DIFFERENTIABLE MANIFOLDS

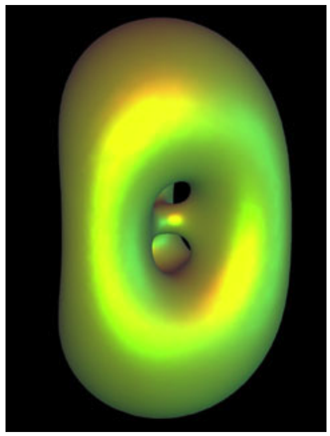
Section c course 2003

Nigel Hitchin

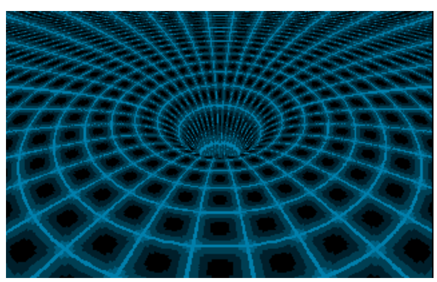
hitchin@maths.ox.ac.uk

1 Introduction

This is an introductory course on differentiable manifolds. These are higher dimensional analogues of surfaces like this:



This is the image to have, but we shouldn't think of a manifold as always sitting inside a fixed Euclidean space like this one, but rather as an abstract object. One of the historical driving forces of the theory was General Relativity, where the manifold is four-dimensional spacetime, wormholes and all:



Spacetime is not part of a bigger Euclidean space, it just exists, but we need to learn how to do analysis on it, which is what this course is about.

Another input to the subject is from mechanics – the dynamics of complicated mechanical systems involve spaces with many degrees of freedom. Just think of the different configurations that an Anglepoise lamp can be put into:

How many degrees of freedom are there? How do we describe the dynamics of this if we hit it?

The first idea we shall meet is really the defining property of a manifold – to be able to describe points locally by n real numbers, *local coordinates*. Then we shall need to define analytical objects (*vector fields, differential forms* for example) which are independent of the choice of coordinates. This has a double advantage: on the one hand it enables us to discuss these objects on topologically non-trivial manifolds like spheres, and on the other it also provides the language for expressing the equations of mathematical physics in a coordinate-free form, one of the fundamental principles of relativity.

The most basic example of analytical techniques on a manifold is the theory of differential forms and the exterior derivative. This generalizes the grad, div and curl of ordinary three-dimensional calculus. A large part of the course will be occupied with this. It provides a very natural generalization of the theorems of Green and Stokes in three dimensions and also gives rise to *de Rham cohomology* which is an analytical way of approaching the algebraic topology of the manifold. This has been important in an enormous range of areas from algebraic geometry to theoretical physics.

More refined use of analysis requires extra data on the manifold and we shall simply define and describe some basic features of Riemannian metrics. These generalize the first fundamental form of a surface and, in their Lorentzian guise, provide the substance of general relativity. A more complete story demands a much longer course, but here we shall consider just two aspects which draw on the theory of differential forms: the study of geodesics via a vector field, the geodesic flow, on the cotangent bundle, and some basic properties of harmonic forms.

Certain standard technical results which we shall require are proved in the Appendix so as not to interrupt the development of the theory.

A good book to accompany the course is: *An Introduction to Differential Manifolds* by Dennis Barden and Charles Thomas (Imperial College Press £19 (paperback)).

2 Manifolds

2.1 Coordinate charts

The concept of a manifold is a bit complicated, but it starts with defining the notion of a *coordinate chart.*

Definition 1 *A coordinate chart on a set X is a subset $U \subseteq X$ together with a bijection*

$$\varphi : U \to \varphi(U) \subseteq \mathbf{R}^n$$

onto an open set $\varphi(U)$ in $\mathbf{R}^n$.

Thus we can parametrize points of U by n coordinates $\varphi(x) = (x_1, \ldots, x_n)$.

We now want to consider the situation where X is covered by such charts and satisfies some consistency conditions. We have

Definition 2 *An n-dimensional atlas on X is a collection of coordinate charts $\{U_\alpha, \varphi_\alpha\}_{\alpha \in I}$ such that*

- *X is covered by the $\{U_\alpha\}_{\alpha \in I}$*
- *for each $\alpha, \beta \in I$, $\varphi_\alpha(U_\alpha \cap U_\beta)$ is open in $\mathbf{R}^n$*
- *the map*

  $$\varphi_\beta \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$$

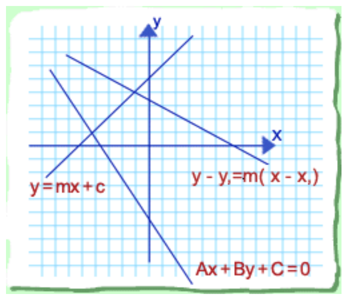
  is C^∞ with C^∞ inverse.

Recall that $F(x_1, \ldots, x_n) \in \mathbf{R}^n$ is C^∞ if it has derivatives of all orders. We shall also say that F is *smooth* in this case. It is perfectly possible to develop the theory of manifolds with less differentiability than this, but this is the normal procedure.

Examples:

1. Let $X = \mathbf{R}^n$ and take $U = X$ with $\varphi = id$. We could also take X to be any open set in $\mathbf{R}^n$.

2. Let X be the set of straight lines in the plane:



Each such line has an equation $Ax + By + C = 0$ where if we multiply A, B, C by a non-zero real number we get the same line. Let U_0 be the set of non-vertical lines. For each line $\ell \in U_0$ we have the equation

$$y = mx + c$$

where m, c are uniquely determined. So $\varphi_0(\ell) = (m, c)$ defines a coordinate chart $\varphi_0 : U_0 \to \mathbf{R}^2$. Similarly if U_1 consists of the non-horizontal lines with equation

$$x = \tilde{m}y + \tilde{c}$$

we have another chart $\varphi_1 : U_1 \to \mathbf{R}^2$.

Now $U_0 \cap U_1$ is the set of lines $y = mx + c$ which are not horizontal, so $m \neq 0$. Thus

$$\varphi_0(U_0 \cap U_1) = \{(m, c) \in \mathbf{R}^2 : m \neq 0\}$$

which is open. Moreover, $y = mx + c$ implies $x = m^{-1}y - cm^{-1}$ and so

$$\varphi_1 \varphi_0^{-1}(m, c) = (m^{-1}, -cm^{-1})$$

which is smooth with smooth inverse. Thus we have an atlas on the space of lines.

3. Consider $\mathbf{R}$ as an additive group, and the subgroup of integers $\mathbf{Z} \subset \mathbf{R}$. Let X be the quotient group $\mathbf{R}/\mathbf{Z}$ and $p : \mathbf{R} \to \mathbf{R}/\mathbf{Z}$ the quotient homomorphism.

Set $U_0 = p(0, 1)$ and $U_1 = p(-1/2, 1/2)$. Since any two elements in the subset $p^{-1}(a)$ differ by an integer, p restricted to $(0, 1)$ or $(-1/2, 1/2)$ is injective and so we have coordinate charts

$$\varphi_0 = p^{-1} : U_0 \to (0, 1), \quad \varphi_1 = p^{-1} : U_1 \to (-1/2, 1/2).$$

Clearly U_0 and U_1 cover $\mathbf{R}/\mathbf{Z}$ since the integer $0 \in U_1$.

