Differential Geometry. Honours 1999

Michael Murray

September 18, 2002

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 derivative of any order. Because a differentiable function is continuous it follows that f has continuous partial derivative of any orders. I am quite deliberately avoiding the notation

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

for the time being.

If $f\colon U \to \mathbb{R}^m$ then we say that f is smooth if the functions $f^i\colon U \to \mathbb{R}$ are smooth where $f(x) = (f^1(x), f^2(x), \ldots, f^m(x))$. Notice that in this case the limit definition of derivative makes sense and we can define

$$f'(x) = df(x) = d_x f = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}.$$

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 as a *linear* map

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

defined by

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

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

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

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 , $f\colon U \to \mathbb{R}^m$ a smooth function, $V \subset \mathbb{R}^m$ open and $g\colon V \to \mathbb{R}^k$ a smooth function with $f(U) \subset V$. Let $x \in U$. Then $f \circ g$ is a smooth function and*

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

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 1.1. *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 in turn. Then x^i is smooth so $x^i \circ \phi = \phi^i$ is smooth. □

1.5 Diffeomorphisms and the inverse function theorem.

A function $f: U \to V$ where U and V are open subsets of $\mathbb{R}^n$ is called a *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 $f: U \to \mathbb{R}^n$ 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 $f: V \to f(V)$ is a diffeomorphism.

With this notion we have the important inverse function theorem:

Theorem 1.2 (Inverse Function Theorem). *Let U be an open subset of $\mathbb{R}^n$ and $f: U \to \mathbb{R}^n$ 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 diffeomorphism:

Lemma 1.2. *Let U and V be open subsets of $\mathbb{R}^n$. A bijection $\phi: U \to V$ is a diffeomorphism if and only if for every function $f: V \to \mathbb{R}$ we have that f is differentiable if and only if $f \circ \phi: U \to \mathbb{R}$ is differentiable.*

Proof. We just apply Lemma **??** 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. First we define:

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

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 $1_{R^n} : \mathbb{R}^n \to \mathbb{R}^n$ be the identity map. That is $1_{R^n}(x) = (x^1, \dots, x^n)$. Then $(\mathbb{R}^n, 1_{\mathbb{R}^n})$ 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

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

the inclusion map defined by $\iota(x) = x$. Then clearly $\iota(U) = U$ which is open 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 $\psi: V \to \mathbb{R}^n$ 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 $\phi: V \to \mathbb{R}^n$ 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
$$U = \mathbb{R}^2 - \{(x, y) \mid y < 0\}$$
and define polar co-ordinates
$$(r, \theta)\colon U \to (0, \infty) \times (-\pi, \pi) \subset \mathbb{R}^2$$
as follows. We define $r(x, y) = \sqrt{x^2 + y^2}$ we define θ 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 $(r, \theta$ 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
$$(x - 0, y - 0, z - 1) = t(\psi_0^1(p), \psi_0^2(p), -1)$$
and hence
$$\psi_0^1(x, y, z) = \frac{x}{1 - z}$$
and
$$\psi_0^2(x, y, z) = \frac{y}{1 - z}.$$

In general a manifold will have lots of co-ordinates. We don't expect a manifold to come with a given set of co-ordinates anymore that 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 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 differentiable. 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 differentiable. However we can ask that f be differentiable 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 differentiable. 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 differentiable 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 differentiable if $f \circ \chi^{-1}$ is is that $\chi \circ \psi^{-1}$ is differentiable. As we want this to work both ways we also require that $\psi \circ \chi^{-1}$ be differentiable. 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 **??** 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.

Example 2.6. If $U \subset \mathbb{R}^2$ is the set in example **??** 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.

