determinant of the jacobian of the diffeomorphism and Lebesgue measure transforms by the absolute value of the determinant of the jacobian of the diffeomorphism. We define the integral of a differential n form by

$$\int_V \omega = \int_V f(x) dx^1 \dots dx^n$$

when $\omega = f(x) dx^1 \wedge \cdots \wedge dx^n$. Alternatively we can write this as

$$\int_V \omega = \int_V \omega(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}) dx^1 \dots dx^n.$$

Call a diffeomorphism $\psi\colon U \to V$ *orientation preserving* if

$$\det\left(\frac{\partial \psi^i}{\partial x^j}\right)(x) > 0$$

for all $x \in U$. Then we have

Proposition 8.7. *If $\psi\colon U \to \psi(U)$ is an orientation preserving diffeomorphism and ω is a differential n form on $\psi(U)$ then*

$$\int_{\psi(U)} \omega = \int_U \psi^*(\omega).$$

We can use this proposition to define the integral of differential forms on a manifold. Let $\{(U_\alpha, \psi_\alpha)\}_{\alpha \in I}$ be a covering of M by co-ordinate charts. Choose a partition of unity ϕ_α subordinate to U_α. Then if ω is a differential n form we can write

$$\omega = \sum_\alpha \phi_\alpha \omega$$

where the support of $\phi_\alpha \omega$ is in U_α. First we define the integral of each of the forms $\phi_\alpha \omega$

$$\int_M \phi_\alpha \omega = \int_{\psi_\alpha(U_\alpha)} (\psi^{-1})^*(\phi_\alpha \omega).$$

Then we define the integral of ω to be

$$\int_M \omega = \sum_{\alpha \in I} \int_{\psi_\alpha(U_\alpha)} (\psi^{-1})^*(\phi_\alpha \omega).$$

We have to show that this is independent of all the choices we have made. So let us take another open cover $\{(V_\beta, \chi_\beta)\}_{\beta\in J}$ with partition of unity ρ_β. Then we have

$$\begin{aligned}\sum_{\alpha\in I}\int_{\psi_\alpha(U_\alpha)}(\psi^{-1})^*(\phi_\alpha\omega) &= \sum_{\alpha\in I}\int_{\psi_\alpha(U_\alpha)}(\psi^{-1})^*(\sum_{\beta\in J}\rho_b\phi_\alpha\omega)\\ &= \sum_{\alpha\in I}\int_{\psi_\alpha(U_\alpha)}(\psi^{-1})^*((\sum_{\beta\in J}\rho_b)\phi_\alpha\omega)\\ &= \sum_{\alpha\in I}\sum_{\beta\in J}\int_{\psi_\alpha(U_\alpha)}(\psi^{-1})^*(\rho_b\phi_\alpha\omega).\end{aligned}$$

The differential forms $\rho_b\phi_\alpha\omega$ have support in $U_\alpha \cap V_\beta$ so we have

$$\begin{aligned}\int_{\psi_\alpha(U_\alpha)}(\psi^{-1})^*(\rho_b\phi_\alpha\omega) &= \int_{\psi_\alpha(U_\alpha\cap V_\beta)}(\psi^{-1})^*(\rho_b\phi_\alpha\omega)\\ &= \int_{\psi_\alpha(U_\alpha\cap V_\beta)}(\psi_\alpha^{-1})^*(\rho_b\phi_\alpha\omega)\end{aligned}$$

If the diffeomorphism

$$\chi_\beta \circ \psi_\alpha^{-1}{}_{|\psi(U_\alpha\cap V_b)}$$

is orientation preserving then we have

$$\int_{\psi_\alpha(U_\alpha\cap V_\beta)}(\psi_\alpha^{-1})^*(\rho_b\phi_\alpha\omega) = \int_{\chi_\beta(U_\alpha\cap V_\beta)}(\chi_\beta^{-1})^*(\rho_b\phi_\alpha\omega) \qquad = \int_{\chi_\beta(U_\beta)}(\chi_\beta^{-1})^*(\rho_b\omega).$$

So we can complete the calculation above and have

$$\sum_{\alpha\in I}\int_{\psi_\alpha(U_\alpha)}(\psi^{-1})^*(\phi_\alpha\omega) = \sum_{\beta\in J}\int_{\chi_\beta(U_\beta)}(\chi^{-1})^*(\rho_\beta\omega).$$

All this calculation rests on the fact that

$$\chi_\beta \circ \psi_\alpha^{-1}{}_{|\psi(U_\alpha\cap V_b)}$$

is an orientation preserving diffeomorphism. In general this will not be the case. We have to introduce the notion of an oriented manifold and and oriented co-ordinate chart. Before we can do that we need to discuss orientations on a vector space.

8.5 Orientation.

Let V be a real vector space of dimension n. Then define $\det(V) = \Lambda^n(V)$. This is a real, one dimensional vector space. So the set $\det(V) = \{0\}$ is *disconnected*. An orientation of the vector space V is a choice of one of these connected components. If X is an invertible linear map from V to V then it induces a linear map from $\det(V) \to \det(V)$ which is therefore multiplication by a complex number. This number is just $\det(X)$ the determinant of X If M is a manifold of dimension n then the same applies to M; $\det(T_xM) - \{0\}$ is a disconnected set. We define

Definition 8.8. A manifold is orientable if there is a non-vanishing n-form on M. Otherwise it is called non-orientable.

If η and ζ are two non-vanishing n forms then $\eta = f\zeta$ for some function f which is either strictly negative or strictly positive. Hence the set of non-vanishing n forms divides into two sets. We have

Definition 8.9 (Orientation)**.** An orientation on M is a choice of one of these two sets.

An orientation defines an orientation on each tangent space T_xM. We call an n form positive if it coincides with the chosen orientation negative otherwise. We say a chart (U, ψ) is positive or oriented if $d\psi^1 \wedge \cdots \wedge d\psi^n$ is positive. Note that if a chart is not positive we can make it so by changing the sign of one co-ordinate function so oriented charts exits. If we chose two oriented charts then we have that

$$\chi \circ \psi^{-1}_{|\psi(U\cap V)}$$

is an oriented diffeomorphism. The converse is also true.

Proposition 8.10. *Assume we have a covering of M by co-ordinate charts* $\{(U_\alpha, \psi_\alpha)\}_{\alpha\in I}$ *such that for any two* (U_α, ψ_α) *and* (U_β, ψ_β) *the diffeomorphism*

$$\psi_\beta \circ \psi_\alpha^{-1}{}_{I\psi_\alpha(U_\alpha\cap U_\beta)}$$

is orientation preserving. Then there is an orientation of M which makes each all these charts oriented.

Proof. Notice that the fact that

$$\psi_\beta \circ \psi_\alpha^{-1}{}_{|\psi_\alpha(U_\alpha\cap U_\beta)}$$

is an oriented diffeomorphism means that if $x \in U_\alpha \cap U_\beta$ then

$$d\psi_\alpha^1 \wedge \ldots d\psi_\alpha^n(x)$$

is a positive multiple of

$$d\psi_\beta^1 \wedge \ldots d\psi_\beta^n(x)$$

Hence if ϕ_α is a partition of unity then

$$\rho = \sum \phi_\alpha d\psi_\alpha^1 \wedge \ldots d\psi_\alpha^n(x)$$

is a non-vanishing n form. Clearly this defines the required orientation. □

8.6 Integration again

We now have the required result that we can integrate differential forms of degree k over a k-dimensional oriented manifold.

9 Stokes theorem.

Recall the Fundamental Theorem of Calculus: If f is a differentiable function then

$$f(b) - f(a) = \int_b^a \frac{df}{dt}(x)dx.$$

In the language we have developed in the previous section this can be written as

$$f(b) - f(a) = \int_{[a,b]} df$$

where we orient the 1-dimensional manifold $[a, b]$ in the positive direction. We want to prove a more general result that will include Stokes theorem, Green's theorem, Gauss' theorem and the Divergence theorem. If M is an oriented manifold of dimension n with boundary ∂M and ω is an $n - 1$ form then Stoke's theorem says that

$$\int_M d\omega = \int_{\partial M} \omega.$$

Before we prove this result we need to make sense of the idea of a manifold with boundary.

