15 Differential geometry

For much of physics and mathematics the concept of a continuous map, provided by topology, is not sufficient. What is often required is a notion of *differentiable* or *smooth* maps between spaces. For this, our spaces will need a structure something like that of a surface in Euclidean space $\mathbb{R}^n$. The key ingredient is the concept of a *differentiable manifold*, which can be thought of as topological space that is 'locally Euclidean' at every point. *Differential geometry* is the area of mathematics dealing with these structures. Of the many excellent books on the subject, the reader is referred in particular to [1–14].

Think of the surface of the Earth. Since it is a sphere, it is neither metrically nor topologically identical with the Euclidean plane $\mathbb{R}^2$. A typical atlas of the world consists of separate pages called *charts*, each representing different regions of the Earth. This representation is not metrically correct since the curved surface of the Earth must be flattened out to conform with a sheet of paper, but it is at least smoothly continuous. Each chart has regions where it connects with other charts – a part of France may find itself on a map of Germany, for example – and the correspondence between the charts in the overlapping regions should be continuous and smooth. Some charts may even find themselves entirely inside others; for example, a map of Italy will reappear on a separate page devoted entirely to Europe. Ideally, the entire surface of the Earth should be covered by the different charts of the atlas, although this may not strictly be the case for a real atlas, since the north and south poles are not always properly covered by some chart. We have here the archetype of a differentiable manifold.

Points of $\mathbb{R}^n$ will usually be denoted from now on by superscripted coordinates, $\mathbf{x} = (x^1, x^2, \dots, x^n)$. In Chapter 12 we defined a function $f : \mathbb{R}^n \to \mathbb{R}$ to be C^r if all its partial derivatives

$$\frac{\partial^s f(x^1, x^2, \dots, x^n)}{\partial x^{i_1}\, \partial x^{i_2} \dots \partial x^{i_s}}$$

exist and are continuous for $s = 1, 2, \dots, r$. A C^0 function is simply a continuous function, while a C^∞ function is one that is C^r for all values of $r = 0, 1, 2, \dots$; such a function will generally be referred to simply as a **differentiable function**. A differentiable function need not be *analytic* (expandable as a power series in a neighbourhood of any point), as illustrated by the function

$$f(x) = \begin{cases} 0 & \text{if } x \le 0, \\ \mathrm{e}^{-1/x^2} & \text{if } x > 0, \end{cases}$$

which is differentiable but not analytic at $x = 0$ since its power series would have all coefficients zero at $x = 0$.

A map $\phi : \mathbb{R}^n \to \mathbb{R}^m$ is said to be C^r if, when expressed in coordinates

$$y^i = \phi_i(x^1, x^2, \ldots, x^n) \quad \text{where} \quad \phi_i = \mathrm{pr}_i \circ \phi \quad (i = 1, 2, \ldots, m),$$

each of the real-valued functions $\phi_i : \mathbb{R}^n \to \mathbb{R}$ is C^r. Similarly, the notion of differentiable and analytic functions can be extended to maps between Euclidean spaces of arbitrary dimensions.

15.1 Differentiable manifolds

A **locally Euclidean space** or **topological manifold** M of **dimension** n is a Hausdorff topological space M in which every point x has a neighbourhood homeomorphic to an open subset of $\mathbb{R}^n$. If p is any point of M then a **(coordinate) chart at** p is a pair (U, ϕ) where U is an open subset of M, called the **domain** of the chart and $\phi : U \to \phi(U) \subset \mathbb{R}^n$ is a homeomorphism between U and its image $\phi(U)$. The image $\phi(U)$ is an open subset of $\mathbb{R}^n$, given the relative topology in $\mathbb{R}^n$. It is also common to call U a **coordinate neighbourhood** of p and ϕ a **coordinate map**. The functions $x^i = \mathrm{pr}_i \circ \phi : U \to \mathbb{R} \quad (i = 1, \ldots, n)$, where $\mathrm{pr}_i : \mathbb{R}^n \to R$ are the standard projection maps, are known as the **coordinate functions** determined by this chart, and the real numbers $x^i(p)$ are called the **coordinates** of p in this chart (see Fig. 15.1). Sometimes, when we wish to emphasize the symbols to be used for the coordinate functions, we denote the chart by $(U, \phi; x^i)$, or simply $(U; x^i)$. Occasionally the term **coordinate system at** p is used for a chart whose domain U covers p. The use of superscripts rather than subscripts for coordinate functions is not universal, but its advantages will become apparent as the tensor formalism on manifolds is developed.

For any pair of coordinate charts $(U, \phi; x^i)$ and $(U', \phi'; x'^j)$ such that $U \cap U' \neq \emptyset$, define the **transition functions**

$$\phi' \circ \phi^{-1} : \phi(U \cap U') \to \phi'(U \cap U'),$$
$$\phi \circ \phi'^{-1} : \phi'(U \cap U') \to \phi(U \cap U'),$$

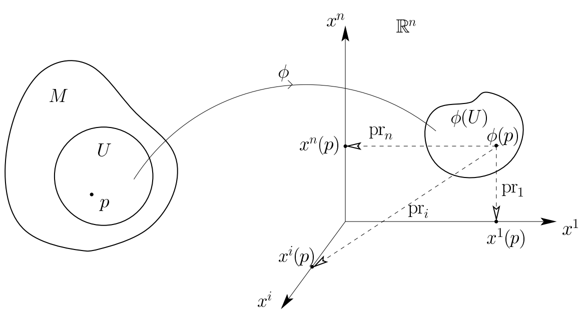

Figure 15.1 Chart at a point p

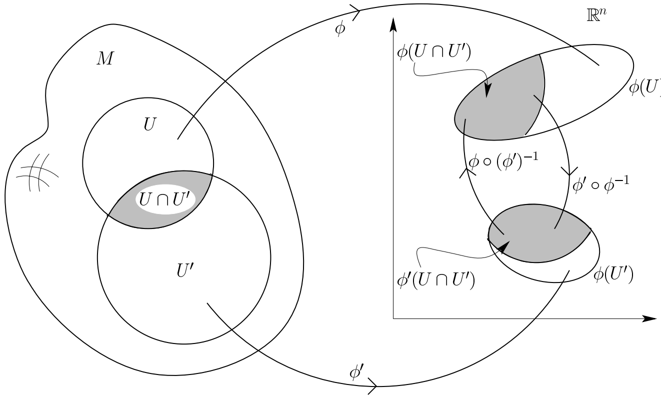

Figure 15.2 Transition functions on compatible charts

which are depicted in Fig. 15.2. The transition functions are often written

$$x'^j = x'^j(x^1, x^2, \dots, x^n) \quad \text{and} \quad x^i = x^i(x'^1, x'^2, \dots, x'^n) \quad (i, j = 1, \dots, n) \tag{15.1}$$

which is an abbreviated form of the awkward, but technically correct,

$$x'^j(p) = \mathrm{pr}_i \circ \phi' \circ \phi^{-1}(x^1(p), x^2(p), \dots, x^n(p)),$$
$$x^i(p) = \mathrm{pr}_i \circ \phi \circ \phi'^{-1}(x'^1(p), x'^2(p), \dots, x'^n(p)).$$

The two charts are said to be C^r**-compatible** where r is a non-negative integer or ∞, if all the functions in (15.1) are C^r. For convenience we will generally assume that the charts are C^∞-compatible.

An **atlas** on M is a family of charts $\mathcal{A} = \{(U_\alpha, \phi_\alpha) \mid \alpha \in A\}$ such that the coordinate neighbourhoods U_α cover M, and any pair of charts from the family are C^∞-compatible. If $\mathcal{A}$ and $\mathcal{A}'$ are two atlases on M then so is their union $\mathcal{A} \cup \mathcal{A}'$.

Exercise: Prove this statement. [*Hint*: A differentiable function of a differentiable function is always differentiable.]

Any atlas $\mathcal{A}$ may thus be extended to a *maximal atlas* by adding to it all charts that are C^∞-compatible with the charts of $\mathcal{A}$. This maximal atlas is called a **differentiable structure** on M. A pair $(M, \mathcal{A})$, where M, is an n-dimensional topological manifold and $\mathcal{A}$ is a differentiable structure on M, is called a **differentiable manifold**; it is usually just denoted M.

The Jacobian matrix $\mathsf{J} = [\partial x'^k / \partial x^j]$ is non-singular since its inverse is $\mathsf{J}^{-1} = [\partial x^i / \partial x'^k]$,

$$\mathsf{J}^{-1}\mathsf{J} = \left[\frac{\partial x^i}{\partial x'^k}\frac{\partial x'^k}{\partial x^j}\right] = \left[\frac{\partial x^i}{\partial x^j}\right] = \left[\delta^i_j\right] = \mathsf{I}.$$

Similarly $\mathsf{J}\mathsf{J}^{-1} = \mathsf{I}$. Hence the Jacobian determinant is non-vanishing, $\det[\partial x'^j / \partial x^i] \neq 0$. We are making a return here and in the rest of this book to the summation convention of earlier chapters.

Example 15.1 Euclidean space $\mathbb{R}^n$ is trivially a manifold, since the single chart ($U = \mathbb{R}^n$, $\phi = \mathrm{id}$) covers it and generates a unique atlas consisting of all charts that are compatible with it. For example, in $\mathbb{R}^2$ it is permissible to use polar coordinates (r, θ) defined by

$$x = r\cos\theta, \qquad y = r\sin\theta,$$

which are compatible with (x, y) on the open set $U = \mathbb{R}^2 - \{(x, y) \mid x \geq 0,\ y = 0\}$. The inverse transformation is

$$r = \sqrt{x^2 + y^2}, \qquad \theta = \begin{cases} \arctan y/x & \text{if } y > 0, \\ \pi & \text{if } y = 0,\ x < 0, \\ \pi + \arctan y/x & \text{if } y < 0. \end{cases}$$

The image set $\phi(U)$ in the (r, θ)-plane is a semi-infinite open strip $r > 0,\ 0 < \theta < 2\pi$.

Example 15.2 Any open region U of $\mathbb{R}^n$ is a differentiable manifold formed by giving it the relative topology and the differentiable structure generated by the single chart $(U, \mathrm{id}_U : U \to \mathbb{R}^n)$. Every chart on U is the restriction of a coordinate neighbourhood and coordinate map on $\mathbb{R}^n$ to the open region U and can be written $(U \cap V, \psi\big|_{U\cap V})$ where (V, ψ) is a chart on $\mathbb{R}^n$. Such a manifold is called an **open submanifold** of $\mathbb{R}^n$.

Exercise: Describe the open region of $\mathbb{R}^3$ and the image set in the $(r,\ \theta,\ \phi)$ on which spherical polar coordinates are defined,

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta. \tag{15.2}$$

Example 15.3 The unit circle $S^1 \subset \mathbb{R}^2$, defined by the equation $x^2 + y^2 = 1$, is a one-dimensional manifold. The coordinate x can be used on either the upper semicircle $y > 0$ or the lower semicircle $y < 0$, but not on all of S^1. Alternatively, setting $r = 1$ in polar coordinates as defined in Example 15.1, a possible chart is $(U, \phi; \theta)$ where $U = S^1 - \{(1, 0)\}$ and $\phi : U \to \mathbb{R}$ is defined by $\phi((x, y)) = \theta$. The image set $\phi(U)$ is the open interval $(0, 2\pi) \subset \mathbb{R}$. These charts are clearly compatible with each other. S^1 is the only one-dimensional manifold that is not homeomorphic to the real line $\mathbb{R}$.

Example 15.4 The 2-sphere S^2 defined as the subset of points (x, y, z) of $\mathbb{R}^3$ satisfying

$$x^2 + y^2 + z^2 = 1$$

is a two-dimensional differentiable manifold. Some possible charts on S^2 are:

(i) Rectangular coordinates (x, y), defined on the upper and lower hemisphere, $z > 0$ and $z < 0$, separately. These two charts are non-intersecting and do not cover the sphere since points on the central plane $z = 0$ are omitted.

(ii) Stereographic projection from the north pole, Eqs. (10.1) and (10.2), defines a chart $(S^2 - \{(0, 0, 1)\}, \mathrm{St}_N)$ where $\mathrm{St}_N : (x, y, z) \mapsto (X, Y)$ is given by

$$X = \frac{x}{1 - z}, \qquad Y = \frac{y}{1 - z}.$$

These coordinates are not defined on the sphere's north pole $N = (0, 0, 1)$, but a similar projection St_S from the south pole $S = (0, 0, -1)$ will cover N,

$$X' = \frac{x}{1+z}, \qquad Y' = \frac{y}{1+z}.$$

Both of these charts are evidently compatible with the rectangular coordinate charts (i) and therefore with each other in their region of overlap.

(iii) Spherical polar coordinates (θ, ϕ) defined by seting $r = 1$ in Eq. (15.2). Simple algebra shows that these are related to the stereographic coordinates (ii) by

$$X = \cot \tfrac{1}{2}\theta \, \cos\phi, \qquad Y = \cot \tfrac{1}{2}\theta \, \sin\phi,$$

and therefore form a compatible chart on their region of definition.

In a similar way the n-sphere S^n,

$$S^n = \left\{ \mathbf{x} \in \mathbb{R}^{n+1} \;\middle|\; \left(x^1\right)^2 + \left(x^2\right)^2 + \cdots + \left(x^{n+1}\right)^2 = 1 \right\}$$

is a differentiable manifold of dimension n. A set of charts providing an atlas is the set of rectangular coordinates on all hemispheres, $(U_i^+, \phi_i^\pm)$ and $(U_i^-, \phi_i^\pm)$, where

$$U_i^+ = \{\mathbf{x} \in S^n \mid x^i > 0\}, \qquad U_i^- = \{\mathbf{x} \in S^n \mid x^i < 0\},$$

and $\phi_i^+ : U_i^+ \to \mathbb{R}^n$ and $\phi_i^- : U_i^- \to \mathbb{R}^n$ are both defined by

$$\phi_i^\pm(\mathbf{x}) = \left(x^1, x^2, \dots, x^{i-1}, x^{i+1}, \dots, x^{n+1}\right).$$

Exercise: Prove that St_N and St_S are compatible, by showing they are related by

$$X' = \frac{X}{X^2 + Y^2}, \qquad Y' = \frac{Y}{X^2 + Y^2}.$$

Example 15.5 The set of $n \times n$ real matrices $M(n, \mathbb{R})$ can be put in one-to-one correspondence with points of $\mathbb{R}^{n^2}$, through the map $\phi : M(n, \mathbb{R}) \to \mathbb{R}^{n^2}$ defined by

$$\phi(A = [a_{ij}]) = \left(a_{11}, a_{12}, \dots, a_{1n}, a_{21}, a_{22}, \dots, a_{nn}\right).$$

This provides $M(n, \mathbb{R})$ with a Hausdorff topology inherited from R^{n^2} in the obvious way. The differentiable structure generated by the chart $(M(n, \mathbb{R}), \phi)$ converts $M(n, \mathbb{R})$ into a differentiable manifold of dimension n^2.