If we calculate the composition

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

we obtain

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

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) where

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

and

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

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

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

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

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

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

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 *non-zero* vector in $\mathbb{R}^{n+1}$ we denote by $[x] = [x^0, \dots, x^n]$ the line through it. 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 $x = \lambda y$. The numbers $x = (x^0, \dots, x^n)$ are often called the *homogeneous co-ordinates* of the line $[x]$. Define a subset $U_i \subset \mathbb{R}P_n$ 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

$$\begin{array}{cccc} \psi_i & : & U_i & \to & \mathbb{R}^n \\ & & [x^0, \dots, x^n] & \mapsto & \left(\frac{x^0}{x^i}, \dots, \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i}, \dots, \frac{x^n}{x^i}\right). \end{array}$$

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 $\mathcal{A} = \{(\mathbb{R}^2 1_{\mathbb{R}^2}\}$. 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 2.1. *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 in $\bar{A}$ 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

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

is a diffeomorphism. Notice that 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

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

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 see that

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

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 2.6. If M is a manifold with atlas $\mathcal{A}$ we define a co-ordinate chart on the manifold M 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 $U \in \mathcal{T}$.

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 2.7. Define a co-ordinate chart (U, ψ) on a set V to be a linear co-ordinate chart if $U = V$, $\psi(U) = \mathbb{R}^n$ and ψ is a bijection. $\psi : V \to \mathbb{R}^n$.

Example 2.13. If V is a vector space and $\psi\colon V \to \mathbb{R}^n$ is a linear isomorphism then (V, ψ) is a linear co-ordinate chart

Definition 2.8. Define two sets of linear co-ordinates ψ and χ to be linearly compatible if $\psi \circ \chi^{-1}$ is a linear isomorphism.

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

Definition 2.9. A linear atlas on a set V is an equivalence class of linear co-ordinates.

Definition 2.10. 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 2.2. *A linear manifold has a natural vector space structure which makes all of the linear co-ordinates linear isomorphisms.*

Because of Proposition **??** 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 we have

Definition 2.11. If V is a linear manifold and $f\colon V \to \mathbb{R}$ is a function we call f *linear* if $f \circ \psi^{-1}$ is linear for some choice of linear co-ordinates ψ.

We leave it as an exercise to prove that if f is linear then $f \circ \chi^{-1}\colon V \to \mathbb{R}$ is linear for any choice of linear co-ordinates χ and that, moreover, f is linear with respect to the addition and scalar multiplication defined in **??**.

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 2.1. *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.

We will in general require a manifold to be Hausdorff and paracompact in the topology.

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 it the dimension of the manifold.

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 do that 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 3.1. *If $f\colon M \to \mathbb{R}$ is a smooth function and (V, χ) is a co-ordinate chart then $f \circ \chi^{-1}\colon \chi(V) \to \mathbb{R}$ 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 into a manifold or paths. We have

Definition 3.2. 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 3.2. *If $\gamma\colon (-\epsilon, \epsilon) \to M$ is a smooth path in M and (V, χ) is a chart with $\gamma((-\epsilon, \epsilon)) \subset V$ then $\chi \circ \gamma$ is smooth.*

Proof. Chain rule and compatibility. □

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

$$(f \circ \gamma)'(0) = df(x)(\gamma'(0)) \tag{4.1}$$

where $\gamma'(0)$ is the tangent vector to γ at $t = 0$. Notice that $\gamma'(0)$ is a vector in $\mathbb{R}^n$ defined by

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

and that $df(x)\colon \mathbb{R}^n \to \mathbb{R}$ is a linear map.

Equation (**??**) tells us that $(f \circ \gamma)'(0)$ 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

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

Now 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 though a point x. We say that γ and ρ are tangent at $t = 0$ if there is a co-ordinate chart (U, ψ) with $\gamma((-\epsilon, \epsilon)) \subset U$ and $\rho((-\delta, \delta)) \subset U$ and

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

Again we have the usual lemma.

Lemma 4.1. *Let $\gamma\colon (-\epsilon, \epsilon) \to M$ and $\rho\colon (-\delta, \delta) \to M$ be paths though a point x which are tangent at $t = 0$. Then if (V, χ) is a chart with $\gamma((-\epsilon, \epsilon)) \subset V$ and $\rho((-\delta, \delta)) \subset V$ then*

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

Proof. Chain rule and compatibility. □

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 4.1. *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 $(\psi \circ \gamma)'(0) = v$. □

Lemma 4.2. *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. □

From the discussion in the previous section the maps $d\chi(x)$ define linear co-ordinates on T_xM and hence by Proposition **??** 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 taking values 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. The relationship is that the equivalence class of paths is the unique one containing the path $t \mapsto x = x + t\gamma'(0)$. Again in this case the extra machinery of equivalence classes of paths adds nothing to what we already know.