9.1 Manifolds with boundary.

We denote by $\mathbb{R}^n_+$ the half-space

$$\mathbb{R}^n_+ = \{(x^1, \ldots, x^n) \mid x^1 > 0\}.$$

We define the boundary of $\mathbb{R}^n_+$ to be

$$\partial\mathbb{R}^n_+ = \{(x^1, \ldots, x^n) \mid x^1 = 0\}.$$

and we identify it with $\mathbb{R}^{n-1}$. Recall that a set $U \subset \mathbb{R}^n_+$ is open if it is of the form $U = V \cap \mathbb{R}^n_+$ where V is open in $\mathbb{R}^n$. If U is open in $\mathbb{R}^n_+$ we say that $f \colon U \to \mathbb{R}$ is smooth if there is an open set $V \subset \mathbb{R}^n$ with $U = V \cap \mathbb{R}^n_+$ and a smooth map $F \colon V \to \mathbb{R}$ such that $F(x) = f(x)$ for all $x \in U$. If $V \subset \mathbb{R}^n$ we define $\partial V = V \cap \partial\mathbb{R}^n_+$.

Let M be a set with a subset denoted by ∂M that we call the boundary of M. We say that (U, ψ) is a co-ordinate chart on M if it is a co-ordinate chart as defined before but in addition $\psi(\partial U) \subset \partial\psi(U)$, $\psi(\partial U)$ is open in $\partial\psi(U)$,and $\psi_{|\partial U} \colon \partial U \to \partial\psi(U)$ is a bijection. We define compatibility of charts in the usual way but with the extended notion of smoothness above. Once we have this we can define the idea of an atlas and the notion of a manifold M with boundary ∂M. Notice that if we discard the boundary points ∂M we immediately see that $M - \partial M$ is a manifold. Similarly ∂M is a manifold. We can extend everything we have done so far to the case of manifolds with boundary.

9.2 Stokes theorem.

Let M be a manifold of dimension n with boundary ∂M and let ω be an $n-1$ form on M with compact support. We want to prove

Theorem 9.1 (Stoke's theorem). *Let M be an oriented manifold with boundary of dimension n and let ω be a differential form of degree $n-1$ with compact support then*

$$\int_M d\omega = \int_{\partial M} \omega.$$

Proof. We cover M with a covering by co-ordinate charts (U_α, ψ_α) and choose a partition of unity ϕ_α subordinate to this cover. Notice that because $\sum_\alpha \phi_\alpha = 1$ we have $\sum_\alpha d\phi_\alpha = 0$ and hence

$$\begin{aligned}
\int_M d\omega &= \sum_\alpha \int_M \phi_\alpha d\omega \\
&= \sum_\alpha \int_M d(\phi_\alpha \omega) \\
&= \sum_\alpha \int_{U_\alpha} d(\phi_\alpha \omega) \\
&= \sum_\alpha \int_{\psi_\alpha(U_\alpha)} (\psi^{-1})^* d(\phi_\alpha \omega) \\
&= \sum_\alpha \int_{\psi_\alpha(U_\alpha)} d((\psi^{-1})^*(\phi_\alpha \omega))
\end{aligned}$$

and

$$\begin{aligned}
\int_{\partial M} \omega &= \sum_\alpha \int_{\partial M} \phi_\alpha \omega \\
&= \sum_\alpha \int_{\partial U_\alpha} \phi_\alpha \omega \\
&= \sum_\alpha \int_{\partial\psi_\alpha(U_\alpha)} (\psi^{-1})^*(\phi_\alpha \omega).
\end{aligned}$$

So it suffices prove that

$$\int_{\psi_\alpha(U_\alpha)} d((\psi^{-1})^*(\phi_\alpha\omega)) = \int_{\partial\psi_\alpha(U_\alpha)} (\psi^{-1})^*(\phi_\alpha\omega)$$

or equivalently Stoke's theorem for differential forms with compact support on $\mathbb{R}^n_+$. Let us assume then that ω is a differential form on U, where U is of the form $U = \mathbb{R}^n_+ \cap V$ for V open in $\mathbb{R}^n$. As as ω is has compact support it has bounded support so there is an $R > 0$ such that if $|x^i| > R$ for all $i = 1, \dots, n$ then $\omega(x) = 0$. Write $x = (t, y)$ for $y \in \mathbb{R}^{n-1}$ and $\omega = f dt \wedge dy^1 \dots dy^{n-1}$ so we have

$$\begin{aligned}
\int_U d\omega &= \int_U \frac{\partial f}{\partial t} dt \wedge dy^1 \dots dy^n + \int_U \sum_{i=1}^{n-1} \frac{\partial f}{\partial x^i} dy^i \wedge dt \wedge dy^1 \dots dy^n. \\
&= \int_U \frac{\partial f}{\partial t} dt \wedge dy^1 \dots dy^n + \int_U \sum_{i=1}^{n-1} [f(t, y^1, \dots, y^{i-1}, R, y^{i+1}, \dots, y^{n-1}) \\
&\quad - \dots f(t, y^1, \dots, y^{i-1}, -R, y^{i+1}, \dots, y^{n-1})] dt \wedge dy^1 \dots dy^n. \\
&= \int_U \frac{\partial f}{\partial t} dt \wedge dy^1 \dots dy^n \\
&= \int_{\partial U} f(0, y^1, \dots, y^{n-1}) dy^1 \dots dy^n \\
&= \int_{\partial U} \omega
\end{aligned}$$

□

9.3 Exercises

Exercise 9.1. If X is a vector field and ω is a differential 1-form show that the differential 1-form defined by

$$L_X(\omega) = \sum_{i,j} (X^i \frac{\partial \omega_j}{\partial \theta^i} + \omega_i \frac{\partial X^i}{\partial \theta^j}) d\theta^j.$$

where

$$X = \sum_i X^i \frac{\partial}{\partial \theta^i} \quad \text{and} \quad \omega = \sum_i \omega_i d\theta^i$$

is actually independent of the choices of co-ordinates. We call $L_X(\omega)$ the Lie derivative of ω by X.

Exercise 9.2. Let α and β be p and q forms, respectively on a manifold M. Show that

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta.$$

Exercise 9.3. Consider the circle $S^1 = \{(x, y) \mid x^2 + y^2 = 1\}$. This is a manifold of dimension 1. The circle has a co-ordinate chart (U, θ) where $U = S^1 - \{(1, 0)\}$ and $\theta \colon U \to (0, 2\pi)$ is defined implicitly by

$$(x, y) = (\cos(\theta(x, y)), \sin(\theta(x, y))).$$

That is θ is the usual angle co-ordinate in polar co-ordinates. Identify the tangent space to the circle at (x, y) with the line in $\mathbb{R}^2$ tangential to the circle at (x, y). Calculate a formula for the vector field $\partial/\partial\theta$ in terms of x and y and hence show that it extends from U to a vector field on all of S^1. Show that $d\theta$ also extends to a differential 1-form ω on all of the circle. Show that there is no function $f : S^1 \to \mathbb{R}$ such that $\omega = df$.

Exercise 9.4. Let $S^2 = \{x \in \mathbb{R}^3 \mid \|x\|^2 = 1\}$ be the two-sphere. Recall that the spherical co-ordinates (θ, ϕ) of the point (x, y, z) on the two-sphere are defined by requiring that:

$$\begin{aligned}
x &= \sin(\psi)\cos(\theta) \\
y &= \sin(\psi)\sin(\theta) \\
z &= \cos(\psi).
\end{aligned}$$

Find an open set $U \subset S^2$ for the domain of the spherical co-ordinates so that $\psi \in (0, \pi)$ and $\theta \in (0, 2\pi)$.

