# 5

# MANIFOLDS

Manifolds are generalizations of our familiar ideas about curves and surfaces to arbitrary dimensional objects. A curve in three-dimensional Euclidean space is parametrized locally by a single number $t$ as $(x(t), y(t), z(t))$, while two numbers $u$ and $v$ parametrize a surface as $(x(u, v), y(u, v), z(u, v))$. A curve and a surface are considered locally homeomorphic to $\mathbb{R}$ and $\mathbb{R}^2$, respectively. A manifold, in general, is a topological space which is homeomorphic to $\mathbb{R}^m$ *locally*; it may be different from $\mathbb{R}^m$ *globally*. The local homeomorphism enables us to give each point in a manifold a set of $m$ numbers called the (local) coordinate. If a manifold is not homeomorphic to $\mathbb{R}^m$ globally, we have to introduce several local coordinates. Then it is possible that a single point has two or more coordinates. We require that the transition from one coordinate to the other be *smooth*. As we will see later, this enables us to develop the usual calculus on a manifold. Just as topology is based on continuity, so the theory of manifolds is based on *smoothness*.

Useful references on this subject are Crampin and Pirani (1986), Matsushima (1972), Schutz (1980) and Warner (1983). Chapter 2 and appendices B and C of Wald (1984) are also recommended. Flanders (1963) is a beautiful introduction to differential forms. Sattinger and Weaver (1986) deals with Lie groups and Lie algebras and contains many applications to problems in physics.

## 5.1 Manifolds

### 5.1.1 Heuristic introduction

To clarify these points, consider the usual sphere of unit radius in $\mathbb{R}^3$. We parametrize the surface of $S^2$, among other possibilities, by two coordinate systems—polar coordinates and stereographic coordinates. Polar coordinates $\theta$ and $\phi$ are usually defined by (figure 5.1)

$$x = \sin\theta\cos\phi \qquad y = \sin\theta\sin\phi \qquad z = \cos\theta, \tag{5.1}$$

where $\phi$ runs from 0 to $2\pi$ and $\theta$ from 0 to $\pi$. They may be inverted on the sphere to yield

$$\theta = \tan^{-1}\frac{\sqrt{x^2+y^2}}{z} \qquad \phi = \tan^{-1}\frac{y}{x}. \tag{5.2}$$

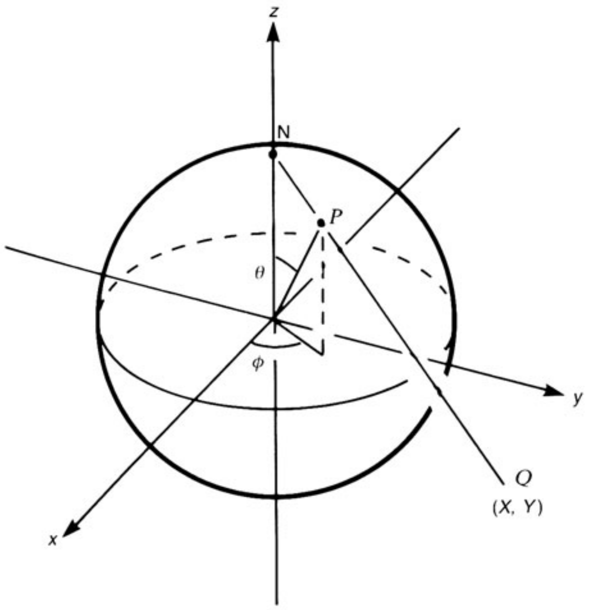


**Figure 5.1.** Polar coordinates $(\theta, \phi)$ and stereographic coordinates $(X, Y)$ of a point $P$ on the sphere $S^2$.

Stereographic coordinates, however, are defined by the projection from the North Pole onto the equatorial plane as in figure 5.1. First, join the North Pole (0, 0, 1) to the point $P(x, y, z)$ on the sphere and then continue in a straight line to the equatorial plane $z = 0$ to intersect at $Q(X, Y, 0)$. Then $X$ and $Y$ are the stereographic coordinates of $P$. We find

$$X = \frac{x}{1-z} \qquad Y = \frac{y}{1-z}. \tag{5.3}$$

The two coordinate systems are related as

$$X = \cot \tfrac{1}{2}\theta \cos \phi \qquad Y = \cot \tfrac{1}{2}\theta \sin \phi. \tag{5.4}$$

Of course, other systems, polar coordinates with different polar axes or projections from different points on $S^2$, could be used. The coordinates on the sphere may be kept arbitrary until some specific calculation is to be carried out. [The longitude is historically measured from Greenwich. However, there is no reason why it cannot be measured from New York or Kyoto.] This arbitrariness of the coordinate choice underlies the theory of manifolds: *all coordinate systems are equally good*. It is also in harmony with the basic principle of physics: *a physical system behaves in the same way whatever coordinates we use to describe it*.

Another point which can be seen from this example is that *no coordinate system may be usable everywhere at once*. Let us look at the polar coordinates on $S^2$. Take the equator ($\theta = \frac{1}{2}\pi$) for definiteness. If we let $\phi$ range from 0 to $2\pi$, then it changes continuously as we go round the equator until we get all the way to $\phi = 2\pi$. There the $\phi$-coordinate has a discontinuity from $2\pi$ to 0 and nearby points have quite different $\phi$-values. Alternatively we could continue $\phi$ through $2\pi$. Then we will encounter another difficulty: at each point we must have infinitely many $\phi$-values, differing from one another by an integral multiple of $2\pi$. A further difficulty arises at the poles, where $\phi$ is not determined at all. [An explorer on the Pole is in a state of timelessness since time is defined by the longitude.] Stereographic coordinates also have difficulties at the North Pole or at any projection point that is not projected to a point on the equatorial plane; and nearby points close to the Pole have widely different stereographic coordinates.

Thus, we cannot label the points on the sphere with a single coordinate system so that both of the following conditions are satisfied.

(i) Nearby points always have nearby coordinates.
(ii) Every point has unique coordinates.

Note, however, that there are infinitely many ways to introduce coordinates that satisfy these requirements on a *part* of $S^2$. We may take advantage of this fact to define coordinates on $S^2$: introduce two or more overlapping coordinate systems, each covering a part of the sphere whose points are to be labelled so that the following conditions hold.

(i′) Nearby points have nearby coordinatcs in at least one coordinate system.
(ii′) Every point has unique coordinates in each system that contains it.

For example, we may introduce two stereographic coordinates on $S^2$, one a projection from the North Pole, the other from the South Pole. Are these conditions (i′) and (ii′) enough to develop sensible theories of the manifold? In fact, we need an extra condition on the coordinate systems.

(iii) If two coordinate systems overlap, they are related to each other in a sufficiently smooth way.

Without this condition, a differentiable function in one coordinate system may not be differentiable in the other system.