We check:

$$\varphi_0(U_0 \cap U_1) = (0, 1/2) \cup (1/2, 1), \quad \varphi_1(U_0 \cap U_1) = (-1/2, 0) \cup (0, 1/2)$$

which are open sets. Finally, if $x \in (0, 1/2)$, $\varphi_1 \varphi_0^{-1}(x) = x$ and if $x \in (1/2, 1)$, $\varphi_1 \varphi_0^{-1}(x) = x - 1$. These maps are certainly smooth with smooth inverse so we have an atlas on $X = \mathbf{R}/\mathbf{Z}$.

4. Let X be the extended complex plane $X = \mathbf{C} \cup \{\infty\}$. Let $U_0 = \mathbf{C}$ with $\varphi_0(z) = z \in \mathbf{C} \cong \mathbf{R}^2$. Now take

$$U_1 = \mathbf{C} \backslash \{0\} \cup \{\infty\}$$

and define $\varphi_1(\tilde{z}) = \tilde{z}^{-1} \in \mathbf{C}$ if $\tilde{z} \neq \infty$ and $\varphi_1(\infty) = 0$. Then

$$\varphi_0(U_0 \cap U_1) = \mathbf{C} \backslash \{0\}$$

which is open, and

$$\varphi_1 \varphi_0^{-1}(z) = z^{-1} = \frac{x}{x^2 + y^2} - i \frac{y}{x^2 + y^2}.$$

This is a smooth and invertible function of (x, y). We now have a 2-dimensional atlas for X, the extended complex plane.

5. Let X be n-dimensional *real projective space*, the set of 1-dimensional vector subspaces of $\mathbf{R}^{n+1}$. Each subspace is spanned by a non-zero vector v, and we define $U_i \subset \mathbf{R}\mathrm{P}^n$ to be the subset for which the i-th component of $v \in \mathbf{R}^{n+1}$ is non-zero. Clearly X is covered by $U_1, \ldots, U_{n+1}$. In U_i we can uniquely choose v such that the ith component is 1, and then U_i is in one-to-one correspondence with the hyperplane $x_i = 1$ in $\mathbf{R}^{n+1}$, which is a copy of $\mathbf{R}^n$. This is therefore a coordinate chart

$$\varphi_i : U_i \to \mathbf{R}^n.$$

The set $\varphi_i(U_i \cap U_j)$ is the subset for which $x_j \neq 0$ and is therefore open. Furthermore

$$\varphi_i \varphi_j^{-1} : \{x \in \mathbf{R}^{n+1} : x_j = 1, x_i \neq 0\} \to \{x \in \mathbf{R}^{n+1} : x_i = 1, x_j \neq 0\}$$

is

$$v \mapsto \frac{1}{x_i} v$$

which is smooth with smooth inverse. We therefore have an atlas for $\mathbf{R}\mathrm{P}^n$.

2.2 The definition of a manifold

All the examples above are actually manifolds, and the existence of an atlas is sufficient to establish that, but there is a minor subtlety in the actual definition of a manifold due to the fact that there are lots of choices of atlases. If we had used a different basis for $\mathbf{R}^2$, our charts on the space X of straight lines would be different, but we would like to think of X as an object independent of the choice of atlas. That's why we make the following definitions:

Definition 3 *Two atlases $\{(U_\alpha, \varphi_\alpha)\}$, $\{(V_i, \psi_i)\}$ are compatible if their union is an atlas.*

What this definition means is that all the extra maps $\psi_i \varphi_\alpha^{-1}$ must be smooth. Compatibility is clearly an equivalence relation, and we then say that:

Definition 4 *A differentiable structure on X is an equivalence class of atlases.*

Finally we come to the definition of a manifold:

Definition 5 *An n-dimensional differentiable manifold is a space X with a differentiable structure.*

The upshot is this: to prove something is a manifold, all you need is to find one atlas. The definition of a manifold takes into account the existence of many more atlases.

Many books give a slightly different definition – they start with a topological space, and insist that the coordinate charts are homeomorphisms. This is fine if you see the world as a hierarchy of more and more sophisticated structures but it suggests that in order to prove something is a manifold you first have to define a topology. As we'll see now, the atlas does that for us.

First recall what a topological space is: a set X with a distinguished collection of subsets V called *open sets* such that

1. $\emptyset$ and X are open
2. an arbitrary union of open sets is open
3. a finite intersection of open sets is open

Now suppose M is a manifold. We shall say that a subset $V \subseteq M$ is open if, for each α, $\varphi_\alpha(V \cap U_\alpha)$ is an open set in $\mathbf{R}^n$. One thing which is immediate is that $V = U_\beta$ is open, from Definition 2.

We need to check that this gives a topology. Condition 1 holds because $\varphi_\alpha(\emptyset) = \emptyset$ and $\varphi_\alpha(M \cap U_\alpha) = \varphi_\alpha(U_\alpha)$ which is open by Definition 1. For the other two, if V_i is a collection of open sets then because φ_α is bijective

$$\varphi_\alpha((\cup V_i) \cap U_\alpha) = \cup \varphi_\alpha(V_i \cap U_\alpha)$$
$$\varphi_\alpha((\cap V_i) \cap U_\alpha) = \cap \varphi_\alpha(V_i \cap U_\alpha)$$

and then the right hand side is a union or intersection of open sets. Slightly less obvious is the following:

Proposition 2.1 *With the topology above $\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha)$ is a homeomorphism.*

Proof: If $V \subseteq U_\alpha$ is open then $\varphi_\alpha(V) = \varphi_\alpha(V \cap U_\alpha)$ is open by the definition of the topology, so φ_α^{-1} is certainly continuous.

Now let $W \subset \varphi_\alpha(U_\alpha)$ be open, then $\varphi_\alpha^{-1}(W) \subseteq U_\alpha$ and U_α is open in M so we need to prove that $\varphi_\alpha^{-1}(W)$ is open in M. But

$$\varphi_\beta(\varphi_\alpha^{-1}(W) \cap U_\beta) = \varphi_\beta \varphi_\alpha^{-1}(W \cap \varphi_\alpha(U_\alpha \cap U_\beta)) \tag{1}$$

From Definition 2 the set $\varphi_\alpha(U_\alpha \cap U_\beta)$ is open and hence its intersection with the open set W is open. Now $\varphi_\beta \varphi_\alpha^{-1}$ is C^∞ with C^∞ inverse and so certainly a homeomorphism, and it follows that the right hand side of (1) is open. Thus the left hand side $\varphi_\beta(\varphi_\alpha^{-1} W \cap U_\beta)$ is open and by the definition of the topology this means that $\varphi_\alpha^{-1}(W)$ is open, i.e. φ_α is continuous. □