For any x in S^2 and $X, Y \in T_x S^2$ define a differential two-form ω on S^2 by $\omega_x(X, Y) = \langle x, X \times Y \rangle$, where $\langle\ ,\ \rangle$ is the usual inner-product on $\mathbb{R}^3$ and x is the cross-product of three vectors. By using suitable co-ordinates (spherical are good) calculate the integral of ω over S^2 and show that it is non-zero.

Exercise 9.5. Show that it is not possible to find a differential one-form μ on the two sphere such that $d\mu$ is the volume form ω defined in exercise (9.4).

Exercise 9.6. Consider the torus T^2 in $\mathbb{R}^3$ with co-ordinates (θ, ϕ) defined implicitly by

$$x = (b + a\sin(\phi))\cos(\theta), (b + a\sin(\phi))\sin(\theta), a\cos(\phi)).$$

Calculate $\partial/\partial\psi$ and $\partial/\partial\theta$. Calculate the (outward) unit normal $n(x)$ to the torus, this is the vector in $\mathbb{R}^3$ orthogonal to the tangent space to the torus at x. You will need to draw a picture or something to check it is the outward normal.

Define vol a two-form by $\mathrm{vol}(X, Y) = \langle n, X \times Y \rangle$ and calculate its integral over T^2 when we orient T^2 in such a way as to make vol positive.

Exercise 9.7. Recall the definition of $\mathbb{R}P_2$ the space all lines through the origin in $\mathbb{R}^2$ and its associated co-ordinate charts given in Exercise 2.4. Calculate the linear relationship between the basis of one forms $d\psi_i^1, d\psi_i^2$ and the basis of one forms $d\psi_j^1, d\psi_j^2$ for $i \neq j$. Hence calculate the relationship between $d\psi_i^1 \wedge d\psi_i^2$ and $d\psi_j^1 \wedge d\psi_j^2$. Show that $\mathbb{R}P_2$ is not orientable.

Exercise 9.8. Let $f \colon M \to N$ be a smooth map. If ω is a p-form on N show that $df^*(\omega) = f^* d\omega$.

10 Complex line bundles

10.1 Introduction

The mathematical motivation for studying vector bundles comes from the example of the tangent bundle TM of a manifold M. Recall that the tangent bundle is the union of all the tangent spaces $T_m M$ for every m in M. As such it is a collection of vector spaces, one for every point of M.

The physical motivation comes from the realisation that the fields in physics may not just be maps $\phi : M \to \mathbb{C}^N$ say, but may take values in *different* vector spaces at each point. Tensors do this for example. The argument for this is partly quantum mechanics because, if ϕ is a wave function on a space-time M say, then what we can know about are expectation values, that is things like:

$$\int_M \langle \phi(x), \phi(x) \rangle dx$$

and to define these all we need to know is that $\phi(x)$ takes its values in a one-dimensional complex vector space with Hermitian inner product. There is no reason for this to be the same one-dimensional Hermitian vector space here as on Alpha Centauri. Functions like ϕ, which are generalisations of complex valued functions, are called *sections* of vector bundles.

We will consider first the simplest theory of vector bundles where the vector space is a one-dimensional complex vector space - line bundles.

10.2 Definition of a line bundle and examples

The simplest example of a line bundle over a manifold M is the *trivial* bundle $\mathbb{C} \times M$. Here the vector space at each point m is $\mathbb{C} \times \{m\}$ which we regard as a copy of $\mathbb{C}$. The general definition uses this as a local model.

Definition 10.1. A complex line bundle over a manifold M is a manifold L and a smooth surjection $\pi \colon L \to M$ such that:

1. Each *fibre* $\pi^{-1}(m) = L_m$ is a a complex one-dimensional vector space.
2. Every $m \in M$ has an open neighbourhood $U \in M$ for which there is a diffeomeorphism $\varphi : \pi^{-1}(U) \to U \times \mathbb{C}$ such that $\varphi(L_m) \subset \{m\} \times \mathbb{C}$ for every m and that moreover the map $\varphi|_{L_m} : L_m \to \{m\} \times \mathbb{C}$ is a linear isomorphism.

Note 10.1. The second condition is called *local triviality* because it says that locally the line bundle looks like $\mathbb{C} \times M$. We leave it as an exercise to show that local triviality makes the map π a submersion (that is it has onto derivative) and the scalar multiplication and vector addition maps smooth. In the quantum mechanical example local triviality means that at least in some local region like the laboratory we can identify the Hermitian vector space where the wave function takes its values with $\mathbb{C}$.

Example 10.1. $\mathbb{C} \times M$ the trivial bundle

Example 10.2. Recall that if $u \in S^2$ then the tangent space at u to S^2 is identified with the set $T_uS^2 = \{v \in \mathbb{R}^3 \mid \langle v, u\rangle = 0\}$. We make this two dimensional real vector space a one dimensional complex vector space by defining $(\alpha + i\beta)v = \alpha.v + \beta.u \times v$. We leave it as an exercise for the reader to show that this does indeed make T_uS^2 into a complex vector space. What needs to be checked is that $[(\alpha + i\beta)\ (\delta + i\gamma)]v = (\alpha + i\beta)\ [(\delta + i\gamma)]v$ and because T_uS^2 is already a real vector space this follows if $i(iv) = -v$. Geometrically this follows from the fact that we have defined multiplication by i to mean rotation through $\pi/2$. We will prove local triviality in a moment.

Example 10.3. If Σ is any surface in $\mathbb{R}^3$ we can use the same construction as in (2). If $x \in \Sigma$ and $\hat{n}_x$ is the unit normal then $T_x\Sigma = \hat{n}_x^{\perp}$. We make this a complex space by defining $(\alpha + i\beta)v = \alpha v + \beta\hat{n}_x \times v$.

Example 10.4 (Hopf bundle). Define $\mathbb{C}P_1$ to be the set of all lines (through the origin) in $\mathbb{C}^2$. Denote the line through the vector $z = (z^0, z^1)$ by $[z] = [z^0, z^1]$. Note that $[\lambda z^0, \lambda z^1] = [z^0, z^1]$ for any non-zero complex number λ. Define two open sets U_i by

$$U_i = \{[z^0, z^1] \mid z^i \neq 0\}$$

and co-ordinates by $\psi_i \colon U_i \to \mathbb{C}$ by $\psi_0([z]) = z^1/z^0$ and $\psi_1([z]) = z^0/z^1$. As a manifold $\mathbb{C}P_1$ is diffeomorphic to S^2. An explicit diffeomorphism $S^2 \to \mathbb{C}P_1$ is given by $(x, y, z) \mapsto [x + iy, 1 - z]$.

We define a line bundle H over $\mathbb{C}P_1$ by $H \subset \mathbb{C}^2 \times \mathbb{C}P_1$ where

$$H = \{(w, [z]) \mid w = \lambda z \text{ for some } \lambda \in \mathbb{C}^\times\}.$$

We define a projection $\pi \colon H \to \mathbb{C}P_1$ by $\pi((w, [z])) = [z]$. The fibre $H_[z] = \pi^{-1}([z])$ is the set

$$\{(\lambda z, [z]) \mid \lambda \in \mathbb{C}^\times\}$$

which is naturally identified with the line through $[z]$. It thereby inherits a vector space structure given by

$$\alpha(w, [z]) + \beta(w', [z]) = (\alpha w + \beta w', [z]).$$

We shall prove later that this is locally trivial.

10.3 Isomorphism of line bundles

It is useful to say that two line bundles $L \to M, J \to M$ are isomorphic if there is a diffeomorphism map $\varphi : L \to J$ such that $\varphi(L_m) \subset J_m$ for every $m \in M$ and such that the induced map $\varphi|_{L_m} : L_m \to J_m$ is a linear isomorphism.