### 5.1.2 Definitions

*Definition 5.1.* $M$ is an $m$-dimensional differentiable manifold if

(i) $M$ is a topological space;
(ii) $M$ is provided with a family of pairs $\{(U_i, \varphi_i)\}$;
(iii) $\{U_i\}$ is a family of open sets which covers $M$, that is, $\cup_i U_i = M$. $\varphi_i$ is a homeomorphism from $U_i$ onto an open subset $U_i'$ of $\mathbb{R}^m$ (figure 5.2); and

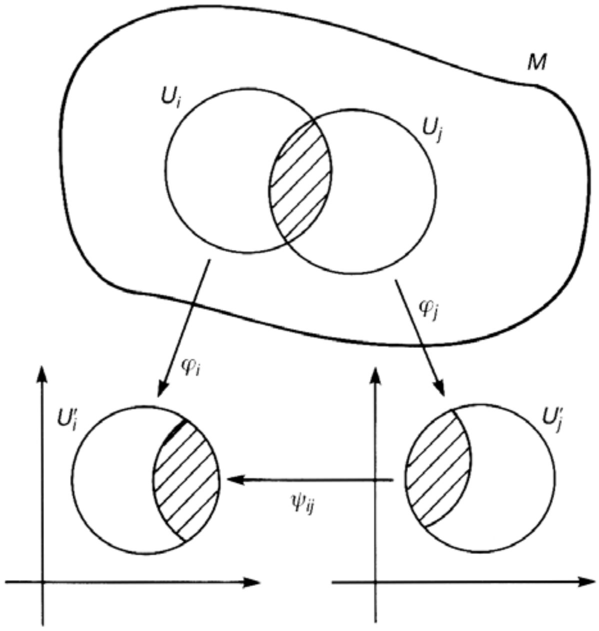


**Figure 5.2.** A homeomorphism $\varphi_i$ maps $U_i$ onto an open subset $U'_i \subset \mathbb{R}^m$, providing coordinates to a point $p \in U_i$. If $U_i \cap U_j \neq \emptyset$, the transition from one coordinate system to another is smooth.

(iv) given $U_i$ and $U_j$ such that $U_i \cap U_j \neq \emptyset$, the map $\psi_{ij} = \varphi_i \circ \varphi_j^{-1}$ from $\varphi_j(U_i \cap U_j)$ to $\varphi_i(U_i \cap U_j)$ is infinitely differentiable.

The pair $(U_i, \varphi_i)$ is called a **chart** while the whole family $\{(U_i, \varphi_i)\}$ is called, for obvious reasons, an **atlas**. The subset $U_i$ is called the **coordinate neighbourhood** while $\varphi_i$ is the **coordinate function** or, simply, the **coordinate**. The homeomorphism $\varphi_i$ is represented by $m$ functions $\{x^1(p), \ldots, x^m(p)\}$. The set $\{x^\mu(p)\}$ is also called the **coordinate**. A point $p \in M$ exists independently of its coordinates; it is up to us how we assign coordinates to a point. We sometimes employ the rather sloppy notation $x$ to denote a point whose coordinates are $\{x^1, \ldots, x^m\}$, unless several coordinate systems are in use. From (ii) and (iii), $M$ is locally Euclidean. In each coordinate neighbourhood $U_i$, $M$ looks like an open subset of $\mathbb{R}^m$ whose element is $\{x^1, \ldots, x^m\}$. Note that we do not require that $M$ be $\mathbb{R}^m$ *globally*. We are living on the earth whose surface is $S^2$, which does not look like $\mathbb{R}^2$ globally. However, it looks like an open subset of $\mathbb{R}^2$ *locally*. Who can tell that we live on the sphere by just looking at a map of London, which, of course, looks like a part of $\mathbb{R}^2$?[1]

[1] Strictly speaking the distance between two longitudes in the northern part of the city is slightly

If $U_i$ and $U_j$ overlap, two coordinate systems are assigned to a point in $U_i \cap U_j$. Axiom (iv) asserts that the transition from one coordinate system to another be *smooth* ($C^\infty$). The map $\varphi_i$ assigns $m$ coordinate values $x^\mu$ ($1 \le \mu \le m$) to a point $p \in U_i \cap U_j$, while $\varphi_j$ assigns $y^\nu$ ($1 \le \nu \le m$) to the same point and the transition from $y$ to $x$, $x^\mu = x^\mu(y)$, is given by $m$ functions of $m$ variables. The coordinate transformation functions $x^\mu = x^\mu(y)$ are the explicit form of the map $\psi_{ji} = \varphi_j \circ \varphi_i^{-1}$. Thus, the differentiability has been defined in the usual sense of calculus: the coordinate transformation is differentiable if each function $x^\mu(y)$ is differentiable with respect to each $y^\nu$. We may restrict ourselves to the differentiability up to $k$th order ($C^k$). However, this does not bring about any interesting conclusions. We simply require, instead, that the coordinate transformations be infinitely differentiable, that is, of class $C^\infty$. Now coordinates have been assigned to $M$ in such a way that if we move over $M$ in whatever fashion, the coordinates we use vary in a smooth manner.

If the union of two atlases $\{(U_i, \varphi_i)\}$ and $\{(V_j, \psi_j)\}$ is again an atlas, these two atlases are said to be **compatible**. The compatibility is an equivalence relation, the equivalence class of which is called the **differentiable structure**. It is also said that mutually compatible atlases define the same differentiable structure on $M$.

Before we give examples, we briefly comment on manifolds *with boundaries*. So far, we have assumed that the coordinate neighbourhood $U_i$ is homeomorphic to an open set of $\mathbb{R}^m$. In some applications, however, this turns out to be too restrictive and we need to relax this condition. If a topological space $M$ is covered by a family of open sets $\{U_i\}$ each of which is homeomorphic to an open set of $H^m \equiv \{(x^1, \ldots, x^m) \in \mathbb{R}^m | x^m \ge 0\}$, $M$ is said to be a **manifold with a boundary**, see figure 5.3. The set of points which are mapped to points with $x^m = 0$ is called the **boundary** of $M$, denoted by $\partial M$. The coordinates of $\partial M$ may be given by $m - 1$ numbers $(x^1, \ldots, x^{m-1}, 0)$. Now we have to be careful when we define the smoothness. The map $\psi_{ij} : \varphi_j(U_i \cap U_j) \to \varphi_i(U_i \cap U_j)$ is defined on an open set of $H^m$ in general, and $\psi_{ij}$ is said to be smooth if it is $C^\infty$ in an open set of $\mathbb{R}^m$ which contains $\varphi_j(U_i \cap U_j)$. Readers are encouraged to use their imagination since our definition is in harmony with our intuitive notions about boundaries. For example, the boundary of the solid ball $D^3$ is the sphere $S^2$ and the boundary of the sphere is an empty set.

### 5.1.3 Examples

We now give several examples to develop our ideas about manifolds. They are also of great relevance to physics.

*Example 5.1.* The Euclidean space $\mathbb{R}^m$ is the most trivial example, where a single chart covers the whole space and $\varphi$ may be the identity map.

shorter than that in the southern part and one may suspect that one lives on a curved surface. Of course, it is the other way around if one lives in a city in the southern hemisphere.