The group of $n \times n$ real non-singular matrices $GL(n, \mathbb{R})$ consists of $n \times n$ real matrices having non-zero determinant. The determinant map $\det : GL(n, \mathbb{R}) \to \mathbb{R}$ is continuous since it is made up purely of polynomial operations, so that $\phi(GL(n, \mathbb{R})) = \det^{-1}(\mathbb{R} - \{0\})$ is an open subset of $\mathbb{R}^{n^2}$. Thus $GL(n, \mathbb{R})$ is a differentiable manifold of dimension n^2, as it is in one-to-one correspondence with an open submanifold of $\mathbb{R}^{n^2}$.

From any two differentiable manifolds M and N of dimensions m and n respectively, it is possible to form their **product** $M \times N$, which is the topological space defined in Section 10.4. Let (U_α, ϕ_α) and (V_β, ψ_β) be any families of mutually compatible charts on M and

N respectively, which generate the differentiable structures on these manifolds. The charts $(U_\alpha \times V_\beta, \phi_\alpha \times \psi_\beta)$, where $\phi_\alpha \times \psi_\beta : U_\alpha \times V_\beta \to \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$ defined by

$$\phi_\alpha \times \psi_\beta((p, q)) = (\phi_\alpha(p), \psi_\beta(q)) = \big(x^1(p), \dots, x^m(p),\ y^1(q), \dots, y^n(q)\big),$$

manifestly cover $M \times N$, and are clearly compatible in their overlaps. The maximal atlas generated by these charts is a differentiable structure on $M \times N$ making it into a differentiable manifold of dimension $m + n$.

Example 15.6 The topological 2-torus $T^2 = S^1 \times S^1$ (see Example 10.13) can be given a differentiable structure as a product manifold in the obvious way from the manifold structure on S^1. Similarly, one can define the n-**torus** to be the product of n circles, $T^n = S^1 \times S^1 \times \cdots \times S^1 = \big(S^1\big)^n$.

Problems

Problem 15.1 Show that the group of unimodular matrices $SL(n, \mathbb{R}) = \{A \in GL(n, \mathbb{R}) \mid \det A = 1\}$ is a differentiable manifold.

Problem 15.2 On the n-sphere S^n find coordinates corresponding to (i) stereographic projection, (ii) spherical polars.

Problem 15.3 Show that the real projective n-space P^n defined in Example 10.15 as the set of straight lines through the origin in $\mathbb{R}^{n+1}$ is a differentiable manifold of dimension n, by finding an atlas of compatible charts that cover it.

Problem 15.4 Define the complex projective n-space CP^n in a similar way to Example 10.15 as lines in $\mathbb{C}^{n+1}$ of the form $\lambda(z^0, z^1, \dots, z^n)$ where $\lambda, z^0, \dots, z^n \in \mathbb{C}$. Show that CP^n is a differentiable (real) manifold of dimension $2n$.

15.2 Differentiable maps and curves

Let M be a differentiable manifold of dimension n. A map $f : M \to \mathbb{R}$ is said to be **differentiable** at a point $p \in M$ if for some coordinate chart $(U, \phi; x^i)$ at p the function $\hat{f} = f \circ \phi^{-1} : \phi(U) \to \mathbb{R}$ is differentiable at $\phi(p) = \mathbf{x}(p) = \big(x^1(p),\ x^2(p), \dots,\ x^n(p)\big)$. This definition is independent of the choice of chart at p, for if (U', ϕ') is a second chart at p that is compatible with (U, ϕ), then

$$\hat{f}' = f \circ \phi'^{-1} = \hat{f} \circ \phi \circ \phi'^{-1}$$

is C^∞ since it is a differentiable function of a differentiable function. We denote by $\mathcal{F}_p(M)$ the set of all real-valued functions on M that are differentiable at $p \in M$.

Given an open set $V \subseteq M$, a real-valued function $f : M \to \mathbb{R}$ is said to be **differentiable** or **smooth** between manifolds on V if it is differentiable at every point $p \in V$. Clearly, the function need only be defined on the open subset V for this definition. We will denote the set of all real-valued functions on M that are differentiable on an open subset V by the symbol $\mathcal{F}(V)$. Since the sum $f + g$ and product fg of any pair of differentiable functions f and

g are differentiable functions, $\mathcal{F}(V)$ is a ring. Furthermore, $\mathcal{F}(V)$ is closed with respect to taking linear combinations $f + ag$ where $a \in \mathbb{R}$ and is therefore also a real vector space that at the same time is a commutative algebra with respect to multiplication of functions fg. All functions in $\mathcal{F}_p(M)$ are differentiable on some open neighbourhood V of the point $p \in M$.

Exercise: Show that $\mathcal{F}_p(M)$ is a real commutative algebra with respect to multiplication of functions.

If M and N are differentiable manifolds, dimensions m and n respectively, then a map $\alpha : M \to N$ is **differentiable at** $p \in M$ if for any pair of coordinate charts $(U, \phi; x^i)$ and $(V, \psi; y^a)$ covering p and $\alpha(p)$ respectively, its coordinate representation

$$\hat{\alpha} = \psi \circ \alpha \circ \phi^{-1} : \phi(U) \to \psi(V)$$

is differentiable at $\phi(p)$. As for differentiable real-valued functions this definition is independent of the choice of charts. The map $\hat{\alpha}$ is represented by n differentiable real-valued functions

$$y^a = \alpha^a\big(x^1, x^2, \dots, x^m\big) \quad (a = 1, 2, \dots, n),$$

where $\alpha^a = \mathrm{pr}_a \circ \hat{\alpha}$.

A **diffeomorphism** is a map $\alpha : M \to N$ that is one-to-one and both α and $\alpha^{-1} : N \to M$ are differentiable. Two manifolds M and N are said to be **diffeomorphic**, written $M \cong N$, if there exists a diffeomorphism $\alpha : M \to N$; the dimensions of the two manifolds must of course be equal, $m = n$. It is a curious and difficult fact that there exist topological manifolds with more than one inequivalent differentiable structure.

A **smooth parametrized curve** on an n-dimensional manifold M is a differentiable map $\gamma : (a, b) \to M$ from an open interval $(a, b) \subseteq \mathbb{R}$ of the real line into M. The curve is said to **pass through** p at $t = t_0$ if $\gamma(t_0) = p$, where $a < t_0 < b$. Note that a parametrized curve consists of a map, not the image points $\gamma(t) \in M$. Changing the parameter from t to $t' = f(t)$, where $f : \mathbb{R} \to \mathbb{R}$ is a monotone differentiable function and $a' = f(a) < t' < b' = f(b)$, changes the parametrized curve to $\gamma' = \gamma \circ f$, but has no effect on the image points in M. Given a chart $(U, \phi; x^i)$ at p, the inverse image of the open set U is an open subset $\gamma^{-1}(U) \subseteq \mathbb{R}$. Let (a_1, b_1) be the connected component of this set that contains the real number t_0 such that $p = \gamma(t_0)$. The 'coordinate representation' of the parametrized curve γ induced by this chart is the smooth curve $\hat{\gamma} = \phi \circ \gamma : (a_1, b_1) \to \mathbb{R}^n$, described by n real-valued functions $x^i = \gamma^i(t)$ where $\gamma^i = \mathrm{pr}_i \circ \phi \circ \gamma$. We often write this simply as $x^i = x^i(t)$ when there is no danger of any misunderstanding (see Fig. 15.3). In another chart $(U'; x'^j)$ the n functions representing the curve change to $x'^j = \gamma'^j(t) = x'^j\big(\gamma(t)\big)$. Assuming compatible charts, these new functions representing the curve are again smooth, although it is possible that the parameter range (a', b') is altered.

Problems

Problem 15.5 Let $\mathbb{R}'$ be the manifold consisting of $\mathbb{R}$ with differentiable structure generated by the chart $(\mathbb{R}; y = x^3)$. Show that the identity map $\mathrm{id}_{\mathbb{R}} : \mathbb{R}' \to \mathbb{R}$ is a differentiable homeomorphism, which is not a diffeomorphism.

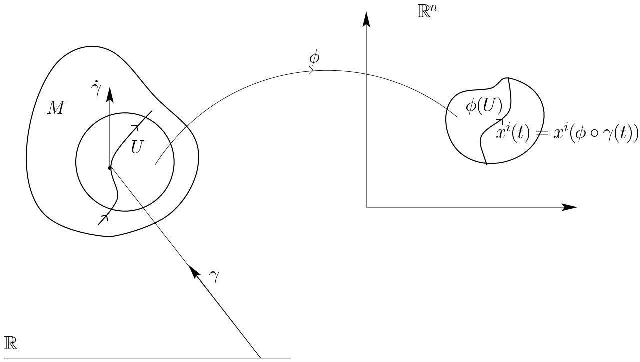

Figure 15.3 Parametrized curve on a differentiable manifold

Problem 15.6 Show that the set of real $m \times m$ matrices $M(m, n; \mathbb{R})$ is a manifold of dimension mn. Show that the matrix multiplication map $M(m, k; \mathbb{R}) \times M(k, n; \mathbb{R}) \to M(m, n; \mathbb{R})$ is differentiable.

15.3 Tangent, cotangent and tensor spaces

Tangent vectors

Let $x^i = x^i(t)$ be a curve in $\mathbb{R}^n$ passing through the point $\mathbf{x}_0 = \mathbf{x}(t_0)$. In elementary mathematics it is common to define the 'tangent' to the curve, or 'velocity', at $\mathbf{x}_0$ as the n-vector $\mathbf{v} = \dot{(\mathbf{x})} = \big(\dot x^1, \dot x^2, \dots, \dot x^n\big)$ where $\dot x^i = (\mathrm{d}x^i/\mathrm{d}t)_{t=t_0}$. In an n-dimensional manifold it is not satisfactory to define the tangent by its components, since general coordinate transformations are permitted. For example, by a rotation of axes in $\mathbb{R}^n$ it is possible to achieve that the tangent vector has components $\mathbf{v} = (v, 0, 0, \dots, 0)$. A coordinate-independent, or *invariant*, approach revolves around the concept of the **directional derivative** of a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ along the curve at $\mathbf{x}_0$,

$$Xf = \frac{\mathrm{d}f(\mathbf{x}(t))}{\mathrm{d}t}\Big|_{t=t_0} = \frac{\mathrm{d}x^i(t)}{\mathrm{d}t}\Big|_{t=t_0}\frac{\partial f(\mathbf{x})}{\partial x^i}\Big|_{\mathbf{x}=\mathbf{x}_0},$$

where X is the linear differential operator

$$X = \frac{\mathrm{d}x^i(t)}{\mathrm{d}t}\Big|_{t=t_0}\frac{\partial}{\partial x^i}\Big|_{\mathbf{x}=\mathbf{x}_0}.$$

The value of the operator X when applied to a function f only depends on the values taken by the function in a neighbourhood of $\mathbf{x}_0$ along the curve in question, and is independent of coordinates chosen for the space $\mathbb{R}^n$. The above expansion demonstrates, however, that the components of the tangent vector in any coordinates on $\mathbb{R}^n$ can be extracted from the directional derivative operator from its coefficients of expansion in terms of coordinate partial derivatives.

The directional derivative operator X is a real-valued map on the algebra of differentiable functions at $\mathbf{x}_0$. Two important properties hold for the map $X : \mathcal{F}_{\mathbf{x}_0}(\mathbb{R}^n) \to \mathbb{R}$:

(i) It is linear on the vector space $\mathcal{F}_{\mathbf{x}_0}(\mathbb{R}^n)$; that is, for any pair of functions f, g and real numbers a, b we have $X(af + bg) = aXf + bXg$.
(ii) The application of X on any product of functions fg in the algebra $\mathcal{F}_{\mathbf{x}_0}(\mathbb{R}^n)$ is determined by the **Leibnitz rule**, $X(fg) = f(\mathbf{x}_0)Xg + g(\mathbf{x}_0)Xf$.

These two properties completely characterize the class of directional derivative operators (see Theorem 15.1), and will be used to motivate the definition of a tangent vector at a point of a general manifold.

A **tangent vector** X_p at any point p of a differentiable manifold M is a linear map from the algebra of differentiable functions at p to the real numbers, $X_p : \mathcal{F}_p(M) \to \mathbb{R}$, which satisfies the Leibnitz rule for products:

$$X_p(af + bg) = aX_p f + bX_p g \qquad \text{(linearity)}, \tag{15.3}$$

$$X_p(fg) = f(p)X_p g + g(p)X_p f \quad \text{(Leibnitz rule)}. \tag{15.4}$$

The set of tangent vectors at p form a vector space $T_p(M)$, since any linear combination $aX_p + bY_p$ of tangent vectors at p, defined by

$$(aX_p + bY_p)f = aX_p f + bY_p f,$$

is a tangent vector at p since it satisfies (15.3) and (15.4). It is called the **tangent space at** p. If (U, ϕ) is any chart at p with coordinate functions x^i, define the operators

$$\left(\partial_{x^i}\right)_p \equiv \frac{\partial}{\partial x^i}\bigg|_p : \mathcal{F}_p(M) \to \mathbb{R}$$

by

$$\left(\partial_{x^i}\right)_p f \equiv \frac{\partial}{\partial x^i}\bigg|_p f = \frac{\partial \hat{f}(x^1, \dots, x^n)}{\partial x^i}\bigg|_{\mathbf{x}=\phi(p)}, \tag{15.5}$$

where $\hat{f} = f \circ \phi^{-1} : \mathbb{R}^n \to \mathbb{R}$. These operators are clearly tangent vectors since they satisfy (15.3) and (15.4). Thus any linear combination

$$X_p = X^i \frac{\partial}{\partial x^i}\bigg|_p \equiv \sum_{i=1}^n X^i \frac{\partial}{\partial x^i}\bigg|_p \quad \text{where} \quad X^i \in \mathbb{R}$$

is a tangent vector. The coefficients X^j can be computed from the action of X on the coordinate functions x^j themselves:

$$X_p x^j = X^i \frac{\partial x^j}{\partial x^i}\bigg|_{\mathbf{x}=\phi(p)} = X^i \delta^j_i = X^j.$$

Theorem 15.1 *If $(U, \phi; x^i)$ is a chart at $p \in M$, then the operators $\left(\partial_{x^i}\right)_p$ defined by (15.5) form a basis of the tangent space $T_p(M)$, and its dimension is $n = \dim M$.*

Proof: Let X_p be a tangent vector at the given fixed point p. Firstly, it follows by the Leibnitz rule (15.4) that X_p applied to a unit constant function $f = 1$ always results in zero, $X_p 1 = 0$, for

$$X_p 1 = X_p(1.1) = 1.X_p 1 + 1.X_p 1 = 2X_p 1.$$

By linearity, X_p applied to any constant function $f = c$ results in zero, $Xc = X(c.1) = cX1 = 0$.