We define a line bundle L to be *trivial* if it is isomorphic to $M \times \mathbb{C}$ the trivial bundle. Any such isomorphism we call a trivialisation of L.

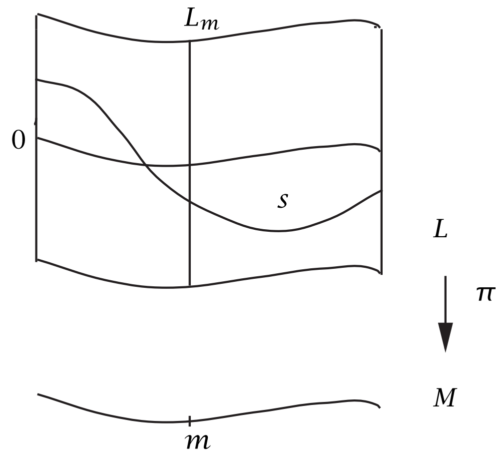

Figure 10.1: A line bundle.

10.4 Sections of line bundles

A section of a line bundle L is like a vector field. That is it is a map $\varphi : M \to L$ such that $\varphi(m) \in L_m$ for all m or more succinctly $\pi \circ \varphi = id_m$.

Example 10.5 (The trivial bundle.). $L = \mathbb{C} \times M$. Every section φ looks like $\varphi(x) = (f(x), x)$ for some function f.

Example 10.6 (The tangent bundle to S^2: TS^2). Sections are vector fields. Alternatively because each $T_xS^2 \subset \mathbb{R}^3$ we can think of a section s as a map $s\colon S^2 \to \mathbb{R}^3$ such that $\langle s(x), x\rangle = 0$ for all $x \in S^2$.

Example 10.7 (The Hopf bundle). By definition a section $s\colon \mathbb{C}P_1 \to H$ is a map

$$s\colon \mathbb{C}P_1 \to H \subset \mathbb{C}^2 \times \mathbb{C}P_1$$

which must have the form $[z] \mapsto ([z], w)$. For convenience we will write it as $s([z]) = ([z], s(z))$ where, for any $[z]$ $s\colon \mathbb{C}P_1 \to \mathbb{C}^2$ satisfies $s([z]) = \lambda z$ for some $\lambda \in \mathbb{C}^\times$.

The set of all sections, denoted by $\Gamma(M, L)$, is a vector space under pointwise addition and scalar multiplication. I like to think of a line bundle as looking like Figure 1.

Here O is the set of all zero vectors or the image of the *zero section*. The curve s is the image of a section and thus generalises the graph of a function.

We have the following result:

Proposition 10.2. *A line bundle L is trivial if and only if it has a nowhere vanishing section.*

Proof. Let $\varphi\colon L \to M \times \mathbb{C}$ be the trivialisation then $\varphi^{-1}(m, 1)$ is a nowhere vanishing section.

Conversely if s is a nowhere vanishing section then define a trivialisation $M \times \mathbb{C} \to L$ by $(m, \lambda) \mapsto \lambda s(m)$. This is an isomorphism. □

Note 10.2. . The condition of local triviality in the definition of a line bundle could be replaced by the existence of local nowhere vanishing sections. This shows that TS^2 is locally trivial as it clearly has *local* nowhere-vanishing vector fields. Recall however the so called 'hairy-ball theorem' from topology which tells us that S^2 has no global nowhere vanishing vector fields. Hence TS^2 is not trivial. We shall prove this result a number of times.

10.5 Transition functions and the clutching construction

Local triviality means that every property of a line bundle can be understood locally. This is like choosing co-ordinates for a manifold. Given $L \to M$ we cover M with open sets U_α on which there are nowhere vanishing sections s_α. If ξ is a global section of L then it satisfies $\xi|_{U_\alpha} = \xi_\alpha s_\alpha$ for some smooth $\xi_\alpha : U_\alpha \to \mathbb{C}$. The converse is also true. If we can find ξ_α such that $\xi_\alpha s_\alpha = \xi_\beta s_\beta$ for all α, β then they fit together to define a global section ξ with $\xi|_{U_\alpha} = \xi_\alpha s_\alpha$.

It is therefore useful to define $g_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathbb{C}^\times$ by $s_\alpha = g_{\alpha\beta} s_\beta$. Then a collection of functions ξ_α define a global section if on any intersection $U_\alpha \cap U_\beta$ we have $\xi_\beta = g_{\alpha\beta}\xi_\alpha$. The functions $g_{\alpha\beta}$ are called the *transition functions* of L. We shall see in a moment that they determine L completely. It is easy to show, from their definition, that the transition functions satisfy three conditions:

(1) $g_{\alpha\alpha} = 1$

(2) $g_{\alpha\beta} g_{\beta\alpha} = 1$

(3) $g_{\alpha\beta}\, g_{\beta\gamma}\, g_{\gamma\alpha} = 1$ on $U_\alpha \cap U_\beta \cap U_\gamma$

We leave it as an exercise to show that these are equivalent to the *cocycle condition*:

$$g_{\beta\gamma}\, g_{\alpha\gamma}^{-1}\, g_{\alpha\beta} = 1 \text{ on } U_\alpha \cap U_\beta \cap U_\gamma$$

Proposition 10.3. *Given an open cover $\{U_\alpha\}$ of M and functions $g_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathbb{C}^\times$ satisfying (1) (2) and (3) above we can find a line bundle $L \to M$ with transition functions the $g_{\alpha\beta}$.*

Proof. Consider the disjoint union $\tilde{M}$ of all the $\mathbb{C} \times U_\alpha$. We stick these together using the $g_{\alpha\beta}$. More precisely let I be the indexing set and define $\tilde{M}$ as the subset of $I \times M$ of pairs (α, m) such that $m \in U_\alpha$. Now consider $\mathbb{C} \times \tilde{M}$ whose elements are triples (λ, m, α) and define $(\lambda, m, \alpha) \sim (\mu, n, \beta)$ if $m = n$ and $g_{\alpha\beta}(m)\lambda = \mu$. We leave it as an exercise to show that $\sim$ is an equivalence relation. Indeed ((1) (2) (3) give reflexivity, symmetry and transitivity respectively.)

Denote equivalence classes by square brackets and define L to be the set of equivalence classes. Define addition by $[(\lambda, m, \alpha)] + [(\mu, m, \alpha)] = [(\lambda + \mu, m, \alpha)]$ and scalar multiplication by $z[(\lambda, m, \alpha)] = [(z\lambda, m, \alpha)]$. The projection map is $\pi([(\lambda, m, \alpha)]) = m$. We leave it as an exercise to show that these are all well-defined. Finally define $s_\alpha(m) = [(1, m, \alpha)]$. Then $s_\alpha\,(m) = [(1, m, \alpha)] = [(g_{\alpha\beta}\,(m), m, \beta)] = g_{\alpha\beta}\,(m) s_\beta\,(m)$ as required.

Finally we need to show that L can be made into a differentiable manifold in such a way that it is a line bundle and the s_α are smooth. Denote by W_α the preimage of U_α under the projection map. There is a bijection $\psi_\alpha \colon W_\alpha \to \mathbb{C} \times U_\alpha$ defined by $\psi_\alpha([\alpha, x, z]) = (z, x)$. This is a local trivialisation. If (V, ϕ) is a co-ordinate chart on $U_\alpha \times \mathbb{C}$ then we can define a chart on L by $(\phi_\alpha^{-1}(V), \phi_\alpha \circ \psi_\alpha)$. We leave it as an exercise to check that these charts define an atlas. This depends on the fact that $g_{\alpha\beta} \colon U_\alpha \cap U_\beta \to \mathbb{C}^\times$ is smooth. □

The construction we have used here is called *the clutching construction.* It follows from this proposition that the transition functions capture all the information contained in L. However they are by no means unique. Even if we leave the cover fixed we could replace each s_α by $h_\alpha s_\alpha$ where $h_\alpha : U_\alpha \to \mathbb{C}^\times$ and then $g_{\alpha\beta}$ becomes $h_\alpha g_{\alpha\beta} h_\beta^{-1}$. If we continued to try and understand this ambiguity and the dependence on the cover we would be forced to invent Cêch cohomology and show that that the isomorphism classes of complex line bundles on M are in bijective correspondence with the Cêch cohomology group $H^1(M, \mathbb{C}^\times)$. We refer the interested reader to [11, 8].