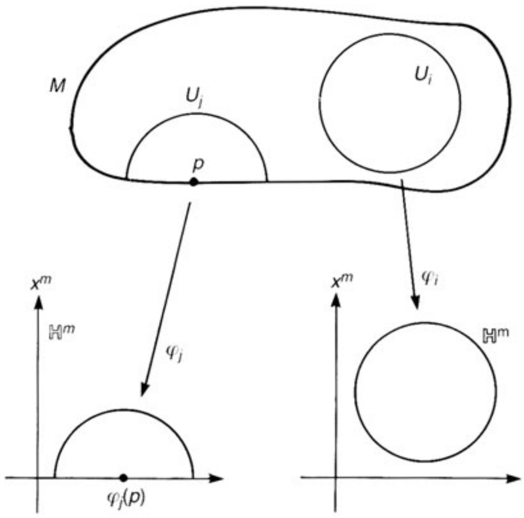


**Figure 5.3.** A manifold with a boundary. The point $p$ is on the boundary.

*Example 5.2.* Let $m = 1$ and require that $M$ be connected. There are only two manifolds possible: a real line $\mathbb{R}$ and the circle $S^1$. Let us work out an atlas of $S^1$. For concreteness take the circle $x^2 + y^1 = 1$ in the $xy$-plane. We need at least two charts. We may take them as in figure 5.4. Define $\varphi_1^{-1} : (0, 2\pi) \to S^1$ by

$$\varphi_1^{-1} : \theta \mapsto (\cos\theta, \sin\theta) \tag{5.5a}$$

whose image is $S^1 - \{(1, 0)\}$. Define also $\psi_2^{-1} : (-\pi, \pi) \to S^1$ by

$$\varphi_2^{-1} : \theta \mapsto (\cos\theta, \sin\theta) \tag{5.5b}$$

whose image is $S^1 - \{(-1, 0)\}$. Clearly $\varphi_1^{-1}$ and $\varphi_2^{-1}$ are invertible and all the maps $\varphi_1, \varphi_2, \varphi_1^{-1}$ and $\varphi_2^{-1}$ are continuous. Thus, $\varphi_1$ and $\varphi_2$ are homeomorphisms. Verify that the maps $\psi_{12} = \varphi_1 \circ \varphi_2^{-1}$ and $\psi_{21} = \varphi_2 \circ \varphi_1^{-1}$ are smooth.

*Example 5.3.* The $n$-dimensional sphere $S^n$ is a differentiable manifold. It is realized in $\mathbb{R}^{n+1}$ as

$$\sum_{i=0}^{n} (x^i)^2 = 1. \tag{5.6}$$

Let us introduce the coordinate neighbourhoods

$$U_{i+} \equiv \{(x^0, x^1, \ldots, x^n) \in S^n | x^i > 0\} \tag{5.7a}$$

$$U_{i-} \equiv \{(x^0, x^1, \ldots, x^n) \in S^n | x^i < 0\}. \tag{5.7b}$$

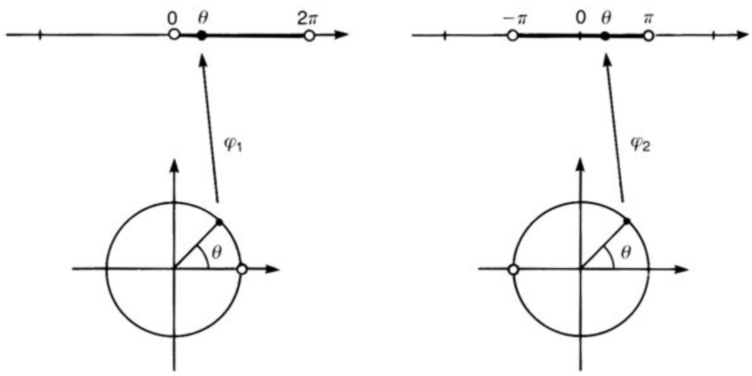


**Figure 5.4.** Two charts of a circle $S^1$.

Define the coordinate map $\varphi_{i+} : U_{i+} \to \mathbb{R}^n$ by

$$\varphi_{i+}(x^0, \ldots, x^n) = (x^0, \ldots, x^{i-1}, x^{i+1}, \ldots, x^n) \tag{5.8a}$$

and $\varphi_{i-} : U_{i-} \to \mathbb{R}^n$ by

$$\varphi_{i-}(x^0, \ldots, x^n) = (x^0, \ldots, x^{i-1}, x^{i+1}, \ldots, x^n). \tag{5.8b}$$

Note that the domains of $\varphi_{i+}$ and $\varphi_{i-}$ are different. $\varphi_{i\pm}$ are the projections of the hemispheres $U_{i\pm}$ to the plane $x^i = 0$. The transition functions are easily obtained from (5.8). Take $S^2$ as an example. The coordinate neighbourhoods are $U_{x\pm}$, $U_{y\pm}$ and $U_{z\pm}$. The transition function $\psi_{y-x+} \equiv \varphi_{y-} \circ \varphi_{x+}^{-1}$ is given by

$$\psi_{y-x+} : (y, z) \mapsto \left(\sqrt{1 - y^2 - z^2}, z\right) \tag{5.9}$$

which is infinitely differentiable on $U_{x+} \cap U_{y-}$.

*Exercise 5.1.* At the beginning of this chapter, we introduced the stereographic coordinates on $S^2$. We may equally define the stereographic coordinates projected from points other than the North Pole. For example, the stereographic coordinates $(U, V)$ of a point in $S^2 - \{\text{South Pole}\}$ projected from the South Pole and $(X, Y)$ for a point in $S^2 - \{\text{North Pole}\}$ projected from the North Pole are shown in figure 5.5. Show that the transition functions between $(U, V)$ and $(X, Y)$ are $C^\infty$ and that they define a differentiable structure on $M$. See also example 8.1.

*Example 5.4.* The real projective space $\mathbb{R}P^n$ is the set of lines through the origin in $\mathbb{R}^{n+1}$. If $x = (x^0, \ldots, x^n) \neq 0$, $x$ defines a line through the origin. Note that $y \in \mathbb{R}^{n+1}$ defines the same line as $x$ if there exists a real number $a \neq 0$ such that $y = ax$. Introduce an equivalence relation $\sim$ by $x \sim y$ if there

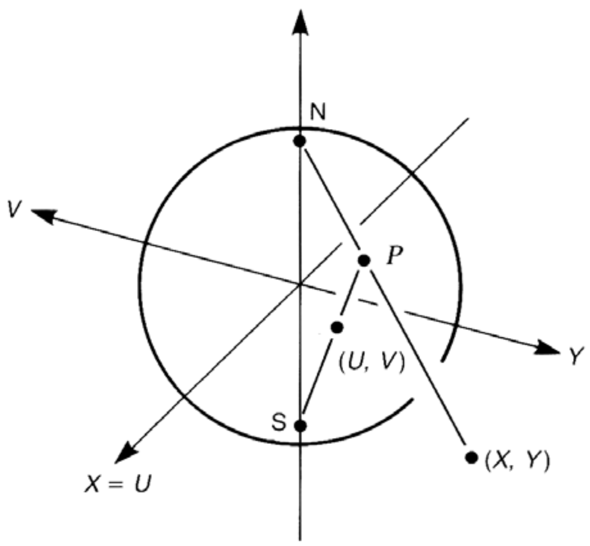


**Figure 5.5.** Two stereographic coordinate systems on $S^2$. The point $P$ may be projected from the North Pole N giving $(X, Y)$ or from the South Pole S giving $(U, V)$.