Set the coordinates of p to be $\phi(p) = \mathbf{a} = (a^1, a^2, \dots, a^n)$, and let $\mathbf{y} = \phi(q)$ be any point in a neighbourhood ball $B_r(\mathbf{a}) \subseteq \phi(U)$. The function $\hat{f} = f \circ \phi^{-1}$ can be written as

$$\begin{aligned}
\hat{f}(y^1, y^2, \dots, y^n) &= \hat{f}(y^1, y^2, \dots, y^n) - \hat{f}(y^1, \dots, y^{n-1}, a^n) \\
&\quad + \hat{f}(y^1, \dots, y^{n-1}, a^n) - \hat{f}(y^1, \dots, y^{n-2}, a^{n-1}, a^n) + \dots \\
&\quad + \hat{f}(y^1, a^2, \dots, a^n) - \hat{f}(a^1, a^2, \dots, a^n) + \hat{f}(a^1, a^2, \dots, a^n) \\
&= \hat{f}(a^1, a^2, \dots, a^n) \\
&\quad + \sum_{i=1}^{n} \int_0^1 \frac{\partial \hat{f}(y^1, \dots, y^{i-1}, a^i + t(y^i - a^i), a^{i+1}, \dots, a^n)}{\partial t} \mathrm{d}t \\
&= \hat{f}(\mathbf{a}) + \sum_{i=1}^{n} \int_0^1 \frac{\partial \hat{f}}{\partial x^i}(y^1, \dots, y^{i-1}, a^i + t(y^i - a^i), a^{i+1}, \dots, a^n) \\
&\quad \times \mathrm{d}t (y^i - a^i).
\end{aligned}$$

Hence, in a neighbourhood of $\mathbf{a} = \phi(p)$, any function $\hat{f}$ can be written in the form

$$\hat{f}(\mathbf{y}) = \hat{f}(\mathbf{a}) + \hat{f}_i(\mathbf{y})(y^i - a^i) \tag{15.6}$$

where the functions $\hat{f}_i(y^1, y^2, \dots, y^n)$ are differentiable at $\mathbf{a}$. Thus, in a neighbourhood of p,

$$f(q) = \hat{f} \circ \phi = f(p) + f_i(q)\big(x^i(q) - a^i\big)$$

where $f_i = \hat{f}_i \circ \phi \in \mathcal{F}_p(M)$. Using the linear and Leibnitz properties of X_p,

$$X_p f = X_p f(p) + X_p f_i \big(x^i(p) - a^i\big) + f_i(p)\big(X_p x^i - X_p a^i\big) = f_i(p) X_p x^i$$

since $X_p c = 0$ for any constant c, and $x^i(p) = a^i$. Furthermore,

$$f_i(p) = \frac{\partial \hat{f}}{\partial x^i}(a^1, \dots, a^n) \int_0^1 \mathrm{d}t = \frac{\partial}{\partial x^i}\bigg|_p f,$$

and the tangent vectors $\big(\partial_{x^i}\big)_p$ span the tangent space $T_p(M)$,

$$X_p = X^i \frac{\partial}{\partial x^i}\bigg|_p = X^i \big(\partial_{x^i}\big)_p \quad \text{where} \quad X^i = X_p x^i. \tag{15.7}$$

To show that they form a basis, we need linear independence. Suppose

$$A^i \frac{\partial}{\partial x^i}\bigg|_p = 0,$$

then the action on the coordinate functions $f = x^j$ gives

$$0 = A^i \frac{\partial}{\partial x^i}\bigg|_p x^j = A^i \frac{\partial x^j}{\partial x^i}\bigg|_{\mathbf{a}} = A^i \delta^j_i = A^j$$

as required. ■

This proof shows that, for every tangent vector X_p, the decomposition given by Eq. (15.7) is unique. The coefficients $X^i = X_p x^i$ are said to be the **components** of the tangent vector X_p in the chart $(U; x^i)$.

How does this definition of tangent vector relate to that given earlier for a curve in $\mathbb{R}^n$? Let $\gamma : (a, b) \to M$ be a smooth parametrized curve passing through the point $p \in M$ at $t = t_0$. Define the **tangent vector to the curve** at p to be the operator $\dot{\gamma}_p$ defined by the action on an arbitrary differentiable function f at p,

$$\dot{\gamma}_p f = \frac{\mathrm{d} f \circ \gamma(t)}{\mathrm{d}t}\Big|_{t=t_0}.$$

It is straightforward to verify that the $\dot{\gamma}_p$ is a tangent vector at p, as it satisfies Eqs. (15.3) and (15.4). In a chart with coordinate functions x^i at p, let the coordinate representation of the curve be $\hat{\gamma} = \phi \circ \gamma = \big(\gamma^1(t), \dots, \gamma^n(t)\big)$. Then

$$\dot{\gamma}_p f = \frac{\mathrm{d}\hat{f} \circ \hat{\gamma}(t)}{\mathrm{d}t}\Big|_{t=t_0} = \frac{\partial \hat{f}}{\partial x^i}\Big|_{\phi(p)} \frac{\mathrm{d}\gamma^i(t)}{\mathrm{d}t}\Big|_{t=t_0}$$

and

$$\dot{\gamma}_p = \dot{\gamma}^i(t_0) \frac{\partial}{\partial x^i}\Big|_p \quad \text{where} \quad \dot{\gamma}^i(t) = \frac{\mathrm{d}\gamma^i(t)}{\mathrm{d}t}.$$

In the case $M = \mathbb{R}^n$ the operator $\dot{\gamma}_p$ is precisely the directional derivative of the curve.

It is also true that every tangent vector is tangent to some curve. For example, the basis vectors $\big(\partial_{x^i}\big)_p$ are tangent to the 'coordinate lines' at $p = \phi^{-1}(\mathbf{a})$,

$$\gamma^i : t \mapsto \phi^{-1}\big((a^1, a^2, \dots, x^i = a^i + t - t_0, \dots, a^n)\big).$$

An arbitrary tangent vector $X_p = X^i \big(\partial_{x^i}\big)_p$ at p is tangent to the curve

$$\gamma : t \mapsto \phi^{-1}\big((a^1 + X^1(t - t_0), \dots, x^i = a^i + X^i(t - t_0), \dots, a^n + X^n(t - t_0))\big).$$

Example 15.7 The curves α, β and γ on $\mathbb{R}^2$ given respectively by

$$\begin{aligned} &\alpha^1(t) = 1 + \sin t \cos t \quad && \alpha^2(t) = 1 + 3t \cos 2t, \\ &\beta^1(t) = 1 + t && \beta^2(t) = 1 + 3t\mathrm{e}^{3t}, \\ &\gamma^1(t) = \mathrm{e}^t && \gamma^2(t) = \mathrm{e}^{3t}, \end{aligned}$$

all pass through the point $p = (1, 1)$ at $t = 0$ and are tangent to each other there,

$$\dot{\alpha}_p = \dot{\beta}_p = \dot{\gamma}_p = \big(\partial_{x^1}\big)_p + 3\big(\partial_{x^2}\big)_p.$$

Cotangent and tensor spaces

The dual space $T_p^*(M)$ associated with the tangent space at $p \in M$ is called the **cotangent space** at p. It consists of all linear functionals on $T_p(M)$, also called **covectors** or **1-forms** at p. The action of a covector ω_p at p on a tangent vector X_p will be denoted by $\omega_p(X_p)$, $\langle \omega_p, X_p \rangle$ or $\langle X_p, \omega_p \rangle$. From Section 3.7 we have that $\dim T_p^*(M) = n = \dim T_p(M) = \dim M$.

If f is any function that is differentiable at p, we define its **differential** at p to be the covector $(\mathrm{d}f)_p$ whose action on any tangent vector X_p at p is given by

$$\langle(\mathrm{d}f)_p, X_p\rangle = X_p f. \tag{15.8}$$

This is a linear functional since, for any tangent vectors X_p, Y_p and scalars $a, b \in \mathbb{R}$,

$$\langle(\mathrm{d}f)_p, aX_p + bY_p\rangle = (aX_p + bY_p)f = aX_p f + bY_p f = a\langle(\mathrm{d}f)_p, X_p\rangle + b\langle(\mathrm{d}f)_p, Y_p\rangle.$$

Given a chart $(U, \phi; x^i)$ at p, the differentials of the coordinate functions have the property

$$\langle(\mathrm{d}x^i)_p, X_p\rangle = X_p x^i = X^i$$

where X^i are the components of the tangent vector, $X_p = X^i\left(\partial_{x^i}\right)_p$. Applying $(\mathrm{d}x^i)_p$ to the basis tangent vectors, we have

$$\langle(\mathrm{d}x^i)_p, \left(\partial_{x^j}\right)_p\rangle = \frac{\partial}{\partial x^j}\Big|_p x^i = \frac{\partial x^i}{\partial x^j}\Big|_{\phi(p)} = \delta^i_j.$$

Hence the linear functionals $(\mathrm{d}x^1)_p$, $(\mathrm{d}x^2)_p, \dots,$ $(\mathrm{d}x^n)_p$ are the dual basis, spanning the cotangent space, and every covector at p has a unique expansion

$$\omega_p = w_i(\mathrm{d}x^i)_p \quad \text{where} \quad w_i = \langle\omega_p, \left(\partial_{x^i}\right)_p\rangle.$$

The w_i are called the **components** of the linear functional ω_p in the chart $(U; x^i)$.

The differential of any function at p has a coordinate expansion

$$(\mathrm{d}f)_p = f_i(\mathrm{d}x^i)_p$$

where

$$f_i = \langle(\mathrm{d}f)_p, \left(\partial_{x^i}\right)_p\rangle = \frac{\partial}{\partial x^i}\Big|_p f = \frac{\partial \hat{f}}{\partial x^i}\Big|_{\phi(p)}.$$

A common way of writing this is the ‘chain rule’

$$(\mathrm{d}f)_p = f_{,i}(p)(\mathrm{d}x^i)_p \tag{15.9}$$

where

$$f_{,i} = \frac{\partial \hat{f}}{\partial x^i} \circ \phi.$$

These components are often referred to as the *gradient* of the function at p. Differentials have never found a comfortable place in calculus as non-vanishing quantities that are ‘arbitrarily small’. The concept of differentials as linear functionals avoids these problems, yet has all the desired properties such as the chain rule of multivariable calculus.

As in Chapter 7, a **tensor of type** (r, s) at p is a multilinear functional

$$A_p : \underbrace{T^*_p(M) \times T^*_p(M) \times \cdots \times T^*_p(M)}_{r} \times \underbrace{T_p(M) \times \cdots \times T_p(M)}_{s} \to \mathbb{R},$$

We denote the tensor space of type (r, s) at p by $T^{(r,s)}_p(M)$. It is a vector space of dimension n^{r+s}.

Vector and tensor fields

A **vector field** X is an assignment of a tangent vector X_p at each point $p \in M$. In other words, X is a map from M to the set $\bigcup_{p \in M} T_p(M)$ with the property that the image of every point, $X(p)$, belongs to the tangent space $T_p(M)$ at p. We may thus write X_p in place of $X(p)$. The vector field is said to be **differentiable** or **smooth** if for every differentiable function $f \in \mathcal{F}(M)$ the function Xf defined by

$$(Xf)(p) = X_p f$$

is differentiable, $Xf \in \mathcal{F}(M)$. The set of all differentiable vector fields on M is denoted $\mathcal{T}(M)$.

Exercise: Show that $\mathcal{T}(M)$ forms a module over the ring of functions $\mathcal{F}(M)$: if X and Y are vector fields, and $f \in \mathcal{F}(M)$ then $X + fY$ is a vector field.

Every smooth vector field defines a map $X : \mathcal{F}(M) \to \mathcal{F}(M)$, which is linear

$$X(af + bg) = aXf + bXg \qquad \text{for all } f, g \in \mathcal{F}(M) \text{ and all } a, b \in \mathbb{R},$$

and satisfies the Leibnitz rule for products

$$F(fg) = fXg + gXf.$$

Conversely, any map X with these properties defines a smooth vector field, since for each point p the map $X_p : \mathcal{F}_p(M) \to \mathcal{F}_p(M)$ defined by $X_p f = (Xf)(p)$ satisfies Eqs. (15.3) and (15.4) and is therefore a tangent vector at p.

We may also define vector fields on any open set U in a similar way as an assignment of a tangent vector at every point of U such that $Xf \in \mathcal{F}(U)$ for all $f \in \mathcal{F}(U)$. By the term **local basis of vector fields** at p we will mean an open neighbourhood U of p and a set of vector fields $\{e_1, e_2, \dots, e_n\}$ on U such that the tangent vectors $(e_i)_q$ span the tangent space $T_q(M)$ at each point $q \in U$. For any chart $(U, \phi; x^i)$, define the vector fields on the domain U

$$\partial_{x^i} \equiv \frac{\partial}{\partial x^i} : \mathcal{F}(U) \to \mathcal{F}(U)$$

by

$$\partial_{x^i} f = \frac{\partial}{\partial x^i} f = \frac{\partial f \circ \phi^{-1}}{\partial x^i}.$$

These vector fields assign the basis tangent vectors $\left(\partial_{x^i}\right)_p$ at each point $p \in U$, and form a local basis of vector fields at any point of U. When it is restricted to the coordinate domain U, every differentiable vector field X on M has a unique expansion in terms of these vector fields

$$X\big|_U = X^i \frac{\partial}{\partial x^i} = X^i \partial_{x^i}$$

where the components $X^i : U \in \mathbb{R}$ are differentiable functions on U. The local vector fields ∂_{x^i} form a module basis on U, but they are not a vector space basis since as a vector space $\mathcal{T}(U)$ is the direct sum of tangent spaces at all points $p \in U$, and is infinite dimensional.