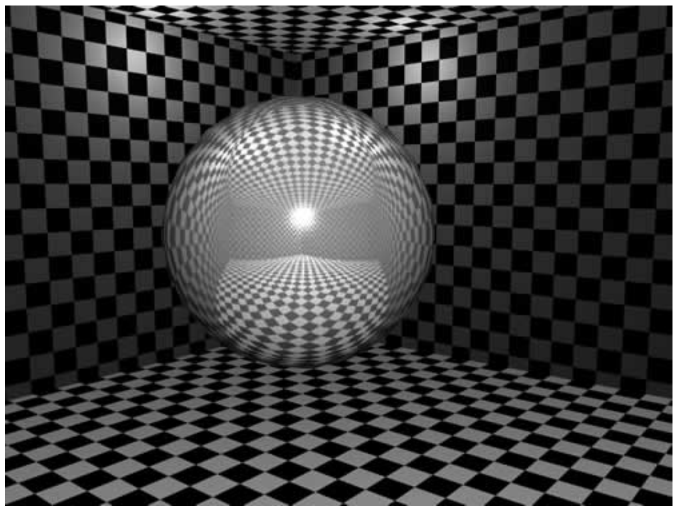
To make any reasonable further progress, we have to make two assumptions about this topology which will hold for the rest of these notes:

- **the manifold topology is Hausdorff**
- **in this topology we have a countable basis of open sets**

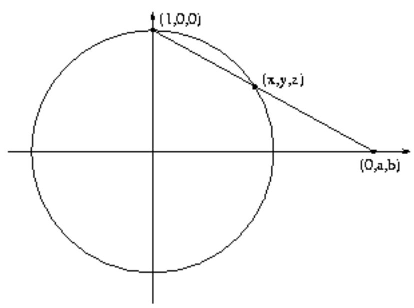
Without these assumptions, manifolds are not even metric spaces, and there is not much analysis that can reasonably be done on them.

2.3 Further examples of manifolds

We need better ways of recognizing manifolds than struggling to find explicit coordinate charts. For example, the sphere is a manifold



and although we can use stereographic projection to get an atlas:



there are other ways. Here is one.

Theorem 2.2 *Let $F : U \to \mathbf{R}^m$ be a C^∞ function on an open set $U \subseteq \mathbf{R}^{n+m}$ and take $c \in \mathbf{R}^m$. Assume that for each $a \in F^{-1}(c)$, the derivative*

$$DF_a : \mathbf{R}^{n+m} \to \mathbf{R}^m$$

is surjective. Then $F^{-1}(c)$ has the structure of an n-dimensional manifold which is Hausdorff and has a countable basis of open sets.

Proof: Recall that the derivative of F at a is the linear map $DF_a : \mathbf{R}^{n+m} \to \mathbf{R}^m$ such that

$$F(a + h) = F(a) + DF_a(h) + R(a, h)$$

where $R(a,h)/\|h\| \to 0$ as $h \to 0$.

If we write $F(x_1,\ldots,x_{n+m}) = (F_1,\ldots,F_m)$ the derivative is the Jacobian matrix

$$\frac{\partial F_i}{\partial x_j}(a) \qquad 1 \le i \le m, 1 \le j \le n+m$$

Now we are given that this is surjective, so the matrix has rank m. Therefore by reordering the coordinates $x_1,\ldots,x_{n+m}$ we may assume that the square matrix

$$\frac{\partial F_i}{\partial x_j}(a) \qquad 1 \le i \le m, 1 \le j \le m$$

is invertible.

Now define

$$G : U \times \mathbf{R}^m \to \mathbf{R}^{n+m}$$

by

$$G(x_1,\ldots,x_{n+m}) = (F_1,\ldots,F_m,x_{m+1},\ldots,x_{n+m}). \tag{2}$$

Then DG_a is invertible.

We now apply the *inverse function theorem* to G, a proof of which is given in the Appendix. It tells us that there is a neighbourhood V of x, and W of $G(x)$ such that $G : V \to W$ is invertible with smooth inverse. Moreover, the formula (2) shows that G maps $V \cap F^{-1}(c)$ to the intersection of W with the copy of $\mathbf{R}^n$ given by $\{x \in \mathbf{R}^{n+m} : x_i = c_i, 1 \le i \le m\}$. This is therefore a coordinate chart φ.

If we take two such charts $\varphi_\alpha, \varphi_\beta$, then $\varphi_\alpha\varphi_\beta^{-1}$ is a map from an open set in $\{x \in \mathbf{R}^{n+m} : x_i = c_1, 1 \le i \le m\}$ to another one which is the restriction of the map $G_\alpha G_\beta^{-1}$ of (an open set in) $\mathbf{R}^{n+m}$ to itself. But this is an invertible C^∞ map and so we have the requisite conditions for an atlas.

Finally, in the induced topology from $\mathbf{R}^{n+m}$, G_α is a homeomorphism, so open sets in the manifold topology are the same as open sets in the induced topology. Since $\mathbf{R}^{n+m}$ is Hausdorff with a countable basis of open sets, so is $F^{-1}(c)$. □

We can now give further examples of manifolds:

Examples: 1. Let

$$S^n = \{x \in \mathbf{R}^{n+1} : \sum_1^{n+1} x_i^2 = 1\}$$

be the unit n-sphere. Define $F : \mathbf{R}^{n+1} \to \mathbf{R}$ by

$$F(x) = \sum_1^{n+1} x_i^2.$$

This is a C^∞ map and

$$DF_a(h) = 2\sum_i a_i h_i$$

is non-zero (and hence surjective in the 1-dimensional case) so long as a is not identically zero. If $F(a) = 1$, then

$$\sum_1^{n+1} a_i^2 = 1 \neq 0$$

so $a \neq 0$ and we can apply Theorem 2.2 and deduce that the sphere is a manifold.

2. Let $O(n)$ be the space of $n \times n$ orthogonal matrices: $AA^T = 1$. Take the vector space M_n of dimension n^2 of all real $n \times n$ matrices and define the function

$$F(A) = AA^T$$

to the vector space of *symmetric* $n \times n$ matrices. This has dimension $n(n+1)/2$. Then $O(n) = F^{-1}(I)$.

Differentiating F we have

$$DF_A(H) = HA^T + AH^T$$

and putting $H = KA$ this is

$$KAA^T + AA^T K^T = K + K^T$$

if $AA^T = I$, i.e. if $A \in F^{-1}(I)$. But given any symmetric matrix S, taking $K = S/2$ shows that DF_I is surjective and so, applying Theorem 2.2 we find that $O(n)$ is a manifold. Its dimension is

$$n^2 - n(n+1)/2 = n(n-1)/2.$$

2.4 Maps between manifolds

We need to know what a smooth map between manifolds is. Here is the definition:

Definition 6 *A map $F : M \to N$ of manifolds is a smooth map if for each point $x \in M$ and chart $(U_\alpha, \varphi_\alpha)$ in M with $x \in U_\alpha$ and chart (V_i, ψ_i) of N with $F(x) \in V_i$, the composite function*

$$\psi_i F \varphi_\alpha^{-1}$$

on $F^{-1}(V_i) \cap U_\alpha$ is a C^∞ function.

Note that it is enough to check that the above holds for one atlas – it will follow from the fact that $\varphi_\alpha \varphi_\beta^{-1}$ is C^∞ that it then holds for all compatible atlases.