Example 10.8. The tangent bundle to the two-sphere. Cover the two sphere by open sets U_0 and U_1 corresponding to the upper and lower hemispheres but slightly overlapping on the equator. The intersection of U_0 and U_1 looks like an annulus. We can find non-vanishing vector fields s_0 and s_1 as in Figure 2.

If we undo the equator to a straightline and restrict s_0 and s_1 to that we obtain Figure 3.

If we solve the equation $s_0 = g_{01} s_1$ then we are finding out how much we have to rotate s_1 to get s_0 and hence defining the map $g_{01} \colon U_0 \cap U_1 \to \mathbb{C}^\times$ with values in the unit circle. Inspection of Figure 3 shows that as we go around the equator once s_0 rotates forwards once and s_1 rotates backwards once so that thought of as a point on the unit circle in $\mathbb{C}^\times$ g_{01} rotates around twice. In other words $g_{01} : U_0 \cap U_1 \to \mathbb{C}^\times$ has winding number 2. This two will be important latter.

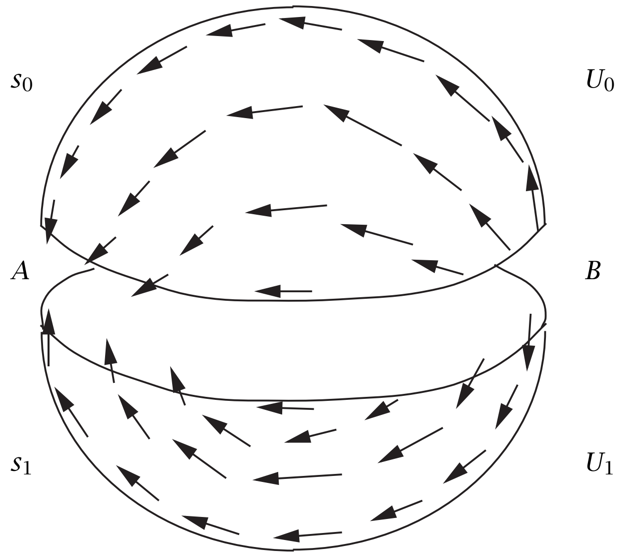

Figure 10.2: Vector fields on the two sphere.

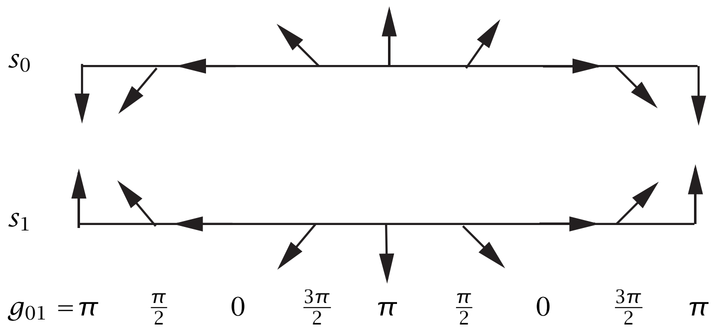

Figure 10.3: The sections s_0 and s_1 restricted to the equator.

Example 10.9 (Hopf bundle.). We can define sections $s_i \colon U_\alpha \to H$ by

$$s_0[z] = ((1, \frac{z^1}{z^0}), [z]) \tag{10.1}$$

$$s_1[z] = ((\frac{z^0}{z^1}, 1), [z]). \tag{10.2}$$

The transition functions are

$$g_{01}([z]) = \frac{z^1}{z^0}.$$

11 Connections, holonomy and curvature

The physical motivation for connections is that you can't do physics if you can't differentiate the fields! So a connection is a rule for differentiating sections of a line bundle. The important thing to remember is that there is no a priori way of doing this - a connection is a *choice* of how to differentiate. Making that choice is something extra, additional structure above and beyond the line bundle itself. The reason for this is that if $L \to M$ is a line bundle and $\gamma : (-\epsilon, \epsilon) \to M$ a path through $\gamma(0) = m$ say and s a section of L then the conventional definition of the rate of change of s in the direction tangent to γ, that is:

$$\lim_{t \to 0} = \frac{s(\gamma(t)) - s(\gamma(0))}{t}$$

makes no sense as $s(\gamma(t))$ is in the vector space $L_{\gamma(t)}$ and $s(\gamma(0))$ is in the *different* vector space $L_{\gamma(0)}$ so that we cannot perform the required subtraction.

So being pure mathematicians we make a definition by abstracting the notion of derivative:

Definition 11.1. A connection ∇ is a linear map

$$\nabla : \Gamma(M, L) \to \Gamma(M, T^*M \otimes L)$$

such that for all s in $\Gamma(M, L)$ and $f \in C^\infty(M, L)$ we have the Liebniz rule:

$$\nabla(fs) = df \otimes s + f \nabla s$$

If $X \in T_x M$ we often use the notation $\nabla_X s = (\nabla s)(X)$.

Example 11.1 (The trivial bundle.). $L = \mathbb{C} \times M$. Then identifying sections with functions we see that (ordinary) differentiation d of functions defines a connection. If ∇ is a general connection then we will see in a moment that $\nabla s - ds$ is a 1-form. So *all* the connections on L are of the form $\nabla = d + A$ for A a 1-form on M (*any* 1-form).

Example 11.2 (The tangent bundle to the sphere.). TS^2. If s is a section then $s : S^2 \to \mathbb{R}^3$ such that $s(u) \in T_u S^2$ that is $\langle s(u), u \rangle = 0$. As $s(u) \in \mathbb{R}^3$ we can differentiate it in $\mathbb{R}^3$ but then ds may not take values in $T_u S^2$ necessarily. We remedy this by defining

$$\nabla(s) = \pi(ds)$$

where π is orthogonal projection from $\mathbb{R}^3$ onto the tangent space to x. That is $\pi(v) = v - \langle x, v \rangle x$.

Example 11.3 (The tangent bundle to a surface.). A surface Σ in $\mathbb{R}^3$. We can do the same orthogonal projection trick as with the previous example.

Example 11.4 (The Hopf bundle.). Because we have $H \subset \mathbb{C}^2 \times \mathbb{C}P_1$ we can apply the same technique as in the previous sections. A section s of H can be identified with a function $s \colon \mathbb{C}P_1 \to \mathbb{C}^2$ such that $s[z] = \lambda z$ for some $\lambda \in \mathbb{C}$. Hence we can differentiate it as a map into $\mathbb{C}^2$. We can then project the result orthogonally using the Hermitian connection on $\mathbb{C}^2$.

The name connection comes from the name infinitesimal connection which was meant to convey the idea that the connection gives an identification of the fibre at a point and the fibre at a nearby 'infinitesimally close' point. Infinitesimally close points are not something we like very much but we shall see in the next section that we can make sense of the 'integrated' version of this idea in as much as a connection, by parallel transport, defines an identification between fibres at endpoints of a path. However this identification is generally path dependent. Before discussing parallel transport we need to consider two technical points.