In a similar way we define a **covector field** or **differentiable 1-form** ω as an assignment of a covector ω_p at each point $p \in M$, such that the function $\langle \omega, X \rangle$ defined by $\langle \omega, X \rangle(p) = \langle \omega_p, X_p \rangle$ is differentiable for every smooth vector field X. The space of differentiable 1-forms will be denoted $\mathcal{T}^*(M)$. Given any smooth function f, let $\mathrm{d}f$ be the differentiable 1-form defined by assigning the differential $\mathrm{d}f_p$ at each point p, so that

$$\langle \mathrm{d}f, X \rangle = Xf \qquad \text{for all } X \in \mathcal{T}(M).$$

We refer to this covector field simply as the **differential** of f. A local module basis on any chart $(U, \phi; x^i)$ consists of the 1-forms $\mathrm{d}x^i$, which have the property

$$\langle \mathrm{d}x^i, \partial_{x^j} \rangle = \frac{\partial x^i}{\partial x^j} = \delta^i_{\ j}.$$

Every differential can be expanded locally by the chain rule,

$$\mathrm{d}f = f_{,i}\, \mathrm{d}x^i \quad \text{where} \quad f_{,i} = \frac{\partial}{\partial x^i} f. \tag{15.10}$$

Tensor fields are defined in a similar way, where the differentiable tensor field A of type (r, s) has a local expansion in any coordinate chart

$$A = A^{i_1 i_2 \dots i_r}{}_{j_1 \dots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \frac{\partial}{\partial x^{i_2}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_r}} \otimes \mathrm{d}x^{j_1} \otimes \dots \otimes \mathrm{d}x^{j_s}. \tag{15.11}$$

The components are differentiable functions over the coordinate domain U given by

$$A^{i_1 i_2 \dots i_r}{}_{j_1 \dots j_s} = A\Big(\mathrm{d}x^{i_1}, \mathrm{d}x^{i_2}, \dots, \mathrm{d}x^{i_r}, \frac{\partial}{\partial x^{j_1}}, \dots, \frac{\partial}{\partial x^{j_s}}\Big).$$

Coordinate transformations

Let $(U, \phi; x^i)$ and $(U', \phi'; x'^j)$ be any two coordinate charts. From the chain rule of partial differentiation

$$\frac{\partial}{\partial x'^j} = \frac{\partial x^i}{\partial x'^j} \frac{\partial}{\partial x^i}, \qquad \frac{\partial}{\partial x^i} = \frac{\partial x'^j}{\partial x^i} \frac{\partial}{\partial x'^j}. \tag{15.12}$$

Exercise: Show these equations by applying both sides to an arbitrary differentiable function f on M.

Substituting the transformations (15.12) into the expression of a tangent vector with respect to either of these bases

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

gives the *contravariant law of transformation* of components

$$X'^j = X^i \frac{\partial x'^j}{\partial x^i}. \tag{15.13}$$

The chain rule (15.10), written in coordinates x'^j and setting $f = x^i$, gives

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

Expressing a differentiable 1-form ω in both coordinate bases,

$$\omega = w_i \, \mathrm{d}x^i = w'_j \, \mathrm{d}x'^j,$$

we obtain the *covariant transformation law* of components

$$w'_j = \frac{\partial x^i}{\partial x'^j} w_i. \tag{15.14}$$

The component transformation laws (15.13) and (15.14) can be identified with similar formulae in Chapter 3 on setting

$$A^j_i = \frac{\partial x'^j}{\partial x^i}, \qquad A'^i_k = \frac{\partial x^i}{\partial x'^k}.$$

The transformation law of a general tensor of type (r, s) follows from Eq. (7.30):

$$T'^{i'_1 \dots i'_r}{}_{j'_1 \dots j'_s} = T^{i_1 \dots i_r}{}_{j_1 \dots j_s} \frac{\partial x'^{i'_1}}{\partial x^{i_1}} \cdots \frac{\partial x'^{i'_r}}{\partial x^{i_r}} \frac{\partial x^{j_1}}{\partial x'^{j'_1}} \cdots \frac{\partial x^{j_s}}{\partial x'^{j'_s}}. \tag{15.15}$$

Tensor bundles

The **tangent bundle** TM on a manifold M consists of the set-theoretical union of all tangent spaces at all points

$$TM = \bigcup_{p \in M} T_p(M).$$

There is a natural **projection map** $\pi : TM \to M$ defined by $\pi(u) = p$ if $u \in T_p(M)$, and for each chart $(U, \phi; x^i)$ on M we can define a chart $(\pi^{-1}(U), \tilde{\phi})$ on TM where the coordinate map $\tilde{\phi} : \pi^{-1}(U) \to \mathbb{R}^{2n}$ is defined by

$$\tilde{\phi}(v) = \big(x^1(p), \dots, x^n(p), v^1, \dots, v^n\big)$$
$$\text{where} \quad p = \pi(v) \quad \text{and} \quad v = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i}\Big|_p .$$

The topology on TM is taken to be the coarsest topology such that all sets $\tilde{\phi}^{-1}(A)$ are open whenever A is an open subset of $\mathbb{R}^{2n}$. With this topology these charts generate a maximal atlas on the tangent bundle TM, making it into a differentiable manifold of dimension $2n$.

Given an open subset $U \subseteq M$, a smooth map $X : U \to TM$ is said to be a **smooth vector field on** U if $\pi \circ X = \mathrm{id}\big|_U$. This agrees with our earlier notion, since it assigns exactly one tangent vector from the tangent space $T_p(M)$ to the point $p \in U$. A similar idea may be used for a **smooth vector field along a parametrized curve** $\gamma : (a, b) \to M$, defined to be a smooth curve $X : (a, b) \to TM$ that *lifts* γ to the tangent bundle in the sense that $\pi \circ X = \gamma$. Essentially, this defines a tangent vector at each point of the curve, not necessarily tangent *to* the curve, in a differentiable manner.

The **cotangent bundle** T^*M is defined in an analogous way, as the union of all cotangent spaces $T^*_p(M)$ at all points $p \in M$. The generating charts have the form $(\pi^{-1}(U), \tilde{\phi})$ on

T^*M where the coordinate map $\tilde{\phi} : \pi^{-1}(U) \to \mathbb{R}^{2n}$ is defined by

$$\tilde{\phi}(\omega_p) = \big(x^1(p), \ldots, x^n(p), w_1, \ldots, w_n\big)$$

$$\text{where} \quad p = \pi(\omega) \quad \text{and} \quad \omega = \sum_{i=1}^{n} w_i (\mathrm{d}x^i)_p,$$

making T^*M into a differentiable manifold of dimension $2n$. This process may be extended to produce the tensor bundle of type $T^{(r,s)}M$, a differentiable manifold of dimension $n + n^{r+s}$.

Problems

Problem 15.7 Let $\gamma : \mathbb{R} \to \mathbb{R}^2$ be the curve $x = 2t + 1$, $y = t^2 - 3t$. Show that at an arbitrary parameter value t the tangent vector to the curve is $X_{\gamma(t)} = \dot{\gamma} = 2\partial_x + (2t - 3)\partial_y = 2\partial_x + (x - 4)\partial_y$. If $f : \mathbb{R}^2 \to \mathbb{R}$ is the function $f = x^2 - y^2$, write f as a function of t along the curve and verify the identities

$$X_{\gamma(t)} f = \frac{\mathrm{d}f(t)}{\mathrm{d}t} = \langle (\mathrm{d}f)_{\gamma(t)}, X_{\gamma(t)} \rangle = C^1_1 (\mathrm{d}f)_{\gamma(t)} \otimes X_{\gamma(t)}.$$

Problem 15.8 Let $x^1 = x$, $x^2 = y$, $x^3 = z$ be ordinary rectangular cartesian coordinates in $\mathbb{R}^3$, and let $x'^1 = r$, $x'^2 = \theta$, $x'^3 = \phi$ be the usual transformation to polar coordinates.

(a) Calculate the Jacobian matrices $[\partial x^i / \partial x'^j]$ and $[\partial x'^i / \partial x^j]$.

(b) In polar coordinates, work out the components of the covariant vector fields having components in rectangular coordinates (i) $(0, 0, 1)$, (ii) $(1, 0, 0)$, (iii) (x, y, z).

(c) In polar coordinates, what are the components of the contravariant vector fields whose components in rectangular coordinates are (i) (x, y, z), (ii) $(0, 0, 1)$, (iii) $(-y, x, 0)$.

(d) If g_{ij} is the covariant tensor field whose components in rectangular coordinates are δ_{ij}, what are its components g'_{ij} in polar coordinates?

Problem 15.9 Show that the curve

$$2x^2 + 2y^2 + 2xy = 1$$

can be converted by a rotation of axes to the standard form for an ellipse

$$x'^2 + 3y'^2 = 1.$$

If $x' = \cos\psi$, $y' = \frac{1}{\sqrt{3}} \sin\psi$ is used as a parametrization of this curve, show that

$$x = \frac{1}{\sqrt{2}}\Big(\cos\psi + \frac{1}{\sqrt{3}}\sin\psi\Big), \qquad y = \frac{1}{\sqrt{2}}\Big(-\cos\psi + \frac{1}{\sqrt{3}}\sin\psi\Big).$$

Compute the components of the tangent vector

$$X = \frac{\mathrm{d}x}{\mathrm{d}\psi}\partial_x + \frac{\mathrm{d}y}{\mathrm{d}\psi}\partial_y.$$

Show that $X(f) = (2/\sqrt{3})\big(x^2 - y^2\big)$.

Problem 15.10 Show that the tangent space $T_{(p,q)}(M \times N)$ at any point (p, q) of a product manifold $M \times N$ is naturally isomorphic to the direct sum of tangent spaces $T_p(M) \oplus T_q(N)$.

Problem 15.11 On the unit 2-sphere express the vector fields ∂_x and ∂_y in terms of the polar coordinate basis ∂_θ and ∂_ϕ. Again in polar coordinates, what are the dual forms to these vector fields?

Problem 15.12 Express the vector field ∂_ϕ in polar coordinates (θ, ϕ) on the unit 2-sphere in terms of stereographic coordinates X and Y.

15.4 Tangent map and submanifolds

The tangent map and pullback of a map

Let $\alpha : M \to N$ be a differentiable map between manifolds M and N, where $\dim M = m$, $\dim N = n$. This induces a map $\alpha_* : T_p(M) \to T_{\alpha(p)}(N)$, called the **tangent map** of α, whereby the tangent vector $Y_{\alpha(p)} = \alpha_* X_p$ is defined by

$$Y_{\alpha(p)} f = (\alpha_* X_p) f = X_p(f \circ \alpha)$$

for any function $f \in \mathcal{F}_{\alpha(p)}(N)$. This map is often called the *differential* of the map α, but this may cause confusion with our earlier use of this term.

Let $(U, \phi; x^i)$ and $(V, \psi; y^a)$ be charts at p and $\alpha(p)$, respectively. The map α has coordinate representation $\hat{\alpha} = \psi \circ \alpha \circ \phi^{-1} : \phi(U) \to \psi(V)$, written

$$y^a = \alpha^a(x^1, x^2, \dots, x^m) \quad (a = 1, \dots, n).$$

To compute the components Y^a of $Y_{\alpha(p)} = \alpha_* X_p = Y^a \big(\partial_{y^a}\big)_{\alpha(p)}$, we perform the following steps:

$$\begin{aligned}
Y_{\alpha(p)} f = Y^a \frac{\partial f \circ \psi^{-1}}{\partial y^a}\Big|_{\alpha(p)} &= X^i \frac{\partial}{\partial x^i}\Big|_p f \circ \alpha \\
&= X^i \frac{\partial f \circ \alpha \circ \phi^{-1}}{\partial x^i}\Big|_{\phi(p)} \\
&= X^i \frac{\partial f \circ \psi^{-1} \circ \hat{\alpha}}{\partial x^i}\Big|_{\phi(p)} \\
&= X^i \frac{\partial f \circ \psi^{-1}\big(\alpha^1(\mathbf{x}), \alpha^2(\mathbf{x}), \dots, \alpha^n(\mathbf{x})\big)}{\partial x^i}\Big|_{\mathbf{x}=\phi(p)} \\
&= X^i \frac{\partial y^a}{\partial x^i}\Big|_{\phi(p)} \frac{\partial f \circ \psi^{-1}}{\partial y^a}\Big|_{\hat{\alpha}(\phi(p))} \\
&= X^i \frac{\partial y^a}{\partial x^i}\Big|_{\phi(p)} \frac{\partial}{\partial y^a}\Big|_{\alpha(p)} f.
\end{aligned}$$

Hence

$$Y^a = X^i \frac{\partial y^a}{\partial x^i}\Big|_{\phi(p)}. \tag{15.16}$$

Exercise: If $\alpha : M \to N$ and $\beta : K \to M$ are differentiable maps between manifolds, show that

$$(\alpha \circ \beta)_* = \alpha_* \circ \beta_* . \tag{15.17}$$

The map $\alpha : M \to N$ also induces a map α^* between cotangent spaces, but in this case it acts in the reverse direction, called the **pullback** induced by α,

$$\alpha^* : T^*_{\alpha(p)}(N) \to T^*_p(M)\,.$$

The pullback of a 1-form $\omega_{\alpha(p)}$ is defined by requiring

$$\langle \alpha^* \omega_{\alpha(p)}, X_p \rangle = \langle \omega_{\alpha(p)}, \alpha_* X_p \rangle \tag{15.18}$$

for arbitrary tangent vectors X_p.

Exercise: Show that this definition uniquely defines the pullback $\alpha^* \omega_{\alpha(p)}$.

Exercise: Show that the pullback of a functional composition of maps is given by

$$(\alpha \circ \beta)^* = \beta^* \circ \alpha^*\,. \tag{15.19}$$

The notion of tangent map or pullback of a map can be extended to totally contravariant or totally covariant tensors, such as r-vectors or r-forms, but is only available for mixed tensors if the map is a diffeomorphism (see Problem 15.13). The tangent map does not in general apply to vector fields, for if it is not one-to-one the tangent vector may not be uniquely defined at the image point $\alpha(p)$ (see Example 15.8 below). However, no such ambiguity in the value of the pullback $\alpha^*\omega$ can ever arise at the inverse image point p, even if ω is a covector field, since its action on every tangent vector X_p is well-defined by Eq. (15.18). The pullback can therefore be applied to arbitrary differentiable 1-forms; the map α need not be either injective or surjective. This is one of the features that makes covector fields more attractive geometrical objects to deal with than vector fields. The following example should make this clear.

Example 15.8 Let $\alpha : \dot{\mathbb{R}}^3 = \mathbb{R}^3 - \{(0, 0, 0)\} \to \mathbb{R}^2$ be the differentiable map

$$\alpha : (x, y, z) \mapsto (u, v) \quad \text{where} \quad u = x + y + z, \quad v = \sqrt{x^2 + z^2}.$$

This map is neither surjective, since the whole lower half plane $v < 0$ is not mapped onto, nor injective, since, for example, the points $\mathbf{p} = (1, y, 0)$ and $\mathbf{q} = (0, y, 1)$ are both mapped to the point $(y + 1, 1)$. Consider a vector field

$$X = X^i \frac{\partial}{\partial x^i} = X^1 \frac{\partial}{\partial x} + X^2 \frac{\partial}{\partial y} + X^3 \frac{\partial}{\partial z}$$

and the action of the tangent map α_* at any point (x, y, z) is

$$\alpha_* X = \big(X^1 + X^2 + X^3\big)\frac{\partial}{\partial u} + \Big(X^1 \frac{x}{\sqrt{x^2 + z^2}} + X^3 \frac{z}{\sqrt{x^2 + z^2}}\Big)\frac{\partial}{\partial v}.$$

While this map is well-defined on the tangent space at any point $\mathbf{x} = (x, y, z)$, it does not in general map the vector field X to a vector field on $\mathbb{R}^2$. For example, no tangent vector can be assigned at $\mathbf{u} = \alpha(\mathbf{p}) = \alpha(\mathbf{q})$ as we would need

$$\big(X^1 + X^2 + X^3\big)(\mathbf{p}) \frac{\partial}{\partial u}\Big|_{\mathbf{u}} + X^1(\mathbf{p}) \frac{\partial}{\partial v}\Big|_{\mathbf{u}} = \big(X^1 + X^2 + X^3\big)(\mathbf{q}) \frac{\partial}{\partial u}\Big|_{\mathbf{u}} + X^3(\mathbf{q}) \frac{\partial}{\partial v}\Big|_{\mathbf{u}}.$$

There is no reason to expect these two tangent vectors at $\mathbf{u}$ to be identical.