Exercise 2.3 *Show that a smooth map is continuous in the manifold topology.*

The natural notion of equivalence between manifolds is the following:

Definition 7 *A diffeomorphism $F : M \to N$ is a smooth map with smooth inverse.*

Example: Take two of our examples above – the quotient group $\mathbf{R}/\mathbf{Z}$ and the 1-sphere, the circle, S^1. We shall show that these are diffeomorphic. First we define a map

$$G : \mathbf{R}/\mathbf{Z} \to S^1$$

by

$$G(x) = (\cos 2\pi x, \sin 2\pi x).$$

This is clearly a bijection. Take $x \in U_0 \subset \mathbf{R}/\mathbf{Z}$ then we can represent the point by $x \in (0, 1)$. Within the range $(0, 1/2)$, $\sin 2\pi x \neq 0$, so with $F = x_1^2 + x_2^2$, we have $\partial F/\partial x_2 \neq 0$. The use of the inverse function theorem in Theorem 2.2 then says that x_1 is a local coordinate for S^1, and in fact on the whole of $(0, 1/2)$ $\cos 2\pi x$ is smooth with smooth inverse. We proceed by taking the other similar open sets to check fully.

3 Tangent vectors and cotangent vectors

3.1 Existence of smooth functions

The most fundamental type of map between manifolds is a smooth map

$$f : M \to \mathbf{R}.$$

We can add these and multiply by constants so they form a vector space $C^\infty(M)$, the space of C^∞ functions on M. In fact, under multiplication it is also a commutative ring. So far, all we can assert is that the constant functions lie in this space, so let's see why there are lots and lots of global C^∞ functions. We shall use bump functions and the Hausdorff property.

First note that the following function of one variable is C^∞:

$$\begin{aligned} f(t) &= e^{-1/t} \quad t > 0 \\ &= 0 \qquad t \leq 0 \end{aligned}$$

Now form

$$g(t) = \frac{f(t)}{f(t) + f(1-t)}$$

so that g is identically 1 when $t \geq 1$ and vanishes if $t \leq 0$. Next write

$$h(t) = g(t+2)g(2-t).$$

This function is completely flat on top.

Finally make an n-dimensional version

$$k(x_1, \ldots, x_n) = h(x_1)h(x_2)\ldots h(x_n).$$

We can rescale the domain of this so that it is zero outside some small ball of radius $2r$ and identically 1 inside the ball of radius r.

We shall use this construction several times later on. For the moment, let M be any manifold and (U, φ_U) a coordinate chart. Choose a function k of the type above whose support (remember $\operatorname{supp} f = \overline{\{x : f(x) \neq 0\}}$) lies in $\varphi_U(U)$ and define

$$f : M \to \mathbf{R}$$

by

$$\begin{aligned} f(x) &= k \circ \varphi_U(x) \quad x \in U \\ &= 0 \qquad x \in M \backslash U. \end{aligned}$$

Is this a smooth function? The answer is yes: clearly $\operatorname{supp} k$ is closed and bounded in $\mathbf{R}^n$ and so compact and since φ_U is a homeomorphism, $\operatorname{supp} f$ is compact. If $y \in M \backslash U$ then y is not in $\operatorname{supp} f$, and if M is *Hausdorff* we can find an open set containing y which does not intersect $\operatorname{supp} f$. Then clearly f is smooth, since it is zero in a neighbourhood of y.

3.2 The derivative of a function

Smooth functions exist in abundance. The question now is: we know what a differentiable function is – so what is its derivative? We need to give some coordinate-independent definition of derivative and this will involve some new concepts. The derivative at a point $a \in M$ will lie in a vector space T_a^* called the cotangent space.

First let's address a simpler question – what does it mean for the derivative to vanish? This is more obviously a coordinate-invariant notion because on a compact manifold any function has a maximum, and in any coordinate system in a neighbourhood of that point, its derivative must vanish. We can check that: if $f : M \to \mathbf{R}$ is smooth then

$$g = f\varphi_\alpha^{-1}$$

is a C^∞ function of $x_1, \ldots x_n$. Suppose its derivative vanishes at $\varphi_U(a)$ and now take a different chart φ_β with $h = f\varphi_\beta^{-1}$. Then

$$g = f\varphi_\alpha^{-1} = f\varphi_\beta^{-1}\varphi_\beta\varphi_\alpha^{-1} = h\varphi_\beta\varphi_\alpha^{-1}.$$

But from the definition of an atlas, $\varphi_\beta\varphi_\alpha^{-1}$ is smooth with smooth inverse, so

$$g(x_1, \ldots, x_n) = h(y_1(x), \ldots, y_n(x))$$

and by the chain rule

$$\frac{\partial g}{\partial x_i} = \sum_j \frac{\partial h}{\partial y_j}(y(a))\frac{\partial y_j}{\partial x_i}(a).$$

Since $y(x)$ is invertible, its Jacobian matrix is invertible, so that $Dg_a = 0$ if and only if $Dh_{y(a)} = 0$. We have checked then that the vanishing of the derivative at a point a is independent of the coordinate chart. We let $Z_a \subset C^\infty(M)$ be the subset of functions whose derivative vanishes at a. Since Df_a is linear in f the subset Z_a is a vector subspace.

Definition 8 *The cotangent space T_a^* at $a \in M$ is the quotient space*

$$T_a^* = C^\infty(M)/Z_a.$$

The derivative of a function f at a is its image in this space and is denoted $(df)_a$.

Here we have simply defined the derivative as all functions modulo those whose derivative vanishes. It's almost a tautology, so to get anywhere we have to prove something about T_a^*. First note that if ψ is a smooth function on a neighbourhood of x, we can multiply it by a bump function to extend it to M and then look at its image in $T_a^* = C^\infty(M)/Z_a$. But its derivative in a coordinate chart around a is independent of the bump function, because all such functions are identically 1 in a neighbourhood of a. Hence we can actually define the derivative at a of smooth functions which are only defined in a neighbourhood of a. In particular we could take the coordinate functions $x_1, \ldots, x_n$. We then have

Proposition 3.1 *Let M be an n-dimensional manifold, then*

- *the cotangent space T_a^* at $a \in M$ is an n-dimensional vector space*
- *if (U, φ) is a coordinate chart around x with coordinates $x_1, \ldots, x_n$, then the elements $(dx_1)_a, \ldots (dx_n)_a$ form a basis for T_a^**
- *if $f \in C^\infty(M)$ and in the coordinate chart, $f\varphi^{-1} = \phi(x_1, \ldots, x_n)$ then*

$$(df)_a = \sum_i \frac{\partial \phi}{\partial x_i}(\varphi(a))(dx_i)_a \tag{3}$$

Proof: If $f \in C^\infty(M)$, with $f\varphi^{-1} = \phi(x_1, \ldots, x_n)$ then

$$f - \sum \frac{\partial \phi}{\partial x_i}(\varphi(a))x_i$$

is a (locally defined) smooth function whose derivative vanishes at a, so

$$(df)_a = \sum \frac{\partial f}{\partial x_i}(\varphi(a))(dx_i)_a$$

and $(dx_1)_a, \ldots (dx_n)_a$ span T_a^*.