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

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

It follows that $(f \circ \gamma)'(0) = (f \circ \rho)'(0)$ if ρ and γ are in the same tangency class. Hence if $X = t_0(X)$ 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*.

4.2 Co-ordinate tangent vectors and one-forms.

Let (U, ψ) be a set of co-ordinates on M where $\psi = (\psi^1, \ldots, \psi^n)$. 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.

Recall that

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

is a linear isomorphism. We denote by

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

the pre-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

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

Notice that if $y = (y^1, \ldots, y^n)$ 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}.

Recall from linear algebra that if $v_1, v_2, \ldots, v_n$ is a basis of a vector space V and $w^1, w^2, \ldots w^n$is a collection of vectors in V^* satisfying $w^i(v_j) = \delta^i_j$ for all i, j then the w_i are a basis of V^* called the basis dual to $v_1, \ldots, v_n$. Hence $d\psi^1(x), \ldots, d\psi^n(x)$ 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 basis. 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 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 $\mathbb{R}^n$. The definition which follows looks a little complicated. The basic idea is that we regard a linear subspace like

$$\{(x^1, \dots, x^k, 0, \dots, 0) \mid (x^1, \dots, x^k) \in \mathbb{R}^k\}$$

as a 'nice' subset of $\mathbb{R}^n$ (for $k \le n$) and a submanifold is a subset which is 'locally' of this form. We begin with

Definition 4.2. Let Z be a subset of $\mathbb{R}^n$. We call a co-ordinate chart (U, ψ) on $\mathbb{R}^n$ d-compatible with Z if there an integer d such that

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

Definition 4.3. Let Z be a subset of $\mathbb{R}^n$. If we can cover Z with domains of co-ordinate charts which are all d-compatible with Z for a fixed d then we call Z a submanifold of dimension d.

Notice that we do not know yet that a submanifold is a manifold ! To rectify this let Z be a submanifold of dimension d and let (U, ψ) be a set of compatible co-ordinates. That is

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

Then consider $(U \cap Z, \bar\psi)$ where $\bar\psi = (\psi^1_{|U\cap Z}, \dots, \psi^d_{|U\cap Z})$. We would like this to be 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 compatible 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

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

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

We have now proved

Theorem 4.1. *If Z is a submanifold the set of charts above is an atlas.*

Lots of examples of submanifolds are provided by the following theorem.

Theorem 4.2. *Let $f \colon \mathbb{R}^n \to \mathbb{R}^{n-d}$ be a smooth map and let $Z = f^{-1}(0)$. Then if $df(z)$ is onto for all $z \in Z$ then Z is a submanifold of dimension d.*

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

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

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

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

and note that $\psi(z) = 0$. Because π is a linear map it is its own derivative, that is for any v we have $d\pi^i(z)(v) = \pi^i(v)$ for all i. So we have

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

Consider v in the kernel of $d\psi(z)$. Then $d\pi(z)(v) = \pi^i(v) = 0$ 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 by 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 $U \cap Z$. □

Example 4.3. Consider the circle in $\mathbb{R}^2$. It is the zero set of the map $f \colon \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x^2 + y^2 - 1$. So $df(x, y)(v, w) = 2xv + 2yw$. If we take $z = (0, 1)$ then the $K_z = \{(v, 0) \mid v \in \mathbb{R}\}$. The projection π is then $\pi(v, w) = (v, 0)$ and the map ψ is

$$\psi(x, y) = (x, x^2 + y^2 - 1).$$

Example 4.4 *(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 maps.

Example 4.5 *(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 $f(X) = X^2 - 1$. 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 $n^2 - d = n(n-1)/2$.