However, if $\omega = w_1\,\mathrm{d}u + w_2\,\mathrm{d}v$ is a differentiable 1-form on $\mathbb{R}^2$ it induces a differentiable 1-form on $\dot{\mathbb{R}}^3$, on substituting $\mathrm{d}u = (\partial u/\partial x)\,\mathrm{d}x + (\partial u/\partial y)\,\mathrm{d}y + (\partial u/\partial z)\,\mathrm{d}z$, etc.

$$\alpha^*\omega = \Big(w_1 + w_2\frac{x}{\sqrt{x^2+z^2}}\Big)\,\mathrm{d}x + w_1\,\mathrm{d}y + \Big(w_1 + w_2\frac{z}{\sqrt{x^2+z^2}}\Big)\,\mathrm{d}z,$$

which is uniquely determined at any point $(x, y, z) \neq (0, 0, 0)$ by the components $w_1(u, v)$ and $w_2(u, v)$ of the differentiable 1-form at $(u, v) = \alpha(x, y, z)$.

Example 15.9 If $\gamma : \mathbb{R} \to M$ is a curve on M and $p = \gamma(t_0)$, the tangent vector to the curve at p is the image under the tangent map induced by γ of the ordinary derivative on the real line,

$$\dot\gamma_p = \gamma_*\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t_0},$$

for if $f : M \to \mathbb{R}$ is any function differentiable at p then

$$\gamma_*\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t_0}(f) = \frac{\mathrm{d}f\circ\gamma}{\mathrm{d}t}\Big|_{t_0} = \dot\gamma_p(f).$$

By a *curve with endpoints* we shall mean the restriction of a parametrized curve $\gamma : (a, b) \to M$ to a closed subinterval of $\gamma : [t_1, t_2] \to M$ where $a < t_1 < t_2 < b$. The **integral** of a 1-form α on the curve with end points γ is defined as

$$\int_\gamma \alpha = \int_{t_1}^{t_2} \alpha(\dot\gamma)\,\mathrm{d}t.$$

In a coordinate representation $x^i = \gamma^i(t)$ and $\alpha = \alpha_i\,\mathrm{d}x^i$,

$$\int_\gamma \alpha = \int_{t_1}^{t_2} \alpha_i(x(t))\frac{\mathrm{d}x^i}{\mathrm{d}t}\,\mathrm{d}t.$$

Let $\gamma' = \gamma\circ f$ be the curve related to γ by a change of parametrization $t' = f(t)$ where $f : \mathbb{R} \to \mathbb{R}$ is a monotone function on the real line. Then

$$\int_\gamma \alpha = \int_{\gamma'} \alpha$$

for, by the standard change of variable formula for a definite integral,

$$\begin{aligned}\int_\gamma \alpha &= \int_{t_1}^{t_2} \alpha_i(x(t))\frac{\mathrm{d}x^i(t(t'))}{\mathrm{d}t'}\frac{\mathrm{d}t'}{\mathrm{d}t}\,\mathrm{d}t\\ &= \int_{t'_1}^{t'_2} \alpha_i(x(t(t')))\frac{\mathrm{d}x'^i(t')}{\mathrm{d}t'}\,\mathrm{d}t'\\ &= \int_{\gamma'} \alpha.\end{aligned}$$

Hence the integral of a 1-form is independent of the parametrization on the curve γ.

The integral of α along γ is zero if its pullback to the real line vanishes, $\gamma^*(\alpha) = 0$, for

$$\int_\gamma \alpha = \int_{t_1}^{t_2} \langle\alpha, \gamma_*\frac{\mathrm{d}}{\mathrm{d}t}\rangle\,\mathrm{d}t = \int_{t_1}^{t_2} \langle\gamma^*(\alpha), \frac{\mathrm{d}}{\mathrm{d}t}\rangle\,\mathrm{d}t = 0.$$

If α is the differential of a scalar field $\alpha = \mathrm{d}f$ it is called an **exact** 1-form. The integral of an exact 1-form is independent of the curve connecting two points $p_1 = \gamma(t_1)$ and $p_2 = \gamma(t_2)$, for

$$\int_\gamma \mathrm{d}f = \int_{t_1}^{t_2} \langle \mathrm{d}f, \dot\gamma\rangle\, \mathrm{d}t = \int_{t_1}^{t_2} \frac{\mathrm{d}f(\gamma(t))}{\mathrm{d}t}\mathrm{d}t = f(\gamma(t_2)) - f(\gamma(t_1)),$$

which only depends on the value of f at the end points. In particular, the integral of an exact 1-form vanishes on any closed circuit, since $\gamma(t_1) = \gamma(t_2)$.

For general 1-forms, the integral is usually curve-dependent. For example, let $\alpha = x\mathrm{d}y$ on the manifold $\mathbb{R}^2$ with coordinates (x, y). Consider the following two curves connecting $p_1 = (-1, 0)$ to $p_2 = (1, 0)$:

$$\begin{aligned} &\gamma_1 : \ x = t, \quad y = 0 && t_1 = -1, \quad t_2 = 1, \\ &\gamma_2 : \ x = \cos t, \quad y = \sin t && t_1 = -\pi, \quad t_2 = 0. \end{aligned}$$

The pullback of α to the first curve vanishes, $\gamma_1^* x\, \mathrm{d}y = t\mathrm{d}0 = 0$, while the pullback to γ_2 is given by

$$\gamma_2^* x\, \mathrm{d}y = \cos t\, \mathrm{d}(y \circ \gamma_2(t)) = \cos t\, \mathrm{d}(\sin t) = \cos^2 t\, \mathrm{d}t.$$

Hence

$$\int_{\gamma_2} x\, \mathrm{d}y = \int_{-\pi}^{0} \cos^2 t\, \mathrm{d}t = \frac{\pi}{2} \neq \int_{\gamma_1} x\, \mathrm{d}y = 0.$$

Submanifolds

Let $\alpha : M \to N$ be a differentiable mapping where $m = \dim M \leq n \dim N$. The map is said to be an **immersion** if the tangent map $\alpha_* : T_p(M) \to T_{\alpha(p)}(N)$ is injective at every point $p \in M$; i.e., α_* is everywhere a non-degenerate linear map. From the inverse function theorem, it is straightforward to show that there exist charts at any point p and its image $\alpha(p)$ such that the map $\hat\alpha$ is represented as

$$\begin{aligned} y^i &= \alpha^i(x^1, x^2, \dots, x^m) = x^i \quad \text{for } i = 1, \dots, m, \\ y^a &= \alpha^a(x^1, x^2, \dots, x^m) = 0 \quad \text{for } a > m. \end{aligned}$$

A detailed proof may be found in [11].

Example 15.10 In general the image $\alpha(M) \subset N$ of an immersion is not a genuine 'submanifold', since there is nothing to prevent self-intersections. For example the mapping $\alpha : \mathbb{R} \to \mathbb{R}^2$ defined by

$$x = \alpha^1(t) = t(t^2 - 1), \qquad y = \alpha^2(t) = t^2 - 1$$

is an immersion since its Jacobian matrix is everywhere non-degenerate,

$$\begin{pmatrix} \dfrac{\partial x}{\partial t} & \dfrac{\partial y}{\partial t} \end{pmatrix} = \begin{pmatrix} 2t^2 - 1 & 2t \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \end{pmatrix} \qquad \text{for all } t \in \mathbb{R}.$$

The subset $\alpha(\mathbb{R}) \subset \mathbb{R}^2$ does not, however, inherit the manifold structure of $\mathbb{R}$ since there is a self-intersection at $t = \pm 1$, as shown in Fig. 15.4.

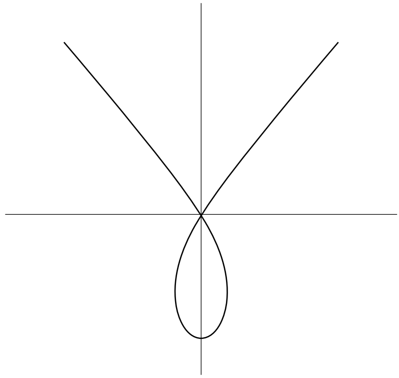

Figure 15.4 Immersion that is not a submanifold

In order to have a natural manifold structure on the subset $\alpha(M)$ we require that the map α is itself injective as well as its tangent map α_*. The map is then called an **embedding**, and the pair (M, α) an **embedded submanifold** of N.

Example 15.11 Let A be any open subset of a manifold M. As in Example 15.2, it inherits a manifold structure from M, whereby a chart is said to be admissible if it has the form $(U \cap A, \phi\big|_{U\cap A})$ for some chart (U, ϕ) on M. With this differentiable structure, A is said to be an **open submanifold** of M. It evidently has the same dimension as M. The pair $(A, \mathrm{id}\big|_A)$ is an embedded submanifold of M.

Example 15.12 Let $T^2 = S^1 \times S^1$ be the 2-torus (see Example 15.6). The space T^2 can also be viewed as the factor space $R^2/\text{mod } 1$, where $(x, y) = (x', y') \mod 1$ if there exist integers k and l such that $x - x' = k$ and $y - y' = l$. Denote equivalence classes mod 1 by the symbol $[(x, y)]$. Consider the curve $\alpha : \mathbb{R} \to T^2$ defined by $\alpha(t) = [(at, bt)]$. This map is an immersion unless $a = b = 0$. If a/b is a rational number it is not an embedding since the curve eventually passes through $(1, 1) = (0, 0)$ for some t and α is not injective. For a/b irrational the curve never passes through any point twice and is therefore an embedding. Figure 15.5 illustrates these properties. When a/b is rational the image $C = \alpha(\mathbb{R})$ has the relative topology in T^2 of a circle. Hence there is an embedding $\beta : S^1 \to C$, making (S^1, β) an embedded submanifold of T^2. It is left to the reader to explicitly construct the map β. In this case the subset $C = \beta(S^1) \subset T^2$ is closed.

The set $\alpha(\mathbb{R})$ is dense in T^2 when a/b is irrational, since the curve eventually passes arbitrarily close to any point of T^2, and cannot be a closed subset. Hence the relative topology on $\alpha(\mathbb{R})$ induced on it as a subset of T^2 is much coarser than the topology it would obtain from $\mathbb{R}$ through the bijective map α. The embedding α is therefore not a homeomorphism from $\mathbb{R}$ to $\alpha(\mathbb{R})$ when the latter is given the relative topology.

In general, an embedding $\alpha : M \to N$ that is also a homeomorphism from M to $\alpha(M)$ when the latter is given the relative topology in N is called a **regular embedding**. A necessary and sufficient condition for this to hold is that there be a coordinate chart $(U, \phi; x^i)$

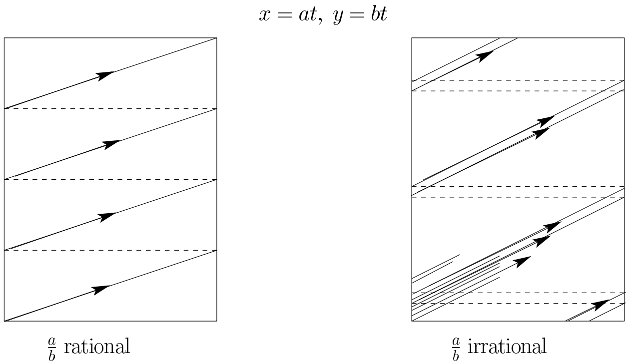

Figure 15.5 Submanifolds of the torus

at every point $p \in \alpha(M)$ such that $\alpha(M) \cap U$ is defined by the 'coordinate slice'

$$x^{m+1} = x^{m+2} = \cdots = x^n = 0.$$

It also follows that the set $\alpha(M)$ must be a closed subset of N for this to occur. The proofs of these statements can be found in [11]. The above embedded submanifold (S^1, β) is a regular embedding when a/b is rational.

Problems

Problem 15.13 Show that if $\rho_p = r_i(\mathrm{d}x^i)_p = \alpha^* \omega_{\alpha(p)}$ then the components are given by

$$r_i = \frac{\partial y^a}{\partial x^i}\bigg|_{\phi(p)} w_a$$

where $\omega_{\alpha(p)} = w_a(\mathrm{d}y^a)_{\alpha(p)}$.

If α is a diffeomorphism, define a map $\alpha_* : T^{(1,1)}_p(M) \to T^{(1,1)}_{\alpha(p)}(N)$ by setting

$$\alpha_* T\big(\omega_{\alpha(p)}, X_{\alpha(p)}\big) = T\big(\alpha^* \omega_{\alpha(p)}, \alpha_*^{-1} X_{\alpha(p)}\big)$$

and show that the components transform as

$$(\alpha_* T)^a{}_b = T^i{}_j \frac{\partial y^a}{\partial x^i} \frac{\partial x^j}{\partial y^b}.$$

Problem 15.14 If $\gamma : \mathbb{R} \to M$ is a curve on M and $p = \gamma(t_0)$ and $\alpha : M \to N$ is a differentiable map show that

$$\alpha_* \dot{\gamma}_p = \dot{\sigma}_{\alpha(p)} \quad \text{where} \quad \sigma = \alpha \circ \gamma : \mathbb{R} \to N.$$

Problem 15.15 Is the map $\alpha : \mathbb{R} \to \mathbb{R}^2$ given by $x = \sin t,\ y = \sin 2t$ (i) an immersion, (ii) an embedded submanifold?

Problem 15.16 Show that the map $\alpha : \dot{\mathbb{R}}^2 \to \mathbb{R}^3$ defined by

$$u = x^2 + y^2, \qquad v = 2xy, \qquad w = x^2 - y^2$$

is an immersion. Is it an embedded submanifold?

Evaluate $\alpha^*(u\mathrm{d}u + v\mathrm{d}v + w\mathrm{d}w)$ and $\alpha_*(\partial_x)_{(a,b)}$. Find a vector field X on $\dot{\mathbb{R}}^2$ for which $\alpha_* X$ is not a well-defined vector field.