If $\sum_i \lambda_i (dx_i)_a = 0$ then $\sum_i \lambda_i x_i$ has vanishing derivative at a and so $\lambda_i = 0$ for all i.
□

Remark: It is rather heavy handed to give two symbols f, ϕ for a function and its representation in a given coordinate system, so often in what follows we shall use just f. Then we can write (3) as

$$df = \sum \frac{\partial f}{\partial x_i} dx_i.$$

With a change of coordinates $(x_1, \ldots, x_n) \to (y_1(x), \ldots, y_n(x))$ the formalism gives

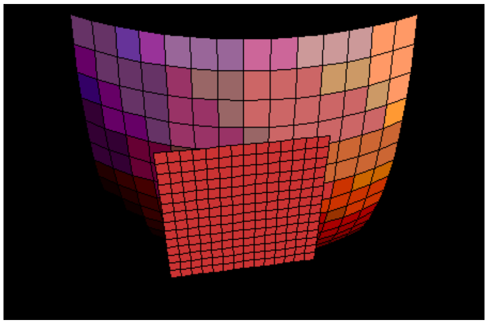
$$df = \sum_j \frac{\partial f}{\partial y_j} dy_j = \sum_{i,j} \frac{\partial f}{\partial y_j} \frac{\partial y_j}{\partial x_i} dx_i.$$

Definition 9 *The tangent space T_a at $a \in M$ is the dual space of the cotangent space T_a^*.*

This is a roundabout way of defining T_a, but since the double dual V^** of a finite dimensional vector space is naturally isomorphic to V the notation is consistent. If $x_1, \ldots, x_n$ is a local coordinate system at a and $(dx_1)_a, \ldots, (dx_n)_a$ the basis of T_a^* defined in (3.1) then the dual basis for the tangent space T_a is denoted

$$\left(\frac{\partial}{\partial x_1}\right)_a, \ldots, \left(\frac{\partial}{\partial x_1}\right)_a.$$

This definition at first sight seems far away from our intuition about the tangent space to a surface in $\mathbf{R}^3$:



The problem arises because our manifold M does not necessarily sit in Euclidean space and we have to define a tangent space intrinsically. The link is provided by the notion of *directional derivative*. If f is a function on a surface in $\mathbf{R}^3$, then for every tangent direction $\mathbf{u}$ at a we can define the derivative of f at a in the direction $\mathbf{u}$, which is a real number: $\mathbf{u} \cdot \nabla f(a)$ or $DF_a(u)$. Imitating this gives the following:

Definition 10 *A tangent vector at a point $a \in M$ is a linear map $X_a : C^\infty(M) \to \mathbf{R}$ such that*

$$X_a(fg) = f(a)X_a g + g(a)X_a f.$$

This is the formal version of the Leibnitz rule for differentiating a product.

Now if $\xi \in T_a$ it lies in the dual space of $T_a^* = C^\infty(M)/Z_a$ and so

$$f \mapsto \xi((df)_a)$$

is a linear map from $C^\infty(M)$ to $\mathbf{R}$. Moreover from (3),

$$d(fg)_a = f(a)(dg)_a + g(a)(df)_a$$

and so

$$X_a(f) = \xi((df)_a)$$

is a tangent vector at a. In fact, any tangent vector is of this form, but the price paid for the nice algebraic definition in (10) which is the usual one in textbooks is that we need a lemma to prove it.

Lemma 3.2 *Let X_a be a tangent vector at a and f a smooth function whose derivative at a vanishes. Then $X_a f = 0$.*

Proof: Use a coordinate system near a. By the fundamental theorem of calculus,

$$\begin{aligned} f(x) - f(a) &= \int_0^1 \frac{\partial}{\partial t} f(a + t(x-a))dt \\ &= \sum_i (x_i - a_i) \int_0^1 \frac{\partial f}{\partial x_i}(a + t(x-a))dt. \end{aligned}$$

If $(df)_a = 0$ then

$$g_i(x) = \int_0^1 \frac{\partial f}{\partial x_i}(a + t(x-a))dt$$

vanishes at $x = a$, as does $x_i - a_i$. Now although these functions are defined locally, using a bump function we can extend them to M, so that

$$f = f(a) + \sum_i g_i h_i \tag{4}$$

where $g_i(a) = h_i(a) = 0$.

By the Leibnitz rule

$$X_a(1) = X_a(1.1) = 2X_a(1)$$

which shows that X_a annihilates constant functions. Applying the rule to (4)

$$X_a(f) = X_a(\sum_i g_i h_i) = \sum_i (g_i(a) X_a h_i + h_i(a) X_a g_i) = 0.$$

This means that $X_a : C^\infty(M) \to \mathbf{R}$ annihilates Z_a and is well defined on $T_a^* = C^\infty(M)/Z_a$ and so $X_a \in T_a$. $\square$

The vectors in the tangent space are therefore the tangent vectors as defined by (10). Locally, in coordinates, we can write

$$X_a = \sum_i^n c_i \left(\frac{\partial}{\partial x_i}\right)_a$$

and then

$$X_a(f) = \sum_i c_i \frac{\partial f}{\partial x_i}(a) \qquad (5)$$

3.3 Derivatives of smooth maps

Suppose $F : M \to N$ is a smooth map and $f \in C^\infty(N)$. Then $f \circ F$ is a smooth function on M.

Definition 11 *The derivative at $a \in M$ of the smooth map $F : M \to N$ is the homomorphism of tangent spaces*

$$DF_a : T_a M \to T_{F(a)} N$$

defined by

$$DF_a(X_a)(f) = X_a(f \circ F).$$

This is an abstract, coordinate-free definition. Concretely, we can use (5) to see that

$$\begin{aligned} DF_a \left(\frac{\partial}{\partial x_i}\right)_a (f) &= \frac{\partial}{\partial x_i}(f \circ F)(a) \\ &= \sum_j \frac{\partial F_j}{\partial x_i}(a) \frac{\partial f}{\partial y_j}(F(a)) = \sum_j \frac{\partial F_j}{\partial x_i}(a) \left(\frac{\partial}{\partial y_j}\right)_{F(a)} f \end{aligned}$$

Thus the derivative of F is an invariant way of defining the Jacobian matrix.

With this definition we can give a generalization of Theorem 2.2 – the proof is virtually the same and is omitted.

Theorem 3.3 *Let $F : M \to N$ be a smooth map and $c \in N$ be such that at each point $a \in F^{-1}(c)$ the derivative DF_a is surjective. Then $F^{-1}(c)$ is a smooth manifold of dimension* $\dim M - \dim N$.

In the course of the proof, it is easy to see that the manifold structure on $F^{-1}(c)$ makes the inclusion

$$\iota : F^{-1}(c) \subset M$$

a smooth map, whose derivative is injective and maps isomorphically to the kernel of DF. So when we construct a manifold like this, its tangent space at a is

$$T_a \cong \operatorname{Ker} DF_a.$$

This helps to understand tangent spaces for the case where F is defined on $\mathbf{R}^n$:

Examples:

1. The sphere S^n is $F^{-1}(1)$ where $F : \mathbf{R}^{n+1} \to \mathbf{R}$ is given by

$$F(x) = \sum_i x_i^2.$$

So here

$$DF_a(x) = 2 \sum_i x_i a_i$$

and the kernel of DF_a consists of the vectors orthogonal to a, which is our usual vision of the tangent space to a sphere.

2. The orthogonal matrices $O(n)$ are given by $F^{-1}(I)$ where $F(A) = AA^T$. At $A = I$, the derivative is

$$DF_I(H) = H + H^T$$

so the tangent space to $O(n)$ at the identity matrix is $\operatorname{Ker} DF_I$, the space of skew-symmetric matrices $H = -H^T$.