4.5 Tangent space to a submanifold

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 $i = 1, \ldots, d$. Because γ has image inside Z we also have that $\psi^i \circ \gamma = 0$ for each $i = d+1, \ldots, n$ 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 $df(x)(\gamma'(0)) = 0$ so that $\gamma'(0) \in K_x$.

We can now define a map $T_xZ \to K_x$. If $X \in T_xZ$ then we choose a path γ in Z 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 $i = 1, \ldots, d$. Hence we also have

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

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

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

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

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

for $i = 1, \ldots, n$. 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.

4.6 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 4.4. Let $f: M \to N$ be a map between manifolds. Then f is called smooth if for every point $x \in M$ there are co-ordinate chart (U, ψ) on M and (V, χ) such that $x \in U$ and $f(U) \subset V$ and

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

is smooth.

Again we have the same sort of lemma:

Lemma 4.3. *Let $f: M \to N$ be a smooth map between manifolds. Assume that there are co-ordinate charts (U, ψ) on M and (V, χ) such that $f(U) \subset V$. Then*

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

is smooth.

4.7 The tangent to a smooth map.

If $f: M \to N$ is a smooth map and $\gamma: (-\epsilon, \epsilon) \to M$ 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

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

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

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

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 4.6. The tangent space to $\mathbb{R}^n$ at $\psi(x)$ is just $\mathbb{R}^n$ again. The map $d\psi(x): T_x M \to \mathbb{R}^n$ is just the map $T_x \psi: T_x M \to T_{\psi(x)} \mathbb{R}^n$.

Example 4.7. If $F: \mathbb{R}^n \to \mathbb{R}^m$ 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 4.2 (Chain Rule). *Let $f: M \to N$ and $g: N \to P$ be smooth functions. Then the map $g \circ f: M \to P$ is smooth and $T_x(g \circ f) = T_{f(x)}(g) \circ T_x(f)$.*

4.8 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 4.5 (Submanifolds). We say that a subset $Z \subset M$ is a submanifold of dimension d of a manifold M of dimension m if for every $\in Z$ we can find a co-ordinate chart (U, ψ) for M with $z \in Z$ and such that

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

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

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

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

Proposition 4.3. *The set consisting of all the charts $(U \cap Z, \bar{\psi})$ constructed in this manner is an atlas. Moreover it makes N a manifold in such a way that the inclusion map $\iota_Z \colon Z \to M$ defined by $\iota_Z(y) = y$ is smooth.*

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

Proposition 4.4. *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 $T_z f$ is onto for all $z \in M$ the set Z is submanifold of M. Moreover the image of $T_z(\iota_Z)$ in $T_z M$ is precisely the kernel of $T_z F$.*

4.9 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 $X(x) \in T_x M$ 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 4.6. 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 4.4. *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 **??** 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

$$X_{|U_\alpha} = \sum_{i=1}^{n} X^i_\alpha \frac{\partial}{\partial \psi^i} \tag{4.2}$$

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

$$X^i_\alpha(x) = \sum_{j=1}^{n} \frac{\partial \psi^i_\alpha}{\partial \psi^j_\beta}(x) X^j_\beta(x). \tag{4.3}$$

The converse is also true. If we have a collection of maps $X^i_\alpha \colon U_\alpha \to \mathbb{R}$ satisfying **??** then we can define a vector field using **??** 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 **??** 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.

4.10 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’.

5 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.

5.1 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

$$\omega(v_1, \ldots, v_{i-1}, \alpha v + \beta w, v_{i+1}, v_k) = \alpha\omega(v_1, \ldots, v_{i-1}, v, v_{i+1}, v_k) \\ +\beta\omega(v_1, \ldots, v_{i-1}, w, v_{i+1}, v_k).$$

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

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

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

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

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

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

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 $v_1, \ldots, v_n$ is a basis of V. Let ω be a k form. Then if $w_1, \ldots, w_k$ 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