exists $a \in \mathbb{R} - \{0\}$ such that $y = ax$. Then $\mathbb{R}P^n = (\mathbb{R}^{n+1} - \{0\})/\sim$. The $n + 1$ numbers $x^0, x^1, \ldots, x^n$ are called the **homogeneous coordinates**. The homogeneous coordinates cannot be a good coordinate system, since $\mathbb{R}P^n$ is an $n$-dimensional manifold (an $(n + 1)$-dimensional space with a one-dimensional degree of freedom killed). The charts are defined as follows. First we take the coordinate neighbourhood $U_i$ as the set of lines with $x^i \neq 0$, and then introduce the **inhomogeneous coordinates** on $U_i$ by

$$\xi^j_{(i)} = x^j/x^i. \tag{5.10}$$

The inhomogeneous coordinates

$$\xi_{(i)} = (\xi^0_{(i)}, \xi^1_{(i)}, \ldots, \xi^{i-1}_{(i)}, \xi^{i+1}_{(i)}, \ldots, \xi^n_{(i)})$$

with $\xi^i_{(i)} = 1$ omitted, are well defined on $U_i$ since $x^i \neq 0$, and furthermore they are independent of the choice of the representative of the equivalence class since $x^j/x^i = y^j/y^i$ if $y = ax$. The inhomogeneous coordinate $\xi_{(i)}$ gives the coordinate map $\varphi_i : U_i \to \mathbb{R}^n$, that is

$$\varphi_i : (x^0, \ldots, x^n) \mapsto (x^0/x^i, \ldots, x^{i-1}/x^i, x^{i+1}/x^i, \ldots, x^n/x^i)$$

where $x^i/x^i = 1$ is omitted. For $x = (x^0, x^1, \ldots, x^n) \in U_i \cap U_j$ we assign two inhomogeneous coordinates, $\xi^k_{(i)} = x^k/x^i$ and $\xi^k_{(j)} = x^k/x^j$. The coordinate transformation $\psi_{ij} = \varphi_i \circ \varphi_j^{-1}$ is

$$\psi_{ij} : \xi^k_{(j)} \mapsto \xi^k_{(i)} = (x^j/x^i)\xi^k_{(j)}. \tag{5.11}$$

This is a multiplication by $x^j/x^i$.

In example 4.12, we defined $\mathbb{R}P^n$ as the sphere $S^n$ with antipodal points identified. This picture is in conformity with the definition here. As a representative of the equivalence class $[x]$, we may take points $|x| = 1$ on a line through the origin. These are points on the unit sphere. Since there are two points on the intersection of a line with $S^n$ we have to take one of them consistently, that is nearby lines are represented by nearby points in $S^n$. This amounts to taking the hemisphere. Note, however, that the antipodal points on the boundary (the equator of $S^n$) are identified by definition, $(x^0, \ldots, x^n) \sim -(x^0, \ldots, x^n)$. This 'hemisphere' is homeomorphic to the ball $D^n$ with antipodal points on the boundary $S^{n-1}$ identified.

*Example 5.5.* A straightforward generalization of $\mathbb{R}P^n$ is the **Grassmann manifold**. An element of $\mathbb{R}P^n$ is a one-dimensional subspace in $\mathbb{R}^{n+1}$. The Grassmann manifold $G_{k,n}(\mathbb{R})$ is the set of $k$-dimensional planes in $\mathbb{R}^n$. Note that $G_{1,n+1}(\mathbb{R})$ is nothing but $\mathbb{R}P^n$. The manifold structure of $G_{k,n}(\mathbb{R})$ is defined in a manner similar to that of $\mathbb{R}P^n$.

Let $M_{k,n}(\mathbb{R})$ be the set of $k \times n$ matrices of rank $k$ ($k \leq n$). Take $A = (a_{ij}) \in M_{k,n}(\mathbb{R})$ and define $k$ vectors $\boldsymbol{a}_i$ $(1 \leq i \leq k)$ in $\mathbb{R}^n$ by $\boldsymbol{a}_i = (a_{ij})$. Since rank $A = k$, $k$ vectors $\boldsymbol{a}_i$ are linearly independent and span a $k$-dimensional plane in $\mathbb{R}^n$. Note, however, that there are infinitely many matrices in $M_{k,n}(\mathbb{R})$ that yield the same $k$-plane. Take $g \in \mathrm{GL}(k, \mathbb{R})$ and consider a matrix $\bar{A} = gA \in M_{k,n}(\mathbb{R})$. $\bar{A}$ defines the same $k$-plane as $A$, since $g$ simply rotates the basis within the $k$-plane. Introduce an equivalence relation $\sim$ by $\bar{A} \sim A$ if there exists $g \in \mathrm{GL}(k, \mathbb{R})$ such that $\bar{A} = gA$. We identify $G_{k,n}(\mathbb{R})$ with the coset space $M_{k,n}(\mathbb{R})/\mathrm{GL}(k, \mathbb{R})$.

Let us find the charts of $G_{k,n}(\mathbb{R})$. Take $A \in M_{k,n}(\mathbb{R})$ and let $\{A_1, \ldots, A_l\}$, $l = \binom{n}{k}$, be the collection of all $k \times k$ minors of $A$. Since rank $A = k$, there exists some $A_\alpha$ $(1 \leq \alpha \leq l)$ such that $\det A \neq 0$. For example, let us assume the minor $A_1$ made of the first $k$ columns has non-vanishing determinant,

$$A = (A_1, \widetilde{A_1}) \tag{5.12}$$

where $\widetilde{A_1}$ is a $k \times (n-k)$ matrix. Let us take the representative of the class to which $A$ belongs to be

$$A_1^{-1} \cdot A = (I_k, A_1^{-1} \cdot \widetilde{A_1}) \tag{5.13}$$

where $I_k$ is the $k \times k$ unit matrix. Note that $A_1^{-1}$ always exists since $\det A_1 \neq 0$. Thus, the real degrees of freedom are given by the entries of the $k \times (n-k)$ matrix $A_1^{-1} \cdot \widetilde{A_1}$. We denote this subset of $G_{k,n}(\mathbb{R})$ by $U_1$. $U_1$ is a coordinate neighbourhood whose coordinates are given by $k(n-k)$ entries of $A_1^{-1} \cdot \widetilde{A_1}$. Since $U_1$ is homeomorphic to $\mathbb{R}^{k(n-k)}$ we find that

$$\dim G_{k,n}(\mathbb{R}) = k(n-k). \tag{5.14}$$

In the case where $\det A_\alpha \neq 0$, where $A_\alpha$ is composed of the columns $(i_1, i_2, \ldots, i_k)$, we multiply $A_\alpha^{-1}$ to obtain the representative

$$\begin{array}{cc} \text{column} \rightarrow & \begin{array}{ccccccccc} \quad & i_1 & & i_2 & & \ldots & & i_k & \end{array} \\ A_\alpha^{-1} \cdot A = & \begin{pmatrix} \ldots & 1 & \ldots & 0 & \ldots\ldots & 0 & \ldots \\ \ldots & 0 & \ldots & 1 & \ldots\ldots & 0 & \ldots \\ \ldots & . & \ldots & . & \ldots\ldots & . & \ldots \\ \ldots & 0 & \ldots & 0 & \ldots\ldots & 1 & \ldots \end{pmatrix} \end{array} \tag{5.15}$$

where the entries not written explicitly form a $k \times (n-k)$ matrix. We denote this subset of $M_{k,n}(\mathbb{R})$ with $\det A_\alpha \neq 0$ by $U_\alpha$. The entries of the $k \times (n-k)$ matrix are the coordinates of $U_\alpha$.