The examples above are of manifolds $F^{-1}(c)$ sitting inside M and are examples of *submanifolds*. Here we shall adopt the following definition of a submanifold, which is often called an *embedded submanifold*:

Definition 12 *A manifold M is a submanifold of N if there is an inclusion map*

$$\iota : M \to N$$

such that

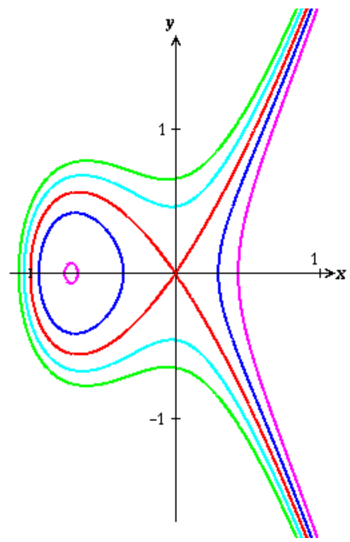
- *ι is smooth*

- *$D\iota_x$ is injective for each $x \in M$*
- *the manifold topology of M is the induced topology from N*

Remark: The topological assumption avoids a situation like this:

$$\iota(t) = (t^2 - 1, t(t^2 - 1)) \in \mathbf{R}^2$$

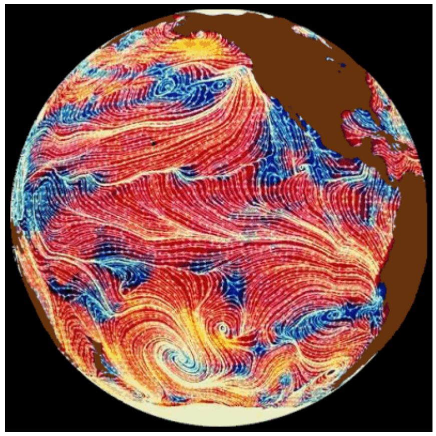
for $t \in (-1, 1)$. This is smooth, injective with injective derivative, but any open set in $\mathbf{R}^2$ containing 0 intersects both ends of the interval. The curve is the left hand loop of the singular cubic: $y^2 = x^2(x + 1)$.



4 Vector fields

4.1 The tangent bundle

Think of the wind velocity at each point of the earth.



This is an example of a vector field on the 2-sphere S^2. Since the sphere sits inside $\mathbf{R}^3$, this is just a smooth map $X : S^2 \to \mathbf{R}^3$ such that $X(x)$ is tangential to the sphere at x.

Our problem now is to define a vector field intrinsically on a general manifold M, without reference to any ambient space. We know what a tangent vector at $a \in M$ is – a vector in T_a – but we want to describe a smoothly varying family of these. To do this we need to fit together all the tangent spaces as a ranges over M into a single manifold called the tangent bundle. We have n degrees of freedom for $a \in M$ and n for each tangent space T_a so we expect to have a $2n$-dimensional manifold. So the *set* to consider is

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

the disjoint union of all the tangent spaces.

First let (U, φ_U) be a coordinate chart for M. Then for $x \in U$ the tangent vectors

$$\left(\frac{\partial}{\partial x_1}\right)_x, \ldots, \left(\frac{\partial}{\partial x_n}\right)_x$$

provide a basis for each T_x. So we have a bijection

$$\psi_U : U \times \mathbf{R}^n \to \bigcup_{x \in U} T_x$$

defined by

$$\psi_U(x, y_1, \ldots, y_n) = \sum_1^n y_i \left(\frac{\partial}{\partial x_i}\right)_x .$$

Thus

$$\Phi_U = (\varphi_U, id) \circ \psi^{-1} : \bigcup_{x \in U} T_x \to \varphi_U(U) \times \mathbf{R}^n$$

is a coordinate chart for

$$V = \bigcup_{x \in U} T_x.$$

Given U_α, U_β coordinate charts on M, clearly

$$\Phi_\alpha(V_\alpha \cap V_\beta) = \varphi_\alpha(U_\alpha \cap U_\beta) \times \mathbf{R}^n$$

which is open in $\mathbf{R}^{2n}$. Also, if $(x_1, \ldots, x_n)$ are coordinates on U_α and $(\tilde{x}_1, \ldots, \tilde{x}_n)$ on U_β then

$$\left(\frac{\partial}{\partial x_i}\right)_x = \sum_j \frac{\partial \tilde{x}_j}{\partial x_i} \left(\frac{\partial}{\partial \tilde{x}_j}\right)_x$$

the dual of (3). It follows that

$$\Phi_\beta \Phi_\alpha^{-1}(x_1, \ldots, x_n, y_1 \ldots, y_n) = (\tilde{x}_1, \ldots, \tilde{x}_n, \sum_j \frac{\partial \tilde{x}_1}{\partial x_i} y_i, \ldots, \sum_i \frac{\partial \tilde{x}_n}{\partial x_i} y_i).$$

and since the Jacobian matrix is smooth in x, linear in y and invertible, $\Phi_\beta \Phi_\alpha^{-1}$ is smooth with smooth inverse and so (V_α, Φ_α) defines an atlas on TM.

Definition 13 *The tangent bundle of a manifold M is the $2n$-dimensional differentiable structure on TM defined by the above atlas.*

The construction brings out a number of properties. First of all the projection map

$$p : TM \to M$$

which assigns to $X_a \in T_a M$ the point a is smooth with surjective derivative, because in our local coordinates it is defined by

$$p(x_1, \ldots, x_n, y_1, \ldots, y_n) = (x_1, \ldots, x_n).$$

The inverse image $p^{-1}(a)$ is the vector space T_a and is called a *fibre* of the projection. Finally, TM is Hausdorff because if X_a, X_b lie in different fibres, since M is Hausdorff we can separate $a, b \in M$ by open sets U, U' and then the open sets $p^{-1}(U), p^{-1}(U')$ separate X_a, X_b in TM. If X_a, Y_a are in the same tangent space then they lie in a coordinate neighbourhood which is homeomorphic to an open set of $\mathbf{R}^{2n}$ and so can be separated there. Since M has a countable basis of open sets and $\mathbf{R}^n$ does, it is easy to see that TM also has a countable basis.

We can now define a vector field:

Definition 14 *A* *vector field* *on a manifold is a smooth map*

$$X : M \to TM$$

such that

$$p \circ X = id_M.$$

This is a clear global definition. What does it mean? We just have to spell things out in local coordinates. Since $p \circ X = id_M$,

$$X(x_1, \ldots, x_n) = (x_1, \ldots, x_n, y_1(x), \ldots, y_n(x))$$

where $y_i(x)$ are smooth functions. Thus the tangent vector $X(x)$ is given by

$$X(x) = \sum_i y_i(x) \left(\frac{\partial}{\partial x_i}\right)_x$$

which is a smoothly varying field of tangent vectors.

Remark: We shall meet other manifolds Q with projections $p : Q \to M$ and the general terminology is that a smooth map $s : M \to Q$ for which $p \circ s = id_M$ is called a *section*. When $Q = TM$ is the tangent bundle we always have the *zero section* given by the vector field $X = 0$. Using a bump function ψ we can easily construct other vector fields by taking a coordinate system, writing

$$X(x) = \sum_i y_i(x) \left(\frac{\partial}{\partial x_i}\right)_x$$

multiplying by ψ and extending.