The first is the question of existence of connections. We have

Proposition 11.2. *Every line bundle has a connection.*

Proof. Let $L \to M$ be the line bundle. Choose an open covering of M by open sets U_α over which there exist nowhere vanishing sections s_α. If ξ is a section of L write it locally as $\xi_{|U_\alpha} = \xi_\alpha s_\alpha$. Choose a partition of unity ρ_α for subordinate to the cover and note that $\rho_\alpha s_\alpha$ extends to a smooth function on all of M. Then define

$$\nabla(\xi) = \sum d\xi_\alpha \rho_\alpha s_\alpha.$$

We leave it as an exercise to check that this defines a connection. □

The second point is that we need to be able to restrict a connection to a open set so that we can work with local trivialisations. We have

Proposition 11.3. *If ∇ is a connection on a line bundle $L \to M$ and $U \subset M$ is an open set then there is a unique connection ∇_U on $L_{|U} \to U$ satisfying*

$$\nabla(s)_{|U} = \nabla_U(s_{|U}).$$

Proof. We first need to show that if s is a section which is zero in a neighbourhood of a point x then $\nabla(s)(x) = 0$. To show this notice that if s is zero on a neighbourhood U of x then we can find a function ρ on M which is 1 outside U and zero in a neighbourhood of x such that $\rho s = s$. Then we have

$$\nabla(s)(x) = \nabla(\rho s)(x) = d\rho(x)s(x) + \rho(x)\nabla(s)(x) = 0.$$

It follows from linearity that if s and t are equal in a neighbourhood of x then $\nabla(s)(x) = \nabla(t)(x)$. If s is a section of L over U and $x \in U$ then we can multiply it by a bump function which is 1 in a neighbourhood of x so that it extends to a section $\hat{s}$ of L over all of M. Then define $\nabla_U(s)(x) = \nabla(\hat{s})(x)$. If we choose a different bump function to extend s to a different section $\tilde{s}$ then $\tilde{s}$ and $\hat{s}$ agree in a neighbourhood of x so that the definition of $\nabla_U(s)(x)$ does not change. □

From now on I will drop the notation $\nabla_{|U}$ and just denote it by ∇.

Let $L \to M$ be a line bundle and $s_\alpha : U_\alpha \to L$ be local nowhere vanishing sections. Define a one-form A_α on U_α by $\nabla s_\alpha = A_\alpha \otimes s_\alpha$. If $\xi \in \Gamma(M, L)$ then $\xi|_{U\alpha} = \xi_\alpha s_\alpha$ where $\xi_\alpha : U_\alpha \to \mathbf{C}$ and

$$\begin{aligned} \nabla\xi|_{U_\alpha} &= d\xi_\alpha s_\alpha + \xi_\alpha \nabla s_\alpha \\ &= (d\xi_\alpha + A_\alpha \xi_\alpha)s_\alpha. \end{aligned} \qquad (11.1)$$

Recall that $s_\alpha = g_{\alpha\beta}s_\beta$ so $\nabla s_\alpha = dg_{\alpha\beta}s_\beta + g_{\alpha\beta}\nabla s_\beta$ and hence $A_\alpha s_\alpha = g_{\alpha\beta}^{-1}dg_{\alpha\beta}g_{\alpha\beta}s_\alpha + s_\alpha A_\beta$. Hence

$$A_\alpha = A_\beta + g_{\alpha\beta}^{-1}dg_{\alpha\beta} \qquad (11.2)$$

The converse is also true. If $\{A_\alpha\}$ is a collection of 1-forms satisfying the equation (11.2) on $U_\alpha \cap U_\beta$ then there is a connection ∇ such that $\nabla s_\alpha = A_\alpha s_\alpha$. The proof is an exercise using equation (11.1) to define the connection.

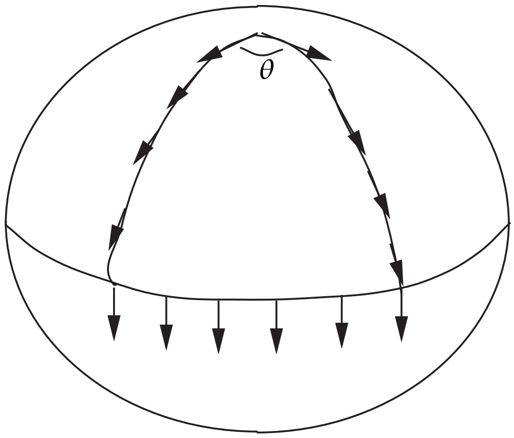

Figure 11.1: Parallel transport on the two sphere.

11.1 Parallel transport and holonomy

If $\gamma : [0,1] \to M$ is a path and ∇ a connection we can consider the notion of moving a vector in $L_{\gamma(0)}$ to $L_{\gamma(1)}$ without changing it, that is *parallel transporting* a vector from $L_{\gamma(0)}, L_{\gamma(1)}$. Here change is measured relative to ∇ so if $\xi(t) \in L_{\gamma(t)}$ is moving without changing it must satisfy the differential equation:

$$\nabla_{\dot{\gamma}} \xi = 0 \tag{11.3}$$

where $\dot{\gamma}$ is the tangent vector field to the curve γ. Assume for the moment that the image of γ is inside an open set U_α over which L has a nowhere vanishing section s_α. Then using (11.3) and letting $\xi(t) = \xi_\alpha(t) s_\alpha(\gamma(t))$ we deduce that

$$\frac{d\xi_\alpha}{dt} = -A_\alpha(\dot{\gamma})\xi_\alpha$$

or

$$\xi_\alpha(t) = \exp\left(-\int_0^t A_\alpha(\dot{\gamma})ds\right)\xi_\alpha(0) \tag{11.4}$$

This is an ordinary differential equation so standard existence and uniqueness theorems tell us that parallel transport defines an isomorphism $L_{\gamma(0)} \cong L_{\gamma(t)}$. Moreover if we choose a curve not inside a special open set like U_α we can still cover it by such open sets and deduce that the parallel transport

$$P_\gamma : L_{\gamma(0)} \to L_{\gamma(1)}$$

is an isomorphism. In general P_γ is dependent on γ and ∇. The most notable example is to take γ a *loop* that is $\gamma(0) = \gamma(1)$. Then we define $\mathrm{hol}(\gamma, \nabla)$, the *holonomy* of the connection ∇ along the curve γ by taking any $s \in L_{\gamma(0)}$ and defining

$$P_\gamma(s) = \mathrm{hol}\ (\gamma, \nabla).s$$

Example 11.5. A little thought shows that ∇ on the two sphere preserves lengths and angles, it corresponds to moving a vector so that the rate of change is normal. If we consider the 'loop' in Figure 4 then we have drawn parallel transport of a vector and the holonomy is $\exp(i\theta)$.

11.2 Curvature

If we have a loop γ whose image is in U_α then we can apply (11.4) to obtain

$$\mathrm{hol}\ (\nabla, \gamma) = \exp\ (-\int_\gamma A_\alpha).$$

If γ is the boundary of a disk D then by Stokes' theorem we have

$$\text{hol}\,(\nabla,\gamma) = \exp - \int_D dA_\alpha. \tag{11.5}$$

Consider the two-forms dA_α. From (11.2) we have

$$\begin{aligned} dA_\alpha &= dA_\beta + d\left(g_{\alpha\beta}^{-1} dg_{\alpha\beta}\right) \\ &= dA_\beta - g_{\alpha\beta}^{-1} dg_{\alpha\beta} g_{\alpha\beta}^{-1} \wedge dg_{\alpha\beta} + g_{\alpha\beta}^{-1} d dg_{\alpha\beta} \\ &= dA_\beta. \end{aligned}$$

So the two-forms dA_α agree on the intersections of the open sets in the cover and hence define a *global* two form that we denote by F and call the *curvature* of ∇. Then we have

Proposition 11.4. *If $L \to M$ is a line bundle with connection ∇ and Σ is a compact submanifold of M with boundary a loop γ then*

$$\text{hol}\,(\nabla,\gamma) = \exp\ - \int_D F$$

Proof. Notice that (11.5) gives the required result if Σ is a disk which is inside one of the U_α. Now consider a general Σ. By compactness we can triangulate Σ in such a way that each of the triangles is in some U_α. Now we can apply (11.5) to each triangle and note that the holonomy up and down the interior edges cancels to give the required result. □