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

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

Proposition 5.1. *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

$$\alpha^1 \wedge \alpha^2 \wedge \cdots \wedge \alpha^m = \sum_{\pi \in S_m} \operatorname{sgn}(\pi)\alpha^1(v_{\pi(1)})\alpha^2(v_{\pi(2)}) \dots \alpha^m(v_{\pi(m)}).$$

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 $v_1, \dots, v_n$ of V. Define the dual basis of V^*, $\alpha^1, \dots, \alpha^n$, 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 $i_1, \dots, i_k$ 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 $1 \leq i_1 < \cdots < i_k \leq n$. We show first that these elements span $\Lambda^k(V^*)$. Let ω be an element of $\Lambda^k(V^*)$. Notice that

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

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 $w_1, \dots, w_k$ and expand them in the basis as

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

Then we have

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

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

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

and consider

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

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

Proposition 5.2. *The vectors $v_{i_1} \wedge \cdots \wedge v_{i_k}$ where $1 \leq i_1 < \cdots < i_k \leq n$ are a basis for $\Lambda^k(V^*)$.*

Proof. We have already seen that these vectors span. It suffices to show that they are linearly independent. We do this by induction on n. If $n = 1$ then the result is clear as the only non-trivial case is $k = 1$ when the result is straightforward. More generally assume we have a linear relation amongst some of the basis vectors. There has to be an index i such that the corresponding v_i does not occur in all the vectors in that linear relation. Otherwise there is only one vector in the linear relation and that is not possible. Then wedge the whole relation with v_i. The terms containing v_i disappear and we obtain a relation between the vectors constructed for the case of a dimension less so by induction that is not possible. □

It is sometimes useful to sum over all k-tuples $i_1, \ldots, i_k$ not just ordered ones. We can do this — and 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 5.1. 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.

5.2 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

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

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

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

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 $\omega_{i_1, \ldots, i_k}$ as

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

We define a smooth differential form by

Definition 5.1 (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 5.1. *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

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

From these calculations we deduce the following Proposition.

Proposition 5.3. *Let $\{(U_\alpha, \psi_\alpha)\}$ be an atlas for a manifold M. Assume that for each U_α we have a collection of functions*

$$\omega^\alpha_{i_1,\dots,i_k} : U_\alpha \to \mathbb{R}$$

(antisymmetric in their indices) and such that for any pair of open sets U_α and U_β we have

$$\omega^\alpha_{i_1,\dots,i_k} = \sum_{a_1,\dots,a_k=1}^{n} \frac{\partial \psi_\beta^{a_1}}{\partial \psi_\alpha^{i_1}} \cdots \frac{\partial \chi_\psi^{a_n}}{\partial \psi_\alpha^{i_n}} \omega^\beta_{i_1,\dots,i_k}.$$

Then there is a globally defined k form ω on M given locally by

$$\omega_{|U_\alpha}(x) = \sum_{i_1,\dots,i_k} \frac{1}{k!} \omega^\alpha_{i_1,\dots,i_k} d\psi_\alpha^{i_1} \wedge \cdots \wedge d\psi_\alpha^{i_k}$$

Note that the notation of subscripting co-ordinates is cumbersome and we usually try very hard to avoid it! However Proposition **??** underpins the following result on the exterior derivative.

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:

$$d(fg) = f dg + (df)g.$$

We want to prove

Proposition 5.4. *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 $p = 0, \dots, n-1$ 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 will be making use of Proposition **??** but will avoid making this explicit because of the unwieldy notation. 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

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

Then define

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

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

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

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

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

then we have

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

or

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

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

Example 5.2. Recall from **??** 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.

5.3 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

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

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 5.5. *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. □

5.4 Integration of differential forms

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,

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

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 on an open set V in $\mathbb{R}^n$ 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

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

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 5.6. *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 $\phi_a \omega$ is in U_α. First we define the integral of each of the forms $\phi_\alpha \omega$

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

Then we define the integral of ω to be

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

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

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

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

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

If the diffeomorphism

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

is orientation preserving then we have

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

So we can complete the calculation above and have

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

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.

5.5 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 5.2. 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 5.3 (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 5.7. *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. □

5.6 Integration again

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

6 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.