The relation between the projective space and the Grassmann manifold is evident. An element of $M_{1,n+1}(\mathbb{R})$ is a vector $A = (x^0, x^1, \ldots, x^n)$. Since the $\alpha$th minor $A_\alpha$ of $A$ is a number $x^\alpha$, the condition $\det A_\alpha \neq 0$ becomes $x^\alpha \neq 0$. The representative (5.15) is just the inhomogeneous coordinate

$$\begin{aligned} &(x^\alpha)^{-1}(x^0, x^1, \ldots, x^\alpha, \ldots, x^n) \\ &= (x^0/x^\alpha, x^1/x^\alpha, \ldots, x^\alpha/x^\alpha = 1, \ldots, x^n/x^\alpha). \end{aligned}$$

Let $M$ be an $m$-dimensional manifold with an atlas $\{(U_i, \varphi_i)\}$ and $N$ be an $n$-dimensional manifold with $\{(V_j, \psi_j)\}$. A **product manifold** $M \times N$ is an $(m+n)$-dimensional manifold whose atlas is $\{(U_i \times V_j), (\varphi_i, \psi_j)\}$. A point in $M \times N$ is written as $(p, q)$, $p \in M$, $q \in N$, and the coordinate function $(\varphi_i, \psi_j)$ acts on $(p, q)$ to yield $(\varphi_i(p), \psi_j(p)) \in \mathbb{R}^{m+n}$. The reader should verify that a product manifold indeed satisfies the axioms of definition 5.1.

*Example 5.6.* The torus $T^2$ is a product manifold of two circles, $T^2 = S^1 \times S^1$. If we denote the polar angle of each circle as $\theta_i \bmod 2\pi$ $(i = 1, 2)$, the coordinates of $T^2$ are $(\theta_1, \theta_2)$. Since each $S^1$ is embedded in $\mathbb{R}^2$, $T^2$ may be embedded in $\mathbb{R}^4$. We often imagine $T^2$ as the surface of a doughnut in $\mathbb{R}^3$, in which case, however, we inevitably have to introduce bending of the surface. This is an extrinsic feature brought about by the 'embedding'. When we say 'a torus is a flat manifold', we refer to the flat surface embedded in $\mathbb{R}^4$. See definition 5.3 for further details.

We may also consider a direct product of $n$ circles,

$$T^n = \underbrace{S^1 \times S^1 \times \cdots \times S^1}_{n}.$$

Clearly $T^n$ is an $n$-dimensional manifold with the coordinates $(\theta_1, \theta_2, \ldots, \theta_n)$ mod$2\pi$. This may be regarded as an $n$-cube whose opposite faces are identified, see figure 2.4 for $n = 2$.

## 5.2 The calculus on manifolds

The significance of differentiable manifolds resides in the fact that we may use the usual calculus developed in $\mathbb{R}^n$. Smoothness of the coordinate transformations

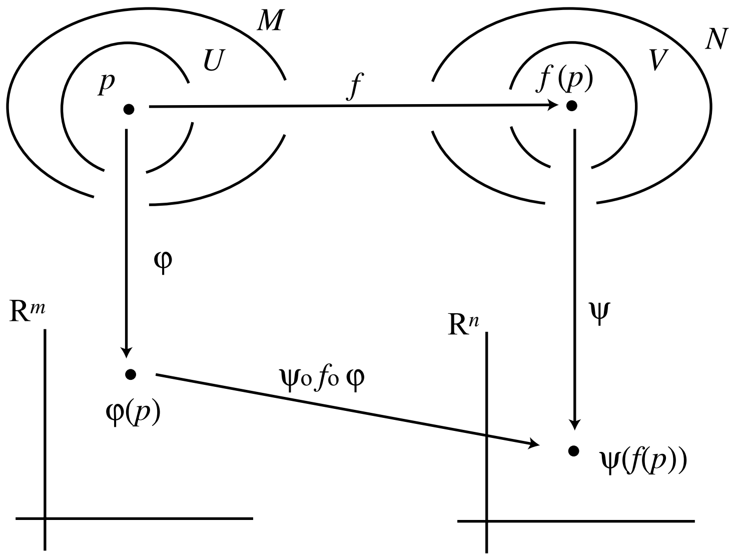


**Figure 5.6.** A map $f : M \to N$ has a coordinate presentation $\psi \circ f \circ \varphi^{-1} : \mathbb{R}^m \to \mathbb{R}^n$.

ensures that the calculus is independent of the coordinates chosen.

### 5.2.1 Differentiable maps

Let $f : M \to N$ be a map from an $m$-dimensional manifold $M$ to an $n$-dimensional manifold $N$. A point $p \in M$ is mapped to a point $f(p) \in N$, namely $f : p \mapsto f(p)$, see figure 5.6. Take a chart $(U, \varphi)$ on $M$ and $(V, \psi)$ on $N$, where $p \in U$ and $f(p) \in V$. Then $f$ has the following coordinate presentation:

$$\psi \circ f \circ \varphi^{-1} : \mathbb{R}^m \to \mathbb{R}^n. \tag{5.16}$$

If we write $\varphi(p) = \{x^\mu\}$ and $\psi(f(p)) = \{y^\alpha\}$, $\psi \circ f \circ \varphi^{-1}$ is just the usual vector-valued function $y = \psi \circ f \circ \varphi^{-1}(x)$ of $m$ variables. We sometimes use (in fact, abuse!) the notation $y = f(x)$ or $y^\alpha = f^\alpha(x^\mu)$, when we know which coordinate systems on $M$ and $N$ are in use. If $y = \psi \circ f \circ \varphi^{-1}(x)$, or simply $y^\alpha = f^\alpha(x^\mu)$, is $C^\infty$ with respect to each $x^\mu$, $f$ is said to be **differentiable** at $p$ or at $x = \varphi(p)$. Differentiable maps are also said to be **smooth**. Note that we require infinite ($C^\infty$) differentiability, in harmony with the smoothness of the transition functions $\psi_{ij}$.

The differentiability of $f$ is independent of the coordinate system. Consider two overlapping charts $(U_1, \varphi_1)$ and $(U_2, \varphi_2)$. Take a point $p \in U_1 \cap U_2$, whose coordinates by $\varphi_1$ are $\{x_1^\mu\}$, while those by $\varphi_2$ are $\{x_2^\nu\}$. When expressed in terms of $\{x_1^\mu\}$, $f$ takes the form $\psi \circ f \circ \varphi_1^{-1}$, while in $\{x_2^\nu\}$, $\psi \circ f \circ \varphi_2^{-1} =$

$\psi \circ f \circ \varphi_1^{-1}(\varphi_1 \circ \varphi_2^{-1})$. By definition, $\psi_{12} = \varphi_1 \circ \varphi_2^{-1}$ is $C^\infty$. In the simpler expressions, they correspond to $y = f(x_1)$ and $y = f(x_1(x_2))$. It is clear that if $f(x_1)$ is $C^\infty$ with respect to $x_1^\mu$ and $x_1(x_2)$ is $C^\infty$ with respect to $x_2^\nu$, then $y = f(x_1(x_2))$ is also $C^\infty$ with respect to $x_2^\nu$.

*Exercise 5.2.* Show that the differentiability of $f$ is also independent of the chart in $N$.

*Definition 5.2.* Let $f : M \to N$ be a homeomorphism and $\psi$ and $\varphi$ be coordinate functions as previously defined. If $\psi \circ f \circ \varphi^{-1}$ is invertible (that is, there exists a map $\varphi \circ f^{-1} \circ \psi^{-1}$) and both $y = \psi \circ f \circ \varphi^{-1}(x)$ and $x = \varphi \circ f^{-1} \circ \psi^{-1}(y)$ are $C^\infty$, $f$ is called a **diffeomorphism** and $M$ is said to be **diffeomorphic** to $N$ and *vice versa*, denoted by $M \equiv N$.