15.5 Commutators, flows and Lie derivatives

Commutators

Let X and Y be smooth vector fields on an open subset U of a differentiable manifold M. We define their **commutator** or **Lie bracket** $[X, Y]$ as the vector field on U defined by

$$[X, Y]f = X(Yf) - Y(Xf) \tag{15.20}$$

for all differentiable functions f on U. This is a vector field since (i) it is linear

$$[X, Y](af + bg) = a[X, Y]f + b[x, y]g$$

for all $f, g \in \mathcal{F}(U)$ and $a, b \in \mathbb{R}$, and (ii) it satisfies the Leibniz rule

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

Linearity is trivial, while the Leibniz rule follows from

$$\begin{aligned}[X, Y](fg) &= X(fYg + gYf) - Y(fXg + gXf)\\ &= Xf\,Yg + fX(Yg) + Xg\,Yf + gX(Yf)\\ &\quad - Yf\,Xg - fY(Xg) - Yg\,Xf - gY(Xf)\\ &= f[X, Y]g + g[X, Y]f.\end{aligned}$$

A number of identities are easily verified for the Lie bracket:

$$[X, Y] = -[Y, X], \tag{15.21}$$

$$[X, aY + bZ] = a[X, Y] + b[X, Z], \tag{15.22}$$

$$[X, fY] = f[X, Y] + Xf\,Y, \tag{15.23}$$

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0. \tag{15.24}$$

Equations (15.21) and (15.22) are trivial, and (15.23) follows from

$$\begin{aligned}[X, fY]g = X(fYg) - fY(Xg) &= fX(Yg) + X(f)Yg - fY(Xg)\\ &= f[X, Y]g + Xf\,Yg.\end{aligned}$$

The **Jacobi identity** (15.24) is proved much as for commutators in matrix theory, Example 6.7.

Exercise: Show that for any functions f, g and vector fields X, Y

$$[fX, gY] = fg[X, Y] + fXg\,Y - gYf\,X.$$

To find a coordinate formula for the Lie product, let $X = X^i(x^1, \dots, x^n)\partial_{x^i}$, $Y = Y^i(x^1, \dots, x^n)\partial_{x^i}$. Then $[X, Y] = [X, Y]^k(x^1, \dots, x^n)\partial_{x^k}$, where

$$[X, Y]^k = [X, Y](x^k) = X(Yx^k) - Y(Xx^k) = X^i\frac{\partial Y^k}{\partial x^i} - Y^i\frac{\partial X^k}{\partial x^i},$$

or in the comma derivative notation

$$[X, Y]^k = X^i Y^k{}_{,i} - Y^i X^k{}_{,i}. \tag{15.25}$$

If we regard the vector field X as acting on the vector field Y by the Lie bracket to produce a new vector field, $X : Y \mapsto \mathcal{L}_X Y = [X, Y]$, this action is remarkably 'derivative-like' in that it is both linear

$$\mathcal{L}_X(aY + bZ) = a\mathcal{L}_X Y + b\mathcal{L}_X Z$$

and has the property

$$\mathcal{L}_X(fY) = Xf\, Y + f\mathcal{L}_X Y. \tag{15.26}$$

These properties follow immediately from (15.22) and (15.23). A geometrical interpretation of this derivative will appear in terms of the concept of a *flow* induced by a vector field.

Integral curves and flows

Let X be a smooth vector field on a manifold M. An **integral curve** of X is a parametrized curve $\sigma : (a, b) \to M$ whose tangent vector $\dot\sigma(t)$ at each point $p = \sigma(t)$ on the curve is equal to the tangent vector X_p assigned to p,

$$\dot\sigma(t) = X_{\sigma(t)}.$$

In a local coordinate chart $(U; x^i)$ at p where the curve can be written as n real functions $x^i(t) = x^i(\sigma(t))$ and the vector field has the form $X = X^i(x^1, \dots, x^n)\partial_{x^i}$, this requirement appears as n ordinary differential equations,

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = X^i\big(x^1(t), \dots, x^n(t)\big). \tag{15.27}$$

The existence and uniqueness theorem of ordinary differential equations asserts that through each point $p \in M$ there exists a unique maximal integral curve $\gamma_p : (a, b) \to M$ such that $a = a(p) < 0 < b = b(p)$ and $p = \gamma_p(0)$ [15, 16]. Uniqueness means that if $\sigma : (c, d) \to M$ is any other integral curve passing through p at $t = 0$ then $a \le c < 0 < d \le b$ and $\sigma = \gamma_p\big|_{(c,d)}$.

By a **transformation** of the manifold M is meant a diffeomorphism $\varphi : M \to M$. A **one-parameter group of transformations** of M, or on M, is a map $\sigma : \mathbb{R} \times M \to M$ such that:

(i) for each $t \in \mathbb{R}$ the map $\sigma_t : M \to M$ defined by $\sigma_t(p) = \sigma(t, p)$ is a transformation of M;
(ii) for all $t, s \in \mathbb{R}$ we have the abelian group property, $\sigma_{t+s} = \sigma_t \circ \sigma_s$.

Since the maps σ_t are one-to-one and onto, every point $p \in M$ is the image of a unique point $q \in M$; that is, we can write $p = \sigma_t(q)$ where $q = \sigma_t^{-1}(p)$. Hence σ_0 is the identity transformation, $\sigma_0 = \mathrm{id}_M$ since $\sigma_0(p) = \sigma_0 \circ \sigma_t(q) = \sigma_t(q) = p$ for all $p \in M$. Furthermore, the inverse of each map σ_t^{-1} is σ_{-t} since $\sigma_t \circ \sigma_{-t} = \sigma_0 = \mathrm{id}_M$.

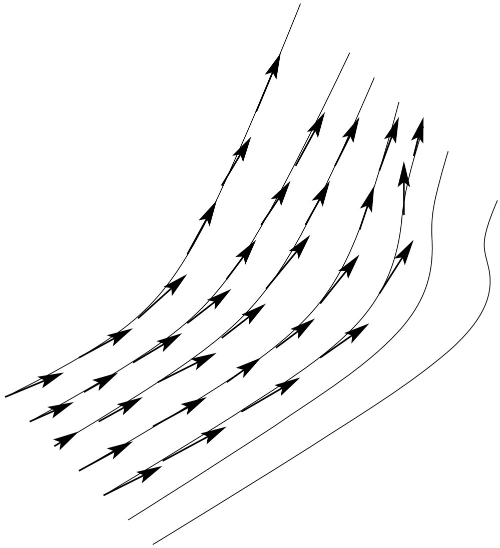

Figure 15.6 Streamlines representing the flow generated by a vector field

The curve $\gamma_p : \mathbb{R} \to M$ defined by $\gamma_p(t) = \sigma_t(p)$ clearly passes through p at $t = 0$. It is called the **orbit** of p under the flow σ and defines a tangent vector X_p at p by

$$X_p f = \frac{\mathrm{d} f(\gamma_p(t))}{\mathrm{d}t}\Big|_{t=0} = \frac{\mathrm{d} f(\sigma_t(p))}{\mathrm{d}t}\Big|_{t=0}.$$

Since p is an arbitrary point of M we have a vector field X on M, said to be the **vector field induced** by the flow σ. Any vector field X induced by a one-parameter group of transformations of M is said to be **complete**. The one-parameter group σ_t can be thought of as 'filling in' the vector field X with a set of curves, which play the role of streamlines for a fluid whose velocity is everywhere given by X (see Fig. 15.6).

Not every vector field is complete, but there is a local concept that is always applicable. A **local one-parameter group of transformations**, or **local flow**, consists of an open subset $U \subseteq M$ and a real interval $I_\epsilon = (-\epsilon, \epsilon)$, together with a map $\sigma : I_\epsilon \times U \to M$ such that:

(i′) for each $t \in I_\epsilon$ the map $\sigma_t : U \to M$ defined by $\sigma_t(p) = \sigma(t, p)$ is a diffeomorphism of U onto $\sigma_t(U)$;

(ii′) if t, s and $t + s \in I_\epsilon$ and $p, \sigma_s(p) \in U$ then $\sigma_{t+s}(p) = \sigma_t\big(\sigma_s(p)\big)$.

A local flow induces a vector field X on U in a similar way to that described above for a flow:

$$X_p f = \frac{\mathrm{d} f(\sigma_t(p))}{\mathrm{d}t}\Big|_{t=0} \qquad \text{for all } p \in U. \tag{15.28}$$

It now turns out that every vector field X corresponds to a local one-parameter group of transformations, which it may be said to **generate**.

Theorem 15.2 *If X is a vector field on M, and $p \in M$ then there exists an interval $I = (-\epsilon, \epsilon)$, a neighbourhood U of p, and a local flow $\sigma : I \times U \to M$ that induces the vector field $X\big|_U$ restricted to U.*

Proof: If $(U, \phi; x^i)$ is a coordinate chart at p we may set

$$X\big|_U = X^i \frac{\partial}{\partial x^i} \quad \text{where} \quad X^i : U \to \mathbb{R}.$$

The existence and uniqueness theorem of ordinary differential equations implies that for any $\mathbf{x} \in \phi(U)$ there exists a unique curve $\mathbf{y} = \mathbf{y}(t; x^1, \dots, x^n)$ on some interval $I = (-\epsilon, \epsilon)$ such that

$$\frac{\mathrm{d}y^i(t; x^1, \dots, x^n)}{\mathrm{d}t} = X^i \circ \phi^{-1}\big(y^1(t; \mathbf{x}), y^2(t; \mathbf{x}), \dots, y^n(t; \mathbf{x})\big)$$

and

$$y^i(0; x^1, \dots, x^n) = x^i.$$

As the solutions of a family of differential equations depend smoothly on the initial coordinates [15, 16], the functions $y(t; x^1, \dots, x^n)$ are differentiable with respect to t and x^i.

For fixed s and fixed $\mathbf{x} \in \phi(U)$ the curves $t \to z^i(t, s; \mathbf{x}) = y^i(t; \mathbf{y}(s; \mathbf{x}))$ and $t \to z(t, s, \mathbf{x}) = y^i(t + s; \mathbf{x})$ satisfy the same differential equation

$$\frac{\mathrm{d}z^i(t, s; x^1, \dots, x^n)}{\mathrm{d}t} = X^i\big(z^1(t, s; \mathbf{x}), \dots, z^n(t, s; \mathbf{x})\big)$$

and have the same initial conditions at $t = 0$,

$$y^i(0; \mathbf{y}(s; \mathbf{x})) = y^i(s; \mathbf{x}) = y^i(0 + s; \mathbf{x}).$$

These solutions are therefore identical and the map $\sigma : I \times U \to U$ defined by $\sigma(t, p) = \phi^{-1}\mathbf{y}\big(t; \phi(p)\big)$ satisfies the local one-parameter group condition

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

■

A useful consequence of this theorem is the local existence of a coordinate system that 'straightens out' any given vector field X so that its components point along the 1-axis, $X^i = (1, 0, \dots, 0)$. The local flow σ_t generated by X is then simply a translation in the 1-direction, $\sigma_t(x^1, x^2, \dots, x^n) = (x^1 + t, x^2, \dots, x^n)$.

Theorem 15.3 *If X is a vector field on a manifold M such that $X_p \neq 0$, then there exists a coordinate chart $(U, \phi; x^i)$ at p such that*

$$X = \frac{\partial}{\partial x^1}. \tag{15.29}$$

Outline proof: The idea behind the proof is not difficult. Pick any coordinate system $(U, \psi; y^i)$ at p such that $y^i(p) = 0$, and $X_p = \big(\partial_{y^1}\big)_p$. Let $\sigma : I_\epsilon \times A \to M$ be a local flow that induces X on the open set A. In a neighbourhood of p consider a small $(n-1)$-dimensional 'open ball' of points through p that cuts across the flow, whose typical point q has coordinates $(0, y^2, \dots, y^n)$, and assign coordinates $(x^1 = t, x^2 = y^2, \dots, x^n = y^n)$

to points on the streamline $\sigma_t(q)$ through q. The coordinates $x^2, \dots, x^n$ are then constant along the curves $t \to \sigma_t(q)$, and the vector field X, being tangent to the streamlines, has coordinates $(1, 0, \dots, 0)$ throughout a neighbourhood of p. A detailed proof may be found in [11, theorem 4.3] or [4, p. 124]. ■

Example 15.13 Let X be the differentiable vector field $X = x^2\partial_x$ on the real line manifold $\mathbb{R}$. To find a coordinate $y = y(x)$ such that $X = \partial_y$, we need to solve the differential equation

$$x^2 \frac{\partial y}{\partial x} = 1.$$

The solution is $y = C - 1/x$.

The local one-parameter group generated by X is found by solving the ordinary differential equation,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x^2.$$

The solution is

$$\sigma_t(x) = \frac{1}{x^{-1} - t} = \frac{x}{1 - tx}.$$

It is straightforward to verify the group property

$$\sigma_t\big(\sigma_s(x)\big) = \frac{x}{1 - (t + s)x} = \sigma_{t+s}(x).$$

Example 15.14 If X and Y are vector fields on M generating flows ϕ_t and ψ_t respectively, let σ be the curve through $p \in M$ defined by

$$\sigma(t) = \psi_{-t} \circ \phi_{-t} \circ \psi_t \circ \phi_t p.$$

Then $\sigma(\sqrt{t})$ is a curve whose tangent vector is the commutator $[X, Y]$ at p. The proof is to let f be any differentiable function at p and show that

$$[X, Y]_p f = \lim_{t \to 0} \frac{f\big[\sigma(\sqrt{t})\big] - f\big[\sigma(0)\big]}{t}.$$

Details may be found in [3, p. 130]. Some interesting geometrophysical applications of this result are discussed in [17].

Lie derivative

Let X be a smooth vector field on a manifold M, which generates a local one-parameter group of transformations σ_t on M. If Y is any differentiable vector field on M, we define its **Lie derivative** along X to be

$$\mathcal{L}_X Y = \lim_{t \to 0} \frac{Y - (\sigma_t)_* Y}{t}. \tag{15.30}$$

Figure 15.7 illustrates the siutation. Essentially, the tangent map of the diffeomorphism σ_t is used to 'drag' the vector field Y forward along the integral curves from a point $\sigma_{-t}(p)$ to p and the result is compared with original value Y_p of the vector field. Equation (15.7)

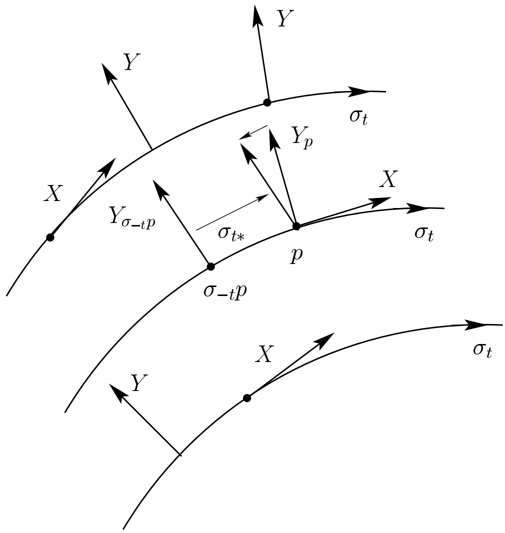

Figure 15.7 Lie derivative of a vector field Y along a vector field X

performs this operation for neighbouring points and takes the limit on dividing by t. We now show that this derivative is identical with the 'derivative-like' operation of taking the commutator of two vector fields.