Remark: Clearly we can do a similar construction using the cotangent spaces T_a^* instead of the tangent spaces T_a, and using the basis

$$(dx_1)_x, \ldots, (dx_n)_x$$

instead of the dual basis

$$\left(\frac{\partial}{\partial x_1}\right)_x, \ldots, \left(\frac{\partial}{\partial x_1}\right)_x.$$

This way we form the *cotangent bundle* T^*M. The derivative of a function f is then a map $df : M \to TM$ satisfying $p \circ df = id_M$, though not every such map of this form is a derivative. The tangent bundle and cotangent bundle are examples of *vector bundles*.

Perhaps we should say here that the tangent bundle and cotangent bundle are examples of *vector bundles*. Here is the general definition:

Definition 15 *A real vector bundle of rank m on a manifold M is a manifold E with a smooth projection $p : E \to M$ such that*

- *each fibre $p^{-1}(x)$ has the structure of an m-dimensional real vector space*
- *each point $x \in M$ has a neighbourhood U and a diffeomorphism*

  $$\psi_U : p^{-1}(U) \cong U \times \mathbf{R}^m$$

  such that ψ_U maps the vector space $p^{-1}(x)$ isomorphically to the vector space $\{x\} \times \mathbf{R}^m$
- *on the intersection $U \cap V$*

  $$\psi_U^{-1}\psi_V : U \cap V \times \mathbf{R}^m \to U \cap V \times \mathbf{R}^m$$

  is of the form

  $$(x, v) \mapsto (x, g_{UV}(x)v)$$

  where $g_{UV}(x)$ is a smooth function on $U \cap V$ with values in the space of invertible $m \times m$ matrices.

For the tangent and cotangent bundle, g_{UV} is the Jacobian matrix of a change of coordinates or its inverse transpose.

4.2 Vector fields as derivations

The algebraic definition of tangent vector in Definition 10 shows that a vector field X maps a C^∞ function to a function on M:

$$X(f)(x) = X_x(f)$$

and the local expression for X means that

$$X(f)(x) = \sum_i y_i(x) \left(\frac{\partial}{\partial x_i}\right)_x (f) = \sum_i y_i(x) \frac{\partial f}{\partial x_i}(x).$$

Since the $y_i(x)$ are smooth, $X(f)$ is again smooth and satisfies the Leibnitz property

$$X(fg) = f(Xg) + g(Xf).$$

In fact, any linear transformation with this property (called a derivation of the algebra $C^\infty(M)$) is a vector field:

Proposition 4.1 *Let $X : C^\infty(M) \to C^\infty(M)$ be a linear map which satisfies*

$$X(fg) = f(Xg) + g(Xf).$$

Then X is a vector field.

Proof: For each $a \in M$, $X_a(f) = X(f)(a)$ satisfies the conditions for a tangent vector at a, so X defines a map $X : M \to TM$ with $p \circ X = id_M$, and so locally can be written as

$$X_x = \sum_i y_i(x) \left(\frac{\partial}{\partial x_i} \right)_x .$$

We just need to check that the $y_i(x)$ are smooth, and for this it suffices to apply X to a coordinate function x_i extended by using a bump function in a coordinate neighbourhood. We get

$$X x_i = y_i(x)$$

and since by assumption X maps smooth functions to smooth functions, this is smooth. $\square$

The characterization of vector fields given by Proposition 4.1 immediately leads to a way of combining two vector fields X, Y to get another. Consider both X and Y as linear maps from $C^\infty(M)$ to itself and compose them. Then

$$\begin{aligned} XY(fg) &= X(f(Yg) + g(Yf)) = (Xf)(Yg) + f(XYg) + (Xg)(Yf) + g(XYf) \\ YX(fg) &= Y(f(Xg) + g(Xf)) = (Yf)(Xg) + f(YXg) + (Yg)(Xf) + g(YXf) \end{aligned}$$

and subtracting and writing $[X, Y] = XY - YX$ we have

$$[X, Y](fg) = f([X, Y]g) + g([X, Y]f)$$

which from Proposition 4.1 means that $[X, Y]$ is a vector field.

Definition 16 *The Lie bracket of two vector fields X, Y is the vector field $[X, Y]$.*

Example: If $M = \mathbf{R}$ then $X = f d/dx, Y = g d/dx$ and so

$$[X, Y] = (fg' - gf') \frac{d}{dx}.$$

We shall later see that there is a geometrical origin for the Lie derivative.

4.3 One-parameter groups of diffeomorphisms

Think of wind velocity (assuming it is constant in time) on the surface of the earth as a vector field on the sphere S^2. There is another interpretation we can make. A particle at position $x \in S^2$ moves after time t seconds to a position $\varphi_t(x) \in S^2$. After a further s seconds it is at

$$\varphi_{t+s}(x) = \varphi_s(\varphi_t(x)).$$

What we get this way is a homomorphism of groups: from the additive group $\mathbf{R}$ to the group of diffeomorphisms of S^2 under the operation of composition. The technical definition is the following:

Definition 17 *A one-parameter group of diffeomorphisms of a manifold M is a smooth map*

$$\varphi : M \times \mathbf{R} \to M$$

such that (writing $\varphi_t(x) = \varphi(x,t)$)

- $\varphi_t : M \to M$ *is a diffeomorphism*
- $\varphi_0 = id$
- $\varphi_{s+t} = \varphi_s \circ \varphi_t$.

We shall show that vector fields generate one-parameter groups of diffeomorphisms, but only under certain hypotheses. If instead of the whole surface of the earth our manifold is just the interior of the UK and the wind is blowing East-West, clearly after however short a time, some particles will be blown offshore, so we cannot hope for $\varphi_t(x)$ that works for all x and t. The fact that the earth is compact is one reason why it works there, and this is one of the results below. The idea, nevertheless, works locally and is a useful way of understanding vector fields as "infinitesimal diffeomorphisms" rather than as abstract derivations of functions.

To make the link with vector fields, suppose φ_t is a one-parameter group of diffeomorphisms and f a smooth function. Then

$$f(\varphi_t(a))$$

is a smooth function of t and we write

$$\frac{\partial}{\partial t} f(\varphi_t(a))|_{t=0} = X_a(f).$$

It is straightforward to see that, since $\varphi_0(a) = a$ the Leibnitz rule holds and this is a tangent vector at a, and so as $a = x$ varies we have a vector field. In local coordinates we have

$$\varphi_t(x_1, \ldots, x_n) = (y_1(x,t), \ldots, y_n(x,t))$$

and

$$\begin{aligned}\frac{\partial}{\partial t} f(y_1, \ldots, y_n) &= \sum_i \frac{\partial f}{\partial y_i}(y) \frac{\partial y_i}{\partial t}(x)|_{t=0} \\ &= \sum_i c_i(x) \frac{\partial f}{\partial x_i}(x)\end{aligned}$$

which yields the vector field

$$X = \sum_i c_i(x) \frac{\partial}{\partial x_i}.$$

We now want to reverse this: go from the vector field to the diffeomorphism. The first point is to track that "trajectory" of a single particle.