Clearly $\dim M = \dim N$ if $M \equiv N$. In chapter 2, we noted that homeomorphisms classify spaces according to whether it is possible to deform one space into another *continuously*. Diffeomorphisms classify spaces into equivalence classes according to whether it is possible to deform one space to another *smoothly*. Two diffeomorphic spaces are regarded as the same manifold. Clearly a diffeomorphism is a homeomorphism. What about the converse? Is a homeomorphism a diffeomorphism? In the previous section, we defined the differentiable structure as an equivalence class of atlases. Is it possible for a topological space to carry many differentiable structures? It is rather difficult to give examples of 'diffeomorphically inequivalent homeomorphisms' since it is known that this is possible only in higher-dimensional spaces ($\dim M \geq 4$). It was believed before 1956 that a topological space admits only one differentiable structure. However, Milnor (1956) pointed out that $S^7$ admits 28 differentiable structures. A recent striking discovery in mathematics is that $\mathbb{R}^4$ admits an infinite number of differentiable structures. Interested readers should consult Donaldson (1983) and Freed and Uhlenbeck (1984). Here we assume that a manifold admits a unique differentiable structure, for simplicity.

The set of diffeomorphisms $f : M \to M$ is a group denoted by Diff$(M)$. Take a point $p$ in a chart $(U, \varphi)$ such that $\varphi(p) = x^\mu(p)$. Under $f \in$ Diff$(M)$, $p$ is mapped to $f(p)$ whose coordinates are $\varphi(f(p)) = y^\mu(f(p))$ (we have assumed $f(p) \in U$). Clearly $y$ is a differentiable function of $x$; this is an *active* point of view to the coordinate transformation. However, if $(U, \varphi)$ and $(V, \psi)$ are overlapping charts, we have two coordinate values $x^\mu = \varphi(p)$ and $y^\mu = \psi(p)$ for a point $p \in U \cap V$. The map $x \mapsto y$ is differentiable by the assumed smoothness of the manifold; this reparametrization is a *passive* point of view to the coordinate transformation. We also denote the group of reparametrizations by Diff$(M)$.

Now we look at special classes of mappings, namely **curves** and **functions**. An open curve in an $m$-dimensional manifold $M$ is a map $c : (a, b) \to M$ where $(a, b)$ is an open interval such that $a < 0 < b$. We assume that the curve does not intersect with itself (figure 5.7). The number $a$ $(b)$ may be $-\infty$ $(+\infty)$ and we have included 0 in the interval for later convenience. If a curve is closed, it is

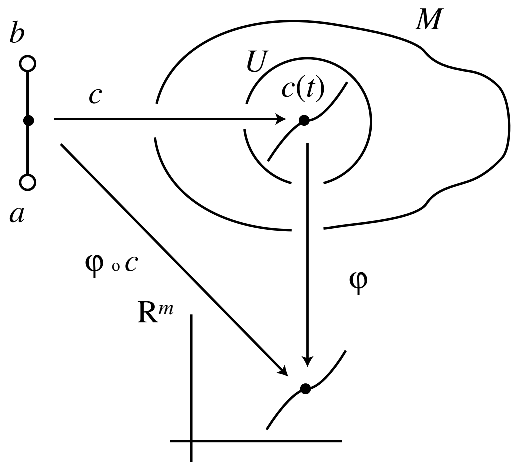


**Figure 5.7.** A curve $c$ in $M$ and its coordinate presentation $\varphi \circ c$.

regarded as a map $c : S^1 \to M$. In both cases, $c$ is locally a map from an open interval to $M$. On a chart $(U, \varphi)$, a curve $c(t)$ has the coordinate presentation $x = \varphi \circ c : \mathbb{R} \to \mathbb{R}^m$.

A function $f$ on $M$ is a smooth map from $M$ to $\mathbb{R}$, see figure 5.8. On a chart $(U, \varphi)$, the coordinate presentation of $f$ is given by $f \circ \varphi^{-1} : \mathbb{R}^m \to \mathbb{R}$ which is a real-valued function of $m$ variables. We denote the set of smooth functions on $M$ by $\mathcal{F}(M)$.

### 5.2.2 Vectors

Now that we have defined maps on a manifold, we are ready to define other geometrical objects: vectors, dual vectors and tensors. In general, an elementary picture of a vector as an arrow connecting a point and the origin does not work in a manifold. [Where is the origin? What is a straight arrow? How do we define a straight arrow that connects London and Los Angeles on the *surface* of the Earth?] On a manifold, a vector is defined to be a **tangent vector** to a curve in $M$.

To begin with, let us look at a tangent line to a curve in the $xy$-plane. If the curve is differentiable, we may approximate the curve in the vicinity of $x_0$ by

$$y - y(x_0) = a(x - x_0) \tag{5.17}$$

where $a = \mathrm{d}y/\mathrm{d}x|_{x=x_0}$. The tangent vectors on a manifold $M$ generalize this tangent line. To define a tangent vector we need a curve $c : (a, b) \to M$ and a function $f : M \to \mathbb{R}$, where $(a, b)$ is an open interval containing $t = 0$, see figure 5.9. We define the tangent vector at $c(0)$ as a directional derivative of a function $f(c(t))$ along the curve $c(t)$ at $t = 0$. The rate of change of $f(c(t))$ at

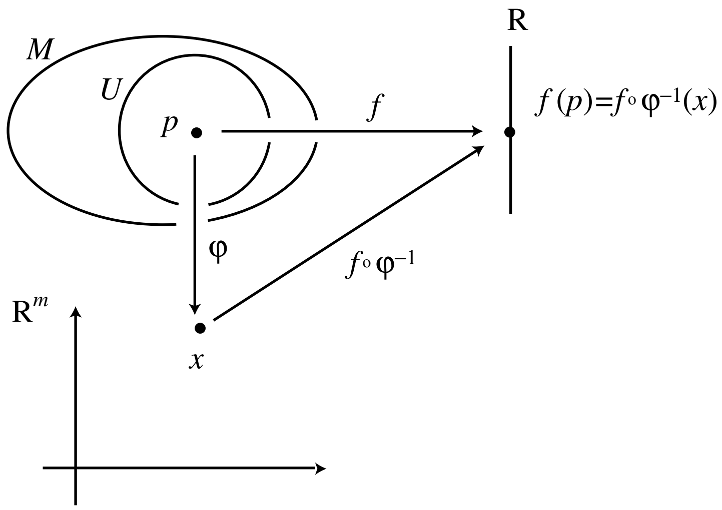


**Figure 5.8.** A function $f : M \to \mathbb{R}$ and its coordinate presentation $f \circ \varphi^{-1}$.

$t = 0$ along the curve is

$$\left.\frac{\mathrm{d}f(c(t))}{\mathrm{d}t}\right|_{t=0}. \tag{5.18}$$

In terms of the local coordinate, this becomes

$$\left.\frac{\partial f}{\partial x^{\mu}}\frac{\mathrm{d}x^{\mu}(c(t))}{\mathrm{d}t}\right|_{t=0}. \tag{5.19}$$

[Note the abuse of the notation! The derivative $\partial f/\partial x^{\mu}$ really means $\partial(f \circ \varphi^{-1}(x))/\partial x^{\mu}$.] In other words, $\mathrm{d}f(c(t))/\mathrm{d}t$ at $t = 0$ is obtained by applying the differential operator $X$ to $f$, where

$$X = X^{\mu}\left(\frac{\partial}{\partial x^{\mu}}\right) \qquad \left(X^{\mu} = \left.\frac{\mathrm{d}x^{\mu}(c(t))}{\mathrm{d}t}\right|_{t=0}\right) \tag{5.20}$$

that is,

$$\left.\frac{\mathrm{d}f(c(t))}{\mathrm{d}t}\right|_{t=0} = X^{\mu}\left(\frac{\partial f}{\partial x^{\mu}}\right) \equiv X[f]. \tag{5.21}$$

Here the last equality defines $X[f]$. It is $X = X^{\mu}\partial/\partial x^{\mu}$ which we now define as the tangent vector to $M$ at $p = c(0)$ along the direction given by the curve $c(t)$.