Let $f : M \to \mathbb{R}$ be a differentiable function, and p any point of M. From (15.30) we have at p,

$$\begin{aligned}(\mathcal{L}_X Y)_p f &= \lim_{t\to 0} \frac{1}{t}\big(Y_p f - ((\sigma_t)_* Y)_p f\big)\\ &= \lim_{t\to 0} \frac{1}{t}\big(Y_p f - Y_{\sigma_{-t}(p)}(f \circ \sigma_t)\big)\\ &= \lim_{t\to 0} \frac{1}{t}\big(Y_p f - Y_{\sigma_{-t}(p)} f - Y_{\sigma_{-t}(p)}(f \circ \sigma_t - f)\big)\\ &= \lim_{t\to 0} \frac{1}{t}\big((Yf - (Yf)\circ\sigma_{-t})(p) - Y_{\sigma_{-t}(p)}(f \circ \sigma_t - f)\big).\end{aligned}$$

On setting $s = -t$ in the first term and using Eq. (15.28), the right-hand side reduces to $X(Yf)(p) - Y(Xf)(p) = [X, Y]_p f$, and we have the desired relation

$$\mathcal{L}_X Y = [X, Y]. \tag{15.31}$$

The concept of Lie derivative can be extended to all tensor fields. First, for any diffeomorphism $\varphi : M \to M$, we define the induced map $\widetilde{\varphi} : \mathcal{T}^{(r,s)}(M) \to \mathcal{T}^{(r,s)}(M)$ in the following way:

(i) for vector fields set $\widetilde{\varphi} = \varphi_*$;
(ii) for scalar fields $f : M \to \mathbb{R}$ set $\widetilde{\varphi} f = f \circ \varphi^{-1}$;
(iii) for covector fields set $\widetilde{\varphi} = \left(\varphi^{-1}\right)^*$;

(iv) the map $\widetilde{\varphi}$ is extended to all tensor fields by demanding linearity and

$$\widetilde{\varphi}(T \otimes S) = \widetilde{\varphi}T \otimes \widetilde{\varphi}S$$

for arbitrary tensor fields T and S.

If ω and X are arbitrary covector and vector fields, then

$$\langle \widetilde{\varphi}\omega, \widetilde{\varphi}X \rangle = \widetilde{\varphi}\langle \omega, X \rangle, \tag{15.32}$$

since

$$\begin{aligned} \langle \widetilde{\varphi}\omega, \widetilde{\varphi}X \rangle(p) &= \langle (\varphi^{-1})^* \omega_{\varphi^{-1}(p)}, \varphi_* X_{\varphi^{-1}(p)} \rangle \\ &= \langle \omega_{\varphi^{-1}(p)}, X_{\varphi^{-1}(p)} \rangle = \langle \omega, X \rangle \big(\varphi^{-1}(p)\big). \end{aligned}$$

Exercise: For arbitrary vector fields X show from (ii) that $\widetilde{\varphi}X(\widetilde{\varphi}f) = \widetilde{\varphi}\big(X(f)\big)$.

Using Eq. (15.11), property (iv) provides a unique definition for the application of the map $\widetilde{\varphi}$ to all higher order tensors. Alternatively, as for covector fields, the following is a characterization of the map $\widetilde{\varphi}$:

$$(\widetilde{\varphi}T)\big(\widetilde{\varphi}\omega^1, \dots, \widetilde{\varphi}\omega^r, \widetilde{\varphi}X_1, \dots, \widetilde{\varphi}X_s\big) = \widetilde{\varphi}(T(\omega^1, \dots, \omega^r, X_1, \dots, X_s))$$

for all vector fields $X_1, \dots X_s$ and covector fields $\omega^1, \dots, \omega^r$.

The **Lie derivative** $\mathcal{L}_X T$ of a smooth tensor field T with respect to the vector field X is defined as

$$\mathcal{L}_X T = \lim_{t \to 0} \frac{1}{t}\big(T - \widetilde{\sigma}_t T\big). \tag{15.33}$$

Exercise: Show that for any tensor field T

$$\mathcal{L}_X T = -\frac{\mathrm{d}\,\widetilde{\sigma}_t T}{\mathrm{d}t}\bigg|_{t=0} \tag{15.34}$$

and prove the Leibnitz rule

$$\mathcal{L}_X(T \otimes S) = T \otimes (\mathcal{L}_X S) + (\mathcal{L}_X T) \otimes S. \tag{15.35}$$

When T is a scalar field f, we find, on changing the limit variable to $s = -t$,

$$\big(\mathcal{L}_X f\big)_p = \frac{\mathrm{d}\, f \circ \sigma_s(p)}{\mathrm{d}s}\bigg|_{s=0} = Xf(p),$$

and in a local coordinate chart $(U; x^i)$

$$\mathcal{L}_X f = Xf = f_{,i} X^i. \tag{15.36}$$

Since for any pair i, j

$$\mathcal{L}_{\partial_{x^i}} \frac{\partial}{\partial x^j} = \left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = \frac{\partial^2}{\partial x^i \partial x^j} - \frac{\partial^2}{\partial x^j \partial x^i} = 0,$$

and $\mathcal{L}_X Y = [X, Y] = -[Y, X] = -\mathcal{L}_Y X$ for any pair of vector fields X, Y, we find

$$\mathcal{L}_X \frac{\partial}{\partial x^j} = -\mathcal{L}_{\partial_{x^j}}\left(X^i \frac{\partial}{\partial x^i}\right) = -X^i{}_{,j} \frac{\partial}{\partial x^i}.$$

Applying the Leibnitz rule (15.35) results in

$$\mathcal{L}_X Y = \mathcal{L}_X\Big(Y^i \frac{\partial}{\partial x^i}\Big) = Y^i{}_{,j} X^j \frac{\partial}{\partial x^i} - Y^j X^i{}_{,j}\frac{\partial}{\partial x^i},$$

in agreement with the component formula for the Lie bracket in Eq. (15.25),

$$\big(\mathcal{L}_X Y\big)^i = Y^i{}_{,j} X^j - Y^j X^i{}_{,j}. \tag{15.37}$$

To find the component formula for the Lie derivative of a 1-form $\omega = w_i \mathrm{d}x^i$, we note that for any pair of vector fields $X,\ Y$

$$\mathcal{L}_X\langle\omega, Y\rangle = X\langle\omega, Y\rangle = \langle\mathcal{L}_X\omega, Y\rangle + \langle\omega, \mathcal{L}_X Y\rangle, \tag{15.38}$$

which follows from Eqs. (15.32) and (15.34),

$$\begin{aligned}\mathcal{L}_X\langle\omega, Y\rangle = X\langle\omega, Y\rangle &= -\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\sigma}_t\langle\omega, Y\rangle\Big|_{t=0}\\ &= -\frac{\mathrm{d}}{\mathrm{d}t}\langle\widetilde{\sigma}_t\omega, \widetilde{\sigma}_t Y\rangle\Big|_{t=0}\\ &= -\langle\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\sigma}_t\omega, Y\rangle\Big|_{t=0} - \langle\omega, \frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\sigma}_t Y\rangle\Big|_{t=0}\\ &= \langle\mathcal{L}_X\omega, Y\rangle + \langle\omega, \mathcal{L}_X Y\rangle.\end{aligned}$$

If $\omega = w_i\, \mathrm{d}x^i$ is a 1-form, then its Lie derivative $\mathcal{L}_X\omega$ with respect to the vector field X has components in a coordinate chart $(U; x^i)$ given by

$$\begin{aligned}\big(\mathcal{L}_X\omega\big)_j &= \langle\mathcal{L}_X\omega, \frac{\partial}{\partial x^j}\rangle\\ &= \mathcal{L}_X\langle\omega, \frac{\partial}{\partial x^j}\rangle - \langle\omega, \mathcal{L}_X\frac{\partial}{\partial x^j}\rangle\\ &= \mathcal{L}_X w_j + \langle\omega, X^i{}_{,j}\frac{\partial}{\partial x^i}\rangle\\ &= w_{j,i}X^i + w_i X^i{}_{,j}.\end{aligned}$$

Extending this argument to a general tensor of type (r, s), we find

$$\begin{aligned}\big(\mathcal{L}_X T\big)^{ij\dots}{}_{kl\dots} = T^{ij\dots}{}_{kl\dots,m}X^m - T^{mj\dots}{}_{kl\dots}X^i{}_{,m} - T^{im\dots}{}_{kl\dots}X^j{}_{,m} - \cdots\\ T^{ij\dots}{}_{ml\dots}X^m{}_{,k} + T^{ij\dots}{}_{km\dots}X^m{}_{,l} + \cdots\end{aligned} \tag{15.39}$$

Example 15.15 In local coordinates such that $X = \partial_{x^1}$ (see Theorem 15.3), all $X^i{}_{,j} = 0$ since the components X^i = consts. and the components of the Lie derivative are simply the derivatives in the 1-direction,

$$\big(\mathcal{L}_X T\big)^{ij\dots}{}_{kl\dots} = T^{ij\dots}{}_{kl\dots,1}.$$

Problems

Problem 15.17 Show that the components of the Lie product $[X, Y]^k$ given by Eq. (15.25) transform as a contravariant vector field under a coordinate transformation $x'^j(x^i)$.

Problem 15.18 Show that the Jacobi identity can be written

$$\mathcal{L}_{[X,Y]}Z = \mathcal{L}_X\mathcal{L}_Y Z - \mathcal{L}_Y\mathcal{L}_X Z,$$

and this property extends to all tensors T:

$$\mathcal{L}_{[X,Y]}T = \mathcal{L}_X\mathcal{L}_Y T - \mathcal{L}_Y\mathcal{L}_X T.$$

Problem 15.19 Let $\alpha : M \to N$ be a diffeomorphism between manifolds M and N and X a vector field on M that generates a local one-parameter group of transformations σ_t on M. Show that the vector field $X' = \alpha_* X$ on N generates the local flow $\sigma'_t = \alpha \circ \sigma_t \circ \alpha^{-1}$.

Problem 15.20 For any real positive number n show that the vector field $X = x^n \partial_x$ is differentiable on the manifold $\mathbb{R}^+$ consisting of the positive real line $\{x \in \mathbb{R} \mid x > 0\}$. Why is this not true in general on the entire real line $\mathbb{R}$? As done for the case $n = 2$ in Example 15.13, find the maximal one-parameter subgroup σ_t generated by this vector field at any point $x > 0$.

Problem 15.21 On the manifold $\mathbb{R}^2$ with coordinates (x, y), let X be the vector field $X = -y\partial_x + x\partial_y$. Determine the integral curve through any point (x, y), and the one-parameter group generated by X. Find coordinates (x', y') such that $X = \partial_{x'}$.

Problem 15.22 Repeat the previous problem for the vector fields, $X = y\partial_x + x\partial_y$ and $X = x\partial_x + y\partial_y$.

Problem 15.23 On a compact manifold show that every vector field X is complete. [*Hint*: Let σ_t be a local flow generating X, and let ϵ be the least bound required on a finite open covering. Set $\sigma_t = (\sigma_{t/N})^N$ for N large enough that $|t| < \epsilon N$.]

Problem 15.24 Show that the Lie derivative $\mathcal{L}_X$ commutes with all operations of contraction C^i_j on a tensor field T,

$$\mathcal{L}_X C^i_j T = C^i_j \mathcal{L}_X T.$$

Problem 15.25 Prove the formula (15.39) for the Lie derivative of a general tensor.

15.6 Distributions and Frobenius theorem

A k**-dimensional distribution** D^k on a manifold M is an assignment of a k-dimensional subspace $D^k(p)$ of the tangent space $T_p(M)$ at every point $p \in M$. The distribution is said to be C^∞ or **smooth** if for all $p \in M$ there is an open neighbourhood U and k smooth vector fields $X_1, \dots, X_k$ on U that span $D^k(q)$ at each point $q \in U$. A vector field X on an open domain A is said to **lie in** or **belong to** the distribution D^k if $X_p \in D^k(p)$ at each point $p \in A$. A one-dimensional distribution is equivalent to a vector field up to an arbitrary scalar factor at every point, and is sometimes called a *direction field*.

An **integral manifold** of a distribution D^k is a k-dimensional submanifold (K, ψ) of M such that all vector fields tangent to the submanifold belong to D^k,

$$\psi_*(T_p(K)) = D^k(\psi(p)).$$

Every one-dimensional distribution has integral manifolds, for if X is any vector field that spans a distribution D^1 then any family of integral curves of X act as integral manifolds of the distribution D^1. We will see, however, that not every distribution of higher dimension has integral manifolds.

A distribution D^k is said to be **involutive** if for any pair of vector fields X, Y lying in D^k, their Lie bracket $[X, Y]$ also belongs to D^k. If $\{e_1, \dots, e_k\}$ is any local basis of vector fields spanning an involutive D^k on an open neighbourhood U, then

$$[e_\alpha, e_\beta] = \sum_{\gamma=1}^{k} C^\gamma_{\alpha\beta} e_\gamma \quad (\alpha, \beta = 1, \dots, k) \tag{15.40}$$

where $C^\gamma_{\alpha\beta} = -C^\gamma_{\beta\alpha}$ are C^∞ functions on U. Conversely, if there exists a local basis $\{e_\alpha\}$ satisfying (15.40) for some scalar structure fields $C^\gamma_{\alpha\beta}$, the distribution is involutive, for if $X = X^\alpha e_\alpha$ and $Y = Y^\beta e_\beta$ then

$$[X, Y] = [X^\alpha e_\alpha,\ Y^\beta e_\beta] = \big(X(Y^\gamma) - Y(X^\gamma) + X^\alpha Y^\beta C^\gamma_{\alpha\beta}\big)e_\gamma,$$

which belongs to D^k as required. For example, if there exists a coordinate chart $(U; x^i)$ such that the distribution D^k is spanned by the first k coordinate basis vector fields

$$e_1 = \partial_{x^1},\ e_2 = \partial_{x^2}, \dots,\ e_k = \partial_{x^k}$$

then D^k is involutive on U since all $[e_\alpha, e_\beta] = 0$, a trivial instance of the relation (15.40). In this case we can restrict the chart to a cubical neighbourhood $U' = \{p \mid -a < x^i(p) < a\}$, and the 'slices' $x^a = \text{const.}$ $(a = k + 1, \dots, n)$ are local integral manifolds of the distribution D^k. The key result is the **Frobenius theorem**:

Theorem 15.4 *A smooth k-dimensional distribution D^k on a manifold M is involutive if and only if every point $p \in M$ lies in a coordinate chart $(U; x^i)$ such that the coordinate vector fields $\partial/\partial x^\alpha$ for $\alpha = 1, \dots, k$ span D^k at each point of U.*