Example 11.6. We calculate the holonomy of the standard connection on the tangent bundle of S^2. Let us use polar co-ordinates: The co-ordinate tangent vectors are:

$$\begin{aligned} \frac{\partial}{\partial\theta} &= (-\sin(\theta)\sin(\phi), \cos(\theta)\sin(\phi), 0) \\ \frac{\partial}{\partial\phi} &= (\cos(\theta)\cos(\phi), \sin(\theta)\cos(\phi), -\sin(\phi)) \end{aligned}$$

Taking the cross product of these and normalising gives the unit normal

$$\begin{aligned} \hat{n} &= (\cos(\theta)\sin(\phi), \sin(\theta)\sin(\phi), \cos(\phi)) \\ &= \sin(\phi)\frac{\partial}{\partial\phi} \times \frac{\partial}{\partial\theta} \end{aligned}$$

To calculate the connection we need a non-vanishing section s we take

$$s = (-\sin(\theta), \cos(\theta), 0)$$

and then

$$ds = (-\cos(\theta), -\sin(\theta), 0)d\theta$$

so that

$$\begin{aligned} \nabla s &= \pi(ds) \\ &= ds - < ds, \hat{n} > \hat{n} \\ &= (-\cos(\theta), -\sin(\theta), 0)d\theta \\ &\quad + \sin(\phi)\ (\cos(\theta)\sin(\phi), \sin(\theta)\sin(\phi), \cos(\phi))d\theta \\ &= (-\cos(\theta)\cos^2(\phi), -\sin(\theta)\cos^2(\phi), \cos(\phi)\sin(\phi))d\theta \\ &= \cos(\phi)\hat{n} \times s \\ &= i\cos(\phi)s \end{aligned}$$

Hence $A = i\cos(\phi)d\theta$ and $F = i\sin(\phi)d\theta \wedge d\phi$. To understand what this two form is note that the volume form on the two-sphere is $\mathrm{vol} = -\sin(\phi)d\theta \wedge d\phi$ and hence $F = i\mathrm{vol}$ The region bounded by the path in Figure 4 has area θ. If we call that region D we conclude that

$$\exp(-\int_D F) = \exp i\theta.$$

Note that this agrees with the previous calculation for the holonomy around this path.

11.3 Curvature as infinitesimal holonomy

The equation $\mathrm{hol}(-\nabla, \partial D) = \exp\,(-\int_D F)$ has an infinitesimal counterpart. If X and Y are two tangent vectors and we let D_t be a parallelogram with sides tX and tY then the holonomy around D_t can be expanded in powers of t as

$$\mathrm{hol}\ (\nabla, D_t) = 1 + t^2\, F(X, Y) + 0(t^3).$$

11.4 Exercises

Exercise 11.1. Let ∇^0 and ∇^1 be connections on a complex line bundle L and define

$$\nabla^t(\phi) = t\nabla^1(\phi) + (1-t)\nabla^0(\phi)$$

for any section ϕ of L. Show that ∇^t is a connection for any real number t. Calculate its curvature.

Exercise 11.2. Show that if $L \to M$ is a trivial bundle then it has zero Chern class.

Exercise 11.3. Consider the Hopf bundle H over $\mathbb{C}P_1$. Define parameters on $U_0 = \mathbb{C}P_1 - [1,0]$ by $(x, y) \mapsto [x + iy, 1]$. Let $s_0([x + iy, 1]) = ([x + iy, 1], (x + iy, 1))$ be the section defined in class. Using (hermitian) orthogonal projection define a connection ∇ on H and calculate the connection one form A_0. Be careful to make the orthogonal projection complex linear. Calculate the curvature over the open set U_0 and integrate it over U_0 to find the Chern class of H. You may find it convenient to work with the complex differential forms $dz = dx + idy$ and $d\bar{z} = dx - idy$.

Exercise 11.4. Consider the tangent bundle to the two-sphere. Give it the connection defined by orthogonal projection and calculate its curvature and hence the chern class of the tangent bundle to the two-sphere.

Exercise 11.5. Repeat Exercise 11.4 for the torus using the co-ordinates defined in Exercise 9.6.

Exercise 11.6. This assumes you are familiar with the Gauss-Bonnet theorem. If Σ is a closed surface in $\mathbb{R}^3$ define a connection on its tangent bundle by using orthogonal projection. Relate the curvature of this connection to the usual Gaussian curvature.

12 Chern classes

In this section we define the Chern class which is a (topological) invariant of a line bundle. Before doing this we collect some facts about the curvature.

Proposition 12.1. *The curvature F of a connection ∇ satisfies:*

(i) $dF = 0$

(ii) If ∇, ∇' are two connections then $\nabla = \nabla' + \eta$ for η a 1-form and $F_\nabla = F_{\nabla'} + d\eta$.

(iii) If Σ is a closed surface then $\frac{1}{2\pi i}\int_\Sigma F_\nabla$ is an integer independent of ∇.

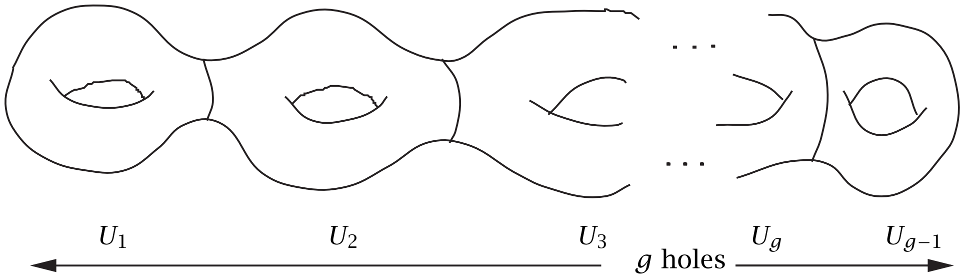

Figure 12.1: A surface of genus g.

Proof. (i) $dF|_{U\alpha} = d(dA_\alpha) = 0$.
(ii) Locally $A'_\alpha = A_\alpha + \eta\alpha$ as $\eta_\alpha = A'_\alpha - A_\alpha$. But $A_\beta = A_\alpha - g_{\alpha\beta}^{-1} dg_{\alpha\beta}$ and $A'_\beta = A'_\alpha - g_{\alpha\beta}^{-1} dg_{\alpha\beta}$ so that $\eta_\beta = \eta_\alpha$. Hence η is a global 1-form and $F_\nabla = dA_\alpha$ so $F'_\nabla = F_\nabla + d\eta$.
(iii) If Σ is a closed surface then $\partial\Sigma = \emptyset$ so by Stokes' theorem $\int_\Sigma F_\nabla = \int_\Sigma F'_\nabla$. Now choose a family of disks D_t in Σ whose limit as $t \to 0$ is a point. For any t the holonomy of the connection around the boundary of D_t can be calculated by integrating the curvature over D_t or over the complement of D_t in Σ and using Proposition 2.1. Taking into account orientation this gives us

$$\exp(\int_{\Sigma - D_t} F) = \exp(-\int_{D_t} F)$$

and taking the limit as $t \to 0$ gives

$$\exp(\int_\Sigma F) = 1$$

which gives the required result. □

The Chern class, $c(L)$, of a line bundle $L \to \Sigma$ where Σ is a surface is defined to be the integer $\frac{1}{2\pi i}\int_\Sigma F_\nabla$ for any connection ∇.