*Example 5.7.* If $X$ is applied to the coordinate functions $\varphi(c(t)) = x^{\mu}(t)$, we have

$$X[x^{\mu}] = \left(\frac{\mathrm{d}x^{\nu}}{\mathrm{d}t}\right)\left(\frac{\partial x^{\mu}}{\partial x^{\nu}}\right) = \left.\frac{\mathrm{d}x^{\mu}(t)}{\mathrm{d}t}\right|_{t=0}$$

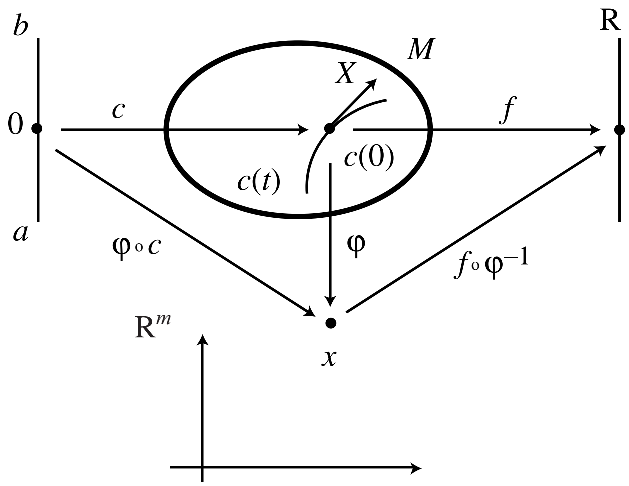


**Figure 5.9.** A curve $c$ and a function $f$ define a tangent vector along the curve in terms of the directional derivative.

which is the $\mu$th component of the velocity vector if $t$ is understood as time.

To be more mathematical, we introduce an equivalence class of curves in $M$. If two curves $c_1(t)$ and $c_2(t)$ satisfy

(i) $c_1(0) = c_2(0) = p$

(ii) $\left.\dfrac{\mathrm{d}x^\mu(c_1(t))}{\mathrm{d}t}\right|_{t=0} = \left.\dfrac{\mathrm{d}x^\mu(c_2(t))}{\mathrm{d}t}\right|_{t=0}$

$c_1(t)$ and $c_2(t)$ yield the same differential operator $X$ at $p$, in which case we define $c_1(t) \sim c_2(t)$. Clearly $\sim$ is an equivalence relation and defines the equivalence classes. We identify the *tangent vector* $X$ with the *equivalence class of curves*

$$[c(t)] = \left\{ \widetilde{c}(t) \middle| \widetilde{c}(0) = c(0) \text{ and } \left.\frac{\mathrm{d}x^\mu(\widetilde{c}(t))}{\mathrm{d}t}\right|_{t=0} = \left.\frac{\mathrm{d}x^\mu(c(t))}{\mathrm{d}t}\right|_{t=0} \right\} \tag{5.22}$$

rather than a curve itself.

All the equivalence classes of curves at $p \in M$, namely all the tangent vectors at $p$, form a vector space called the **tangent space** of $M$ at $p$, denoted by $T_pM$. To analyse $T_pM$, we may use the theory of vector spaces developed in section 2.2. Evidently, $e_\mu = \partial/\partial x^\mu$ $(1 \le \mu \le m)$ are the basis vectors of $T_pM$, see (5.20), and $\dim T_pM = \dim M$. The basis $\{e_\mu\}$ is called the **coordinate basis**. If a vector $V \in T_pM$ is written as $V = V^\mu e_\mu$, the numbers $V^\mu$ are called the components of $V$ with respect to $e_\mu$. By construction, it is obvious that a vector $X$ exists without specifying the coordinate, see (5.21). The assignment of the coordinate is simply for our convenience. This coordinate independence of a vector enables us to find the transformation property of the *components* of the vector. Let $p \in U_i \cap U_j$ and $x = \varphi_i(p)$, $y = \varphi_j(p)$. We have two expressions for $X \in T_pM$,

$$X = X^\mu \frac{\partial}{\partial x^\mu} = \widetilde{X}^\mu \frac{\partial}{\partial y^\mu}.$$

This shows that $X^\mu$ and $\widetilde{X}^\mu$ are related as

$$\widetilde{X}^\mu = X^\nu \frac{\partial y^\mu}{\partial x^\nu}. \tag{5.23}$$

Note again that the components of the vector transform in such a way that the vector itself is left invariant.

The basis of $T_pM$ need not be $\{e_\mu\}$, and we may think of the linear combinations $\hat{e}_i \equiv A_i{}^\mu e_\mu$, where $A = (A_i{}^\mu) \in \mathrm{GL}(m, \mathbb{R})$. The basis $\{\hat{e}_i\}$ is known as the **non-coordinate basis**.

### 5.2.3 One-forms

Since $T_pM$ is a vector space, there exists a dual vector space to $T_pM$, whose element is a linear function from $T_pM$ to $\mathbb{R}$, see section 2.2. The dual space is called the **cotangent space** at $p$, denoted by $T_p^*M$. An element $\omega : T_pM \to \mathbb{R}$ of $T_p^*M$ is called a **dual vector**, **cotangent vector** or, in the context of differential forms, a **one-form**. The simplest example of a one-form is the differential $\mathrm{d}f$ of a function $f \in \mathcal{F}(M)$. The action of a vector $V$ on $f$ is $V[f] = V^\mu \partial f/\partial x^\mu \in \mathbb{R}$. Then the action of $\mathrm{d}f \in T_p^*M$ on $V \in T_pM$ is defined by

$$\langle \mathrm{d}f, V\rangle \equiv V[f] = V^\mu \frac{\partial f}{\partial x^\mu} \in \mathbb{R}. \tag{5.24}$$

Clearly $\langle \mathrm{d}f, V\rangle$ is $\mathbb{R}$-linear in both $V$ and $f$.

Noting that $\mathrm{d}f$ is expressed in terms of the coordinate $x = \varphi(p)$ as $\mathrm{d}f = (\partial f/\partial x^\mu)\mathrm{d}x^\mu$, it is natural to regard $\{\mathrm{d}x^\mu\}$ as a basis of $T_p^*M$. Moreover, this is a dual basis, since

$$\left\langle \mathrm{d}x^\mu, \frac{\partial}{\partial x^\mu}\right\rangle = \frac{\partial x^\nu}{\partial x^\mu} = \delta^\nu_\mu. \tag{5.25}$$

An arbitrary one-form $\omega$ is written as

$$\omega = \omega_\mu \, \mathrm{d}x^\mu \tag{5.26}$$

where the $\omega_\mu$ are the components of $\omega$. Take a vector $V = V^\mu \partial/\partial x^\mu$ and a one-form $\omega = \omega_\mu \mathrm{d}x^\mu$. The **inner product** $\langle \ , \ \rangle : T_p^*M \times T_pM \to \mathbb{R}$ is defined by

$$\langle \omega, V\rangle = \omega_\mu V^\nu \left\langle \mathrm{d}x^\mu, \frac{\partial}{\partial x^\nu}\right\rangle = \omega_\mu V^\nu \delta^\mu_\nu = \omega_\mu V^\mu. \tag{5.27}$$

Note that the inner product is defined between a vector and a dual vector and not between two vectors or two dual vectors.