Proof: The *if* part follows from the above remarks. The converse will be shown by induction on the dimension k. The case $k = 1$ follows immediately from Theorem 15.3. Suppose now that the statement is true for all $(k-1)$-dimensional distributions, and let D^k be a k-dimensional involutive distribution spanned at all points of an open set A by vector fields $\{X_1, \dots, X_k\}$. At any point $p \in A$ there exist coordinates $(V; y^i)$ such that $X_k = \partial_{y^k}$. Set

$$Y_\alpha = X_\alpha - (X_\alpha y^k)X_k, \quad Y_k = X_k$$

where Greek indices $\alpha, \beta, \dots$ range from 1 to $k - 1$. The vector fields $Y_1, Y_2, \dots, Y_k$ clearly span D^k on V, and

$$Y_\alpha y^k = 0, \qquad Y_k y^k = 1\,. \tag{15.41}$$

Since D^k is an involutive we can write

$$[Y_\alpha,\ Y_\beta] = C^\gamma_{\alpha\beta} Y_\gamma + a_{\alpha\beta} Y_k,$$
$$[Y_\alpha,\ Y_k] = C^\gamma_\alpha Y_\gamma + a_\alpha Y_k.$$

Applying both sides of these equations to the coordinate function y^k and using (15.41), we find $a_{\alpha\beta} = a_\alpha = 0$, whence

$$[Y_\alpha,\ Y_\beta] = C^\gamma_{\alpha\beta} Y_\gamma, \tag{15.42}$$

$$[Y_\alpha,\ Y_k] = C^\gamma_\alpha Y_\gamma. \tag{15.43}$$

The distribution D^{k-1} spanned by $Y_1,\ Y_2, \dots, Y_{k-1}$ is therefore involutive on V, and by the induction hypothesis there exists a coordinate chart $(W; z^i)$ such that D^{k-1} is spanned by $\{\partial_{z^1}, \dots,\ \partial_{z^{k-1}}\}$. Set

$$\frac{\partial}{\partial z^\alpha} = A^\beta_\alpha Y_\beta$$

where $[A^\beta_\alpha]$ is a non-singular matrix of functions on W. The original distribution D^k is spanned on W by the set of vector fields

$$\{\partial_{z^1},\ \partial_{z^2}, \dots,\ \partial_{z^{k-1}},\ Y_k\}.$$

It follows then from (15.43) that

$$[\partial_{z^\alpha},\ Y_k] = K^\beta_\alpha \partial_{z^\beta} \tag{15.44}$$

for some functions K^β_α. If we write

$$Y_k = \sum_{\alpha=1}^{k-1} \xi^\alpha \partial_{z^\alpha} + \sum_{a=k}^{n} \xi^a \partial_{z^a}$$

and apply Eq. (15.44) to the coordinate functions z^a $(a = k, \dots, n)$, we find

$$\frac{\partial \xi^a}{\partial z^\alpha} = 0.$$

Hence $\xi^a = \xi^a(z^k, \dots, z^n)$ for all $a \geq k$. Since Y_k is linearly independent of the vectors ∂_{z^α}, the distribution D^k is spanned by the set of vectors $\{\partial_{z^1},\ \partial_{z^2}, \dots,\ \partial_{z^{k-1}},\ Z\}$, where

$$Z = Y_k - \xi^\alpha \partial_{z^\alpha} = \xi^a(z^k, \dots, z^n)\partial_{z^a}.$$

By Theorem 15.3 there exists a coordinate transformation not involving the first $(k-1)$ coordinates,

$$x^k = x^k(z^k, \dots, z^n),\ x^{k+1} = x^{k+1}(z^k, \dots, z^n), \dots,\ x^n = x^n(z^k, \dots, z^n)$$

such that $Z = \partial_{x^k}$. Setting $x^1 = z^1, \dots,\ x^{k-1} = z^{k-1}$, we have coordinates $(U; x^i)$ in which D^k is spanned by $\{\partial_{x^1}, \dots,\ \partial_{x^{k-1}},\ \partial_{x^k}\}$. ■

Theorem 15.5 *A set of vector fields $\{X_1, X_2, \dots, X_k\}$ is equal to the first k basis fields of a local coordinate system, $X_1 = \partial_{x^1}, \dots, X_k = \partial_{x^k}$ if and only if they commute with each other, $[X_\alpha, X_\beta] = 0$.*

Proof: The vanishing of all commutators is clearly a necessary condition for the vector fields to be local basis fields of a coordinate system, for if $X_\alpha = \partial_{x^\alpha}$ $(\alpha = 1, \dots, r)$ then $[X_\alpha, X_\beta] = [\partial_{x^\alpha}, \partial_{x^\beta}] = 0$.

To prove sufficiency, we again use induction on k. The case $k = 1$ is essentially Theorem 15.3. By the induction hypothesis, there exists local coordinates $(U; x^i)$ such that $X_\alpha = \partial_{x^\alpha}$ for $\alpha = 1, \dots, k-1$. Set $Y = X_k = Y^i(x^1, \dots, x^n)\partial_{x^i}$, and by Example 15.15 $Y^i{}_{,\alpha} = 0$, so that we may write

$$Y = \sum_{\alpha=1}^{k-1} Y^\alpha(x^k, \dots, x^n)\partial_{x^\alpha} + \sum_{a=k}^{n} Y^a(x^k, \dots, x^n)\partial_{x^a}.$$

Using Theorem 15.3 we may perform a coordinate transformation on the last $n - k + 1$ coordinates such that

$$Y = \sum_{\alpha=1}^{k-1} Y^\alpha(x^k, \dots, x^n)\partial_{x^\alpha} + \partial_{x^k}.$$

A coordinate transformation

$$\begin{aligned} x'^\alpha &= x^\alpha + f^\alpha(x^k, \dots, x^n) \quad && (\alpha = 1, \dots, k-1) \\ x'^a &= x^a && (a = k, \dots, n) \end{aligned}$$

has the effect

$$Y = \sum_{\beta=1}^{k-1}\Big(Y^\beta + \frac{\partial f^\beta}{\partial x^k}\Big)\frac{\partial}{\partial x'^\beta} + \frac{\partial}{\partial x'^k}.$$

Solving the differential equations

$$\frac{\partial f^\beta}{\partial x^k} = -Y^\beta(x^k, \dots, x^n)$$

by a straightforward integration leads to $Y = \partial_{x'^k}$ as required. ■

Example 15.16 On $\dot{\mathbb{R}}^3 = \mathbb{R}^3 - \{(0, 0, 0)\}$ let X_1, X_2, X_3 be the three vector fields

$$X_1 = y\partial_z - z\partial_y, \qquad X_2 = z\partial_x - x\partial_z, \qquad X_3 = x\partial_y - y\partial_x.$$

These three vector fields generate a two-dimensional distribution D^2, as they are not linearly independent

$$xX_1 + yY_2 + zX_3 = 0.$$

The Lie bracket of any pair of these vector fields is easily calculated,

$$\begin{aligned} [X_1, X_2]f &= [y\partial_z - z\partial_y, z\partial_x - x\partial_z]f \\ &= yz[\partial_z, \partial_x]f + yf_{,x} - yx[\partial_z, \partial_z]f - z^2[\partial_y, \partial_x]f + zx[\partial_y, \partial_z]f - xf_{,y} \\ &= \big(-x\partial_y + y\partial_x\big)f = -X_3 f. \end{aligned}$$

There are similar identities for the other commutators,

$$[X_1, X_2] = -X_3, \qquad [X_2, X_3] = -X_1, \qquad [X_3, X_1] = -X_2. \tag{15.45}$$

Hence the distribution D^2 is involutive and by the Frobenius theorem it is possible to find a local transformation to coordinates $y^1,\ y^2,\ y^3$ such that ∂_{y^1} and ∂_{y^2} span all three vector fields X_1, X_2 and X_3.

The vector field $X = x\partial_x + y\partial_y + z\partial_z$ commutes with all X_i: for example,

$$\begin{aligned}
[X_3, X]f &= [x\partial_y - y\partial_x,\ xf_{,x} + yf_{,y} + zf_{,z}]f \\
&= x^2[\partial_y, \partial_x]f - x\partial_y f + x\partial_y f + xy[\partial_y, \partial_y]f - y^2[\partial_x, \partial_y]f \\
&\quad + xz[\partial_y, \partial_z]f - yz[\partial_x, \partial_z]f \\
&= 0\,.
\end{aligned}$$

Hence the distribution E^2 generated by the pair of vector fields $\{X_3, X\}$ is also involutive. Let us consider spherical polar coordinates, Eq. (15.2), having inverse transformations

$$r = \sqrt{x^2 + y^2 + z^2}, \qquad \theta = \cos^{-1}\left(\frac{z}{r}\right), \qquad \phi = \tan^{-1}\left(\frac{y}{x}\right).$$

Express the basis vector fields in terms of these coordinates

$$\begin{aligned}
\partial_x &= \frac{\partial r}{\partial x}\partial_r + \frac{\partial \theta}{\partial x}\partial_\theta + \frac{\partial \phi}{\partial x}\partial_\phi = \sin\theta\cos\phi\,\partial_r + \frac{\cos\theta\cos\phi}{r}\partial_\theta - \frac{\sin\phi}{r\sin\theta}\partial_\phi, \\
\partial_y &= \frac{\partial r}{\partial y}\partial_r + \frac{\partial \theta}{\partial y}\partial_\theta + \frac{\partial \phi}{\partial y}\partial_\phi = \sin\theta\sin\phi\,\partial_r + \frac{\cos\theta\sin\phi}{r}\partial_\theta + \frac{\cos\phi}{r\sin\theta}\partial_\phi, \\
\partial_z &= \frac{\partial r}{\partial z}\partial_r + \frac{\partial \theta}{\partial z}\partial_\theta + \frac{\partial \phi}{\partial z}\partial_\phi = \cos\theta\,\partial_r - \frac{\sin\theta}{r}\partial_\theta,
\end{aligned}$$

and a simple calculation gives

$$\begin{aligned}
X_1 &= y\partial_z - z\partial_y = -\sin\phi\,\partial_\theta - \cot\theta\cos\phi\,\partial_\phi, \\
X_2 &= z\partial_x - x\partial_z = -\cos\phi\,\partial_\theta - \cot\theta\sin\phi\,\partial_\phi, \\
X_3 &= x\partial_y - y\partial_x = \partial_\phi, \\
X &= x\partial_x + y\partial_y + z\partial_z = r\partial_r = \partial_{r'} \quad \text{where } r' = \ln r.
\end{aligned}$$

The distribution D^2 is spanned by the basis vector fields ∂_θ and ∂_ϕ, while the distribution E^2 is spanned by the vector fields ∂_r and ∂_ϕ in spherical polars.

Exercise: Find a chart, two of whose basis vector fields span the distribution generated by X_1 and X. Do the same for the distribution generated by X_2 and X.

Problems

Problem 15.26 Let D_k be an involutive distribution spanned locally by coordinate vector fields $e_\alpha = \partial/\partial x^\alpha$, where Greek indices α, β, etc. all range from 1 to k. If $X_\alpha = A^\beta{}_\alpha e_\beta$ is any local basis spanning a distribution D^k, show that the matrix of functions $[A^\beta{}_\alpha]$ is non-singular everywhere on its region of definition, and that $[X_\alpha, X_\beta] = C^\gamma_{\alpha\beta} X_\gamma$ where

$$C^\gamma_{\alpha\beta} = \left(A^\delta{}_\alpha A^\eta{}_{\beta,\delta} - A^\delta{}_\beta A^\eta{}_{\alpha,\delta}\right)\left(A^{-1}\right)^\gamma{}_\eta.$$

Problem 15.27 There is a classical version of the Frobenius theorem stating that a system of partial differential equations of the form

$$\frac{\partial f^\beta}{\partial x^j} = A^\beta_j\big(x^1, \dots, x^k, f^1(x), \dots, f^r(x)\big)$$

where $i, j = 1, \dots, k$ and $\alpha, \beta = 1, \dots, r$ has a unique local solution through any point $(a^1, \dots, a^k, b^1, \dots, b^r)$ if and only if

$$\frac{\partial A^\beta_j}{\partial x^i} - \frac{\partial A^\beta_i}{\partial x^j} + A^\alpha_i \frac{\partial A^\beta_j}{\partial y^\alpha} - A^\alpha_j \frac{\partial A^\beta_i}{\partial y^\alpha} = 0$$

where $A^\beta_j = A^\beta_j\big(x^1, \dots, x^k, y^1, \dots, y^r\big)$. Show that this statement is equivalent to the version given in Theorem 15.4. [*Hint*: On $\mathbb{R}^n$ where $n = r + k$ consider the distribution spanned by vectors

$$Y_i = \frac{\partial}{\partial x^i} + A^\beta_i \frac{\partial}{\partial y^\beta} \quad (i = 1, \dots, k)$$

and show that the integrability condition is precisely the involutive condition $[Y_i, Y_j] = 0$, while the condition for an integral submanifold of the form $y^\beta = f^\beta(x^1, \dots, x^k)$ is $A^\beta_j = f^\beta_{,j}$.]

References

[1] L. Auslander and R. E. MacKenzie. *Introduction to Differentiable Manifolds*. New York, McGraw-Hill, 1963.

[2] R. W. R. Darling. *Differential Forms and Connections*. New York, Cambridge University Press, 1994.

[3] T. Frankel. *The Geometry of Physics*. New York, Cambridge University Press, 1997.

[4] N. J. Hicks. *Notes on Differential Geometry*. New York, D. Van Nostrand Company, 1965.

[5] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*. New York, Interscience Publishers, 1963.

[6] L. H. Loomis and S. Sternberg. *Advanced Calculus*. Reading, Mass., Addison-Wesley, 1968.

[7] M. Nakahara. *Geometry, Topology and Physics*. Bristol, Adam Hilger, 1990.

[8] C. Nash and S. Sen. *Topology and Geometry for Physicists*. London, Academic Press, 1983.

[9] I. M. Singer and J. A. Thorpe. *Lecture Notes on Elementary Topology and Geometry*. Glenview, Ill., Scott Foresman, 1967.

[10] M. Spivak. *Differential Geometry, Vols. 1–5*. Boston, Publish or Perish Inc., 1979.

[11] W. H. Chen, S. S. Chern, and K. S. Lam. *Lectures on Differential Geometry*. Singapore, World Scientific, 1999.

[12] S. Sternberg. *Lectures on Differential Geometry*. Englewood Cliffs, N.J., Prentice-Hall, 1964.

[13] F. W. Warner. *Foundations of Differential Manifolds and Lie Groups*. New York, Springer-Verlag, 1983.

[14] C. de Witt-Morette, Y. Choquet-Bruhat and M. Dillard-Bleick. *Analysis, Manifolds and Physics*. Amsterdam, North-Holland, 1977.

[15] E. Coddington and N. Levinson. *Theory of Ordinary Differential Equations*. New York, McGraw-Hill, 1955.

[16] W. Hurewicz. *Lectures on Ordinary Differential Equations*. New York, John Wiley & Sons, 1958.

[17] E. Nelson. *Tensor Analysis*. Princeton, N.J., Princeton University Press, 1967.