6.1 Manifolds with boundary.

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

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

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

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

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.

6.2 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. To state Stoke's theorem precisely we need to orient M and ∂M. Assume that M is oriented so that $(\psi^0, \dots, \psi^n)$ are positive co-ordinates near the boundary chosen so that they are defined for $\psi^0 \geq 0$. Then we orient ∂M so that $(\psi^1_{|\partial M}, \dots, \psi^{n-1}_{|\partial M})$ are *negatively* oriented co-ordinates on ∂M. We want to prove

Theorem 6.1 (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

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

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

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

or equivalently to prove 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 $\mathbb{R}^n$. Write $x = (t, y)$ for $y \in \mathbb{R}^{n-1}$ and let

$$\omega = f(t, y) dy^1 \dots dy^{n-1} + \sum_{i=1}^{n-1} g_i(t, y) dt \wedge dy^1 \wedge \dots \wedge \widehat{dy^i} \wedge \dots \wedge dy^{n-1}$$

where we follow the usual convention that the hat symbol $\widehat{\ }$ marks a term that is *missing*. Note that because ω is compactly supported there is an $R > 0$ that $f(t, y)$ and each $g_i(t, y)$ are zero if $t > R$ and each $|y^i| > R$. If we orient the boundary using the outward normal then $-dy^1 \wedge \dots \wedge dy^{n-1}$ is negative on the boundary so that we have

$$\int_{\partial U} \omega = -\int f(0, y) dy^1 \wedge \dots \wedge dy^{n-1}.$$

Now we have

$$d\omega = \left(\frac{\partial f(t, y)}{\partial t} + \sum_{i=1}^{n-1} \frac{\partial g_i(t, y)}{\partial y^i} \right) dt \wedge dy^1 \wedge \dots \wedge dy^{n-1}.$$

Note that

$$\int \frac{\partial g_i(t, y)}{\partial y^i} dy^i = 0$$

by the compactness of support for any i. Hence we have

$$\begin{aligned}\int_U d\omega &= \int_U \frac{\partial f}{\partial t} dt \wedge dy^1 \dots dy^n \\ &= -\int_{\partial U} f(0, y^1, \dots, y^{n-1}) dy^1 \dots dy^n \\ &= \int_{\partial U} \omega\end{aligned}$$

as required. □

Note 6.1. To see that the orientations make sense consider a function $f(x, y)$ defined on $[0, \infty)$. Then the fundamental theorem of calculus says that

$$\int_{[0,\infty)} df = \int_0^\infty \frac{\partial f}{\partial x} dx = -f(0).$$

A Partitions of unity.

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.

B Vector fields and the tangent bundle.

We have seen how to define tangent vectors at a point of a manifold. In many problems we are interested in vector fields, that is a choice of vector at every point of a manifold. We can think of this in the following manner. 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 $\pi \colon TM \to M$ 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 $X \colon M \to TM$ 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 $X \colon M \to TM$ 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 that immediately necessary.

Let E be a set with a surjection $\pi \colon E \to M$ 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

$$\phi_a(x) \colon E_x \to V$$

such that the map

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

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

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

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 $(x, \phi_a(x) \circ \phi_b(x)^{-1} v)$ 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 ??. 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 $V = R^n$ and $\phi_a(x) = d\psi_a(x)$. The condition we require to hold is that the map

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

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

We can now define

Definition B.1. A smooth vector field on a manifold M 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 $i = 1, \dots, n$. 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

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

This proves

Proposition B.1. *LetX 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.

C Vector fields and derivations.

Let us now define derivations of $C^\infty(M)$.

Definition C.1. A derivation of $C^\infty(M)$ is a linear map

$$D\colon C^\infty(M) \to C^\infty(M)$$

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 C.1. *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 above 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 by ?? $X(x)$ is smooth. $\square$

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

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

D Tensor products

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 D.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

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

$\square$

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 D.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 **??** 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 D.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 $X : V \to V'$ is linear and $Y : W \to W'$ 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 D.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

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