Definition 18 *An integral curve of a vector field X is a smooth map $\varphi : (\alpha, \beta) \subset \mathbf{R} \to M$ such that*

$$D\varphi_t \left(\frac{d}{dt} \right) = X_{\varphi(t)}.$$

Example: Suppose $M = \mathbf{R}^2$ with coordinates (x, y) and $X = \partial/\partial x$. The derivative $D\varphi$ of the smooth function $\varphi(t) = (x(t), y(t))$ is

$$D\varphi \left(\frac{d}{dt} \right) = \frac{dx}{dt} \frac{\partial}{\partial x} + \frac{dy}{dt} \frac{\partial}{\partial y}$$

so the equation for an integral curve of X is

$$\begin{aligned}\frac{dx}{dt} &= 1 \\ \frac{dy}{dt} &= 0\end{aligned}$$

which gives

$$\varphi(t) = (t + a_1, a_2).$$

In our wind analogy, the particle at (a_1, a_2) is transported to $(t + a_1, a_2)$.

In general we have:

Theorem 4.2 *Given a vector field X on a manifold M and $a \in M$ there exists a maximal integral curve of X through a.*

By "maximal" we mean that the interval (α, β) is maximal – as we saw above it may not be the whole of the real numbers.

Proof: First consider a coordinate chart $(U_\alpha, \varphi_\alpha)$ around a then if

$$X = \sum_i c_i(x) \frac{\partial}{\partial x_i}$$

the equation

$$D\varphi_t \left(\frac{d}{dt} \right) = X_{\varphi(t)}$$

can be written as the system of ordinary differential equations

$$\frac{dx_i}{dt} = c_i(x_1, \ldots, x_n).$$

The existence and uniqueness theorem for ODE's (see Appendix) asserts that there is some interval on which there is a unique solution with initial condition

$$(x_1(0), \ldots, x_n(0)) = \varphi_\alpha(a).$$

Suppose $\varphi : (\alpha, \beta) \to M$ is *any* integral curve with $\varphi(0) = a$. For each $x \in (\alpha, \beta)$ the subset $\varphi([0, x]) \subset M$ is compact, so it can be covered by a finite number of coordinate charts, in each of which we can apply the existence and uniqueness theorem to intervals $[0, \alpha_1], [\alpha_1, \alpha_2], \ldots, [\alpha_n, x]$. Uniqueness implies that these local solutions agree with φ on any subinterval containing 0.

We then take the maximal open interval on which we can define φ. □

To find the one-parameter group of diffeomorphisms we now let $a \in M$ vary. In the example above, the integral curve through (a_1, a_2) was $t \mapsto (t + a_1, a_2)$ and this defines the group of diffeomorphisms

$$\varphi_t(x_1, x_2) = (t + x_1, x_2).$$

Theorem 4.3 *Let X be a vector field on a manifold M and for $(t, x) \in \mathbf{R} \times M$, let $\varphi_(x, t) = \varphi_t(x)$ be the maximal integral curve of X through x. Then*

- *the map* $(t, x) \mapsto \varphi_t(x)$ *is smooth*
- $\varphi_t \circ \varphi_s = \varphi_{t+s}$ *wherever the maps are defined*
- *if* M *is compact, then* $\varphi_t(x)$ *is defined on* $\mathbf{R} \times M$ *and gives a one-parameter group of diffeomorphisms.*

Proof: The previous theorem tells us that for each $a \in M$ we have an open interval $(\alpha(a), \beta(a))$ on which the maximal integral curve is defined. The local existence theorem also gives us that there is a solution for initial conditions in a neighbourhood of a so the set

$$\{(t, x) \in \mathbf{R} \times M : t \in (\alpha(x), \beta(x))\}$$

is open. This is the set on which $\varphi_t(x)$ is maximally defined.

The theorem (see Appendix) on smooth dependence on initial conditions tells us that $(t, x) \mapsto \varphi_t(x)$ is smooth.

Consider $\varphi_t \circ \varphi_s(x)$. If we fix s and vary t, then this is the unique integral curve of X through $\varphi_s(x)$. But $\varphi_{t+s}(x)$ is an integral curve which at $t = 0$ passes through $\varphi_s(x)$. By uniqueness they must agree so that $\varphi_t \circ \varphi_s = \varphi_{t+s}$. (Note that $\varphi_t \circ \varphi_{-t} = id$ shows that we have a diffeomorphism wherever it is defined).

Now consider the case where M is compact. For each $x \in M$, we have an open interval $(\alpha(x), \beta(x))$ containing 0 and an open set $U_x \subseteq M$ on which $\varphi_t(x)$ is defined. Cover M by $\{U_x\}_{x \in M}$ and take a finite subcovering $U_{x_1}, \ldots, U_{x_N}$, and set

$$I = \bigcap_1^N (\alpha(x_i), \beta(x_i))$$

which is an open interval containing 0. By construction, for $t \in I$ we get

$$\varphi_t : I \times M \to M$$

which defines an integral curve (though not necessarily maximal) through each point $x \in M$ and with $\varphi_0(x) = x$. We need to extend to all real values of t.

If $s, t \in \mathbf{R}$, choose n such that $(|s| + |t|)/n \in I$ and define (where multiplication is composition)

$$\varphi_t = (\varphi_{t/n})^n, \qquad \varphi_s = (\varphi_{s/n})^n.$$

Now because $t/n, s/n$ and $(s + t)/n$ lie in I we have

$$\varphi_{t/n}\varphi_{s/n} = \varphi_{(s+t)/n} = \varphi_{s/n}\varphi_{t/n}$$

and so because $\varphi_{t/n}$ and $\varphi_{s/n}$ commute, we also have

$$\begin{aligned}\varphi_t\varphi_s &= (\varphi_{t/n})^n(\varphi_{s/n})^n \\ &= (\varphi_{(s+t)/n})^n \\ &= \varphi_{s+t}\end{aligned}$$

which completes the proof. □

4.4 The Lie bracket revisited

All the objects we shall consider will have the property that they can be transformed naturally by a diffeomorphism, and the link between vector fields and diffeomorphisms we have just observed provides an "infinitesimal' version of this.

Given a diffeomorphism $F : M \to M$ and a smooth function f we get the transformed function $f \circ F$. When $F = \varphi_t$, generated according to the theorems above by a vector field X, we then saw that

$$\frac{\partial}{\partial t} f(\varphi_t)|_{t=0} = X(f).$$

So: *the natural action of diffeomorphisms on functions specializes through one-parameter groups to the derivation of a function by a vector field.*

Now suppose Y is a vector field, considered as a map $Y : M \to TM$. With a diffeomorphism $F : M \to M$, its derivative $DF_x : T_x \to T_{F(x)}$ gives

$$DF_x(Y_x) \in T_{F(x)}.$$

This defines a new vector field $\tilde{Y}$ by

$$\tilde{Y}_{F(x)} = DF_x(Y_x) \tag{6}$$

Thus for a function f,

$$(\tilde{Y})(f \circ F) = (Yf) \circ F \tag{7}$$

Now if $F = \varphi_t$ for a one-parameter group, we have $\tilde{Y}_t$ and we can differentiate to get

$$\dot{Y} = \frac{\partial}{\partial t}\tilde{Y}_t\bigg|_{t=0}$$

From (7) this gives

$$\dot{Y}f + Y(Xf) = XYf$$

so that $\dot{Y} = XY - YX$ is the *Lie derivative* defined above. Thus *the natural action of diffeomorphisms on vector fields specializes through one-parameter groups to the Lie bracket* $[X, Y]$.