Example 12.1. For the case of the two sphere previous results showed that $F = -i\,\mathrm{vol}_{S^2}$. Hence

$$c(TS^2) = \frac{-i}{2\pi i}\int_{S^2} \mathrm{vol} = \frac{-i}{2\pi i} 4\pi = -2.$$

Some further insight into the Chern class can be obtained by considering a covering of S^2 by two open sets U_0, U_1 as in Figure 2. Let $L \to S^2$ be given by a transition for $g_{01} : U_0 \cap U_1 \to \mathbf{C}^\times$. Then a connection is a pair of 1-forms A_0, A_1, on U_0, U_1 respectively, such that

$$A_1 = A_0 + dg_{10}g_{10}^{-1} \text{ on } U_0 \cap U_1.$$

Take $A_0 = 0$ and A_1 to be any extension of $dg_{10}g_{10}^{-1}$ to U_1. Such an extension can be made by shrinking U_0 and U_1 a little and using a cut-off function. Then $F = dA_0 = 0$ on U_0 and $F = dA_1$ on U_1. To find $c(L)$ we note that by Stokes theorem:

$$\int_{S^2} F = \int_{U_1} F = \int_{\partial U_1} A_1 = \int_{\partial U_1} dg_{10}g_{10}^{-1}.$$

But this is just $2\pi i$ the winding number of g_{10}. Hence the Chern class of L is the winding number of g_{10}. Note that we have already seen that for TS^2 the winding number and Chern class are both -2. It is not difficult to go further now and prove that isomorphism classes of line bundles on S^2 are in one to one correspondence with the integers via the Chern class but will not do this here.

Example 12.2. Another example is a surface Σ_g of genus g as in Figure 5. We cover it with g open sets $U_1, \dots, U_g$ as indicated. Each of these open sets is diffeomorphic to either a torus with a disk removed or a torus with two disks removed. A torus has a non-vanishing vector field on it. If we imagine a rotating bicycle wheel then the inner tube of the tyre (ignoring the valve!) is a torus and the tangent vector field generated by the rotation defines a non-vanishing vector field. Hence the same is true of the open sets in Figure 5. There are corresponding transition functions $g_{12}, g_{23}, \ldots, g_{g-1g}$ and we can define a connection in a manner analogous to the two-sphere case and we find that

$$c(T\Sigma_g) = \sum_{i=1}^{g-1} \text{winding number}(g_{i,i+1}).$$

All the transition functions have winding number -2 so that

$$c(T\Sigma_g) = 2 - 2g.$$

This is a form of the Gauss-Bonnet theorem. It would be a good exercise for the reader familiar with the classical Riemannian geometry of surfaces in $\mathbb{R}^3$ to relate this result to the Gauss-Bonnet theorem. In the classical Gauss-Bonnet theorem we integrate the Gaussian curvature which is the trace of the curvature of the Levi-Civita connection.

So far we have only defined the Chern class for a surface. To define it for manifolds of higher dimension we need to recall the definition of de Rham cohomology [4]. If M is a manifold we have the de Rham complex

$$0 \to \Omega^0(M) \to \Omega^1(M) \to \ldots \to \Omega^m(M) \to 0.$$

where $\Omega^p(M)$ is the space of all p forms on M, the horizontal maps are d the exterior derivative and $m = \dim(M)$. Then $d^2 = 0$ and it makes sense to define:

$$H^p(M) = \frac{\text{kernel } d : \Omega^p\ (M) \to \Omega^{p+1}\ (M)}{\text{image } d : \Omega^{p-1}\ (M) \to \Omega^p\ (M)}$$

This is the pth de Rham cohomology group of M - a finite dimensional vector space if M is compact or otherwise well behaved.

The general definition of $c(L)$ is to take the cohomology class in $H^2(M)$ containing $\frac{1}{2\pi i}F_\nabla$ for some connection.

It is a standard result [4] that if M is oriented, compact, connected and two dimensional integrating representatives of degree two cohomology classes defines an isomorphism

$$\begin{array}{ccc} H^2(M) & \to & \mathbb{R} \\ [\omega] & \mapsto & \int_M \omega \end{array}$$

where $[\omega]$ is a cohomology class with representative form ω. Hence we recover the definition for surfaces.

13 Vector bundles and gauge theories

Line bundles occur in physics in electromagnetism. The electro-magnetic tensor can be interpreted as the curvature form of a line bundle. A very nice account of this and related material is given by Bott in [3]. More interesting however are so-called non-abelian gauge theories which involve vector bundles.

To generalize the previous sections to a vector bundles E one needs to work through replacing $\mathbb{C}$ by $\mathbb{C}^N$ and $\mathbb{C}^\times$ by $GL(n, \mathbb{C})$. Now non-vanishing sections and local trivialisations are not the same thing. A local trivialisation corresponds to a *local frame*, that is n local sections $s_1, \ldots, s_n$ such that $s_1(m), \ldots, s_n(m)$ are a basis for E_m all m. The transition function is then matrix valued

$$g_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(n, \mathbb{C}).$$

The clutching construction still works.

A connection is defined the same way but locally corresponds to matrix valued one-forms A_α. That is

$$\nabla|_{U\alpha}(\Sigma_i \xi^i s_i) = \Sigma_i (d\xi i + \Sigma_j A^i_{\alpha j} \xi^j) s_i$$

and the relationship between A_β and A_α is

$$A_\beta = g^{-1}_{\alpha\beta} \, A_\alpha \, g_{\alpha\beta} + g^{-1}_{\alpha\beta} \, dg_{\alpha\beta}.$$

The correct definition of curvature is

$$F_\alpha = dA_\alpha + A_\alpha \wedge A_\alpha$$

where the wedge product involves matrix multiplication as well as wedging of one forms. We find that

$$F_\beta = g^{-1}_{\alpha\beta} \, F_\alpha \, g_{\alpha\beta}$$

and that F is properly thought of as a two-form with values in the linear operators on E. That is if X and Y are vectors in the tangent space to M at m then $F(X,Y)$ is a linear map from E_m to itself.

We have no time here to even begin to explore the rich geometrical theory that has been built out of gauge theories and instead refer the reader to some references [1, 2, 6, 7].

We conclude with some remarks about the relationship of the theory we have developed here and classical Riemannian differential geometry. This is of course where all this theory began not where it ends! There is no reason in the above discussion to work with complex vector spaces, real vector spaces would do just as well. In that case we can consider the classical example of tangent bundle TM of a Riemannian manifold. For that situation there is a special connection, the Levi-Civita connection. If $(x^1, \ldots, x^n)$ are local co-ordinates on the manifold then the Levi-Civita connection is often written in terms of the Christoffel symbols as

$$\nabla_{\frac{\partial}{\partial x^i}} \left(\frac{\partial}{\partial x^j}\right) = \sum_k \Gamma^k_{ij} \frac{\partial}{\partial x^k}.$$

The connection one-forms are supposed to be matrix valued and they are

$$\sum_i \Gamma^k_{ij} dx^i.$$

The curvature F is the Riemann curvature tensor R. As a two-form with values in matrices it is

$$\sum_{ij} R^k_{ijk} dx^i \wedge dx^j.$$

References

[1] M.F. Atiyah: *Geometry of Yang-Mills Fields*, Lezione Fermione, Pisa 1979. (Also appears in Atiyah's collected works.)

[2] M.F. Atiyah and N.J. Hitchin: *The geometry and dynamics of magnetic monopoles* Princeton University Press, Princeton, 1988.

[3] R. Bott. *On some recent interactions between mathematics and physics.* Canadian Mathematical Bulletin, **28** (1985), no. 2, 129–164.

[4] R. Bott and L.W. Tu: *Differential forms in algebraic topology.* Springer-Verlag, New York.

[5] Y. Choquet-Bruhat, C. DeWitt-Morette and M. Dillard-Bleick: *Analysis, manifolds, and physics* North-Holland, Amsterdam, (1982)

[6] S.K. Donaldson and P.B. Kronheimer *The geometry of four-manifolds* Oxford University Press, (1990).

[7] D. Freed and K. Uhlenbeck. *Instantons and Four-Manifolds* Springer-Verlag, New York (1984).

[8] P. Griffiths and J. Harris: *Principles of algebraic geometry*, Wiley, 1978, New York.

[9] S. Kobayashi and K. Nomizu: *Foundations of Differential Geometry.* Interscience,

[10] S. Lang: *Differential manifolds* (1972)

[11] R.O. Wells: *Differential Analysis on Complex Manifolds* Springer-Verlag, 1973, New York.