Since $\omega$ is defined without reference to any coordinate system, for a point $p \in U_i \cap U_j$, we have

$$\omega = \omega_\mu \mathrm{d}x^\mu = \widetilde{\omega}_\nu \, \mathrm{d}y^\nu$$

where $x = \varphi_i(p)$ and $y = \varphi_j(p)$. From $\mathrm{d}y^\nu = (\partial y^\nu / \partial x^\mu)\mathrm{d}x^\mu$ we find that

$$\widetilde{\omega}_\nu = \omega_\mu \frac{\partial x^\mu}{\partial y^\nu}. \tag{5.28}$$

### 5.2.4 Tensors

A **tensor** of type $(q, r)$ is a multilinear object which maps $q$ elements of $T_p^* M$ and $r$ elements of $T_p M$ to a real number. $\mathcal{T}^q_{r,p}(M)$ denotes the set of type $(q, r)$ tensors at $p \in M$. An element of $\mathcal{T}^q_{r,p}(M)$ is written in terms of the bases described earlier as

$$T = T^{\mu_1 \ldots \mu_q}{}_{\nu_1 \ldots \nu_r} \frac{\partial}{\partial x^{\mu_1}} \cdots \frac{\partial}{\partial x^{\mu_q}} \mathrm{d}x^{\nu_1} \ldots \mathrm{d}x^{\nu_r}. \tag{5.29}$$

Clearly this is a linear function from

$$\otimes^q T_p^* M \otimes^r T_p M$$

to $\mathbb{R}$. Let $V_i = V_i^\mu \partial / \partial x^\mu$ $(1 \le i \le r)$ and $\omega_i = \omega_{i\mu} \mathrm{d}x^\mu$ $(1 \le i \le q)$. The action of $T$ on them yields a number

$$T(\omega_1, \ldots, \omega_q; V_1, \ldots, V_r) = T^{\mu_1 \ldots \mu_q}{}_{\nu_1 \ldots \nu_r} \omega_{1\mu_1} \ldots \omega_{q\mu_q} V_1^{\nu_1} \ldots V_r^{\nu_r}.$$

In the present notation, the inner product is $\langle \omega, X \rangle = \omega(X)$.

### 5.2.5 Tensor fields

If a vector is assigned *smoothly* to each point of $M$, it is called a **vector field** over $M$. In other words, $V$ is a vector field if $V[f] \in \mathcal{F}(M)$ for any $f \in \mathcal{F}(M)$. Clearly each component of a vector field is a smooth function from $M$ to $\mathbb{R}$. The set of the vector fields on $M$ is denoted as $\mathcal{X}(M)$. A vector field $X$ at $p \in M$ is denoted by $X|_p$, which is an element of $T_p M$. Similarly, we define a **tensor field** of type $(q, r)$ by a smooth assignment of an element of $\mathcal{T}^q_{r,p}(M)$ at each point $p \in M$. The set of the tensor fields of type $(q, r)$ on $M$ is denoted by $\mathcal{T}^q_r(M)$. For example, $\mathcal{T}^0_1(M)$ is the set of the dual vector fields, which is also denoted by $\Omega^1(M)$ in the context of differential forms, see section 5.4. Similarly, $\mathcal{T}^0_0(M) = \mathcal{F}(M)$ is denoted by $\Omega^0(M)$ in the same context.

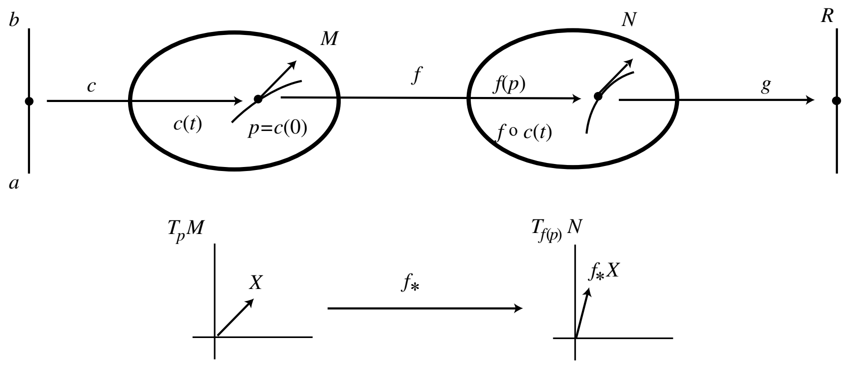


**Figure 5.10.** A map $f : M \to N$ induces the differential map $f_* : T_pM \to T_{f(p)}N$.

### 5.2.6 Induced maps

A smooth map $f : M \to N$ naturally induces a map $f_*$ called the **differential map** (figure 5.10),

$$f_* : T_pM \to T_{f(p)}N. \tag{5.30}$$

The explicit form of $f_*$ is obtained by the definition of a tangent vector as a directional derivative along a curve. If $g \in \mathcal{F}(N)$, then $g \circ f \in \mathcal{F}(M)$. A vector $V \in T_pM$ acts on $g \circ f$ to give a number $V[g \circ f]$. Now we define $f_*V \in T_{f(p)}N$ by

$$(f_*V)[g] \equiv V[g \circ f] \tag{5.31}$$

or, in terms of charts $(U, \varphi)$ on $M$ and $(V.\psi)$ on $N$,

$$(f_*V)[g \circ \psi^{-1}(y)] \equiv V[g \circ f \circ \varphi^{-1}(x)] \tag{5.32}$$

where $x = \varphi(p)$ and $y = \psi(f(p))$. Let $V = V^\mu \partial/\partial x^\mu$ and $f_*V = W^\alpha \partial/\partial y^\alpha$. Then (5.32) yields

$$W^\alpha \frac{\partial}{\partial y^\alpha}[g \circ \psi^{-1}(y)] = V^\mu \frac{\partial}{\partial x^\mu}[g \circ f \circ \varphi^{-1}(x)].$$

If we take $g = y^\alpha$, we obtain the relation between $W^\alpha$ and $V^\mu$,

$$W^\alpha = V^\mu \frac{\partial}{\partial x^\mu} y^\alpha(x). \tag{5.33}$$

Note that the matrix $(\partial y^\alpha/\partial x^\mu)$ is nothing but the Jacobian of the map $f : M \to N$. The differential map $f_*$ is naturally extended to tensors of type $(q, 0)$, $f_* : \mathcal{T}^q_{0,p}(M) \to \mathcal{T}^q_{0,f(p)}(N)$.

*Example 5.8.* Let $(x^1, x^2)$ and $(y^1, y^2, y^3)$ be the coordinates in $M$ and $N$, respectively, and let $V = a\partial/\partial x^1 + b\partial/\partial x^2$ be a tangent vector at $(x^1, x^2)$.

Let $f : M \to N$ be a map whose coordinate presentation is $y = (x^1, x^2, \sqrt{1-(x^1)^2-(x^2)^2})$. Then

$$f_* V = V^\mu \frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial}{\partial y^\alpha} = a\frac{\partial}{\partial y^1} + b\frac{\partial}{\partial y^2} - \left(a\frac{y^1}{y^3} + b\frac{y^2}{y^3}\right)\frac{\partial}{\partial y^3}.$$

*Exercise 5.3.* Let $f : M \to N$ and $g : N \to P$. Show that the differential map of the composite map $g \circ f : M \to P$ is

$$(g \circ f)_* = g_* \circ f_*. \tag{5.34}$$

A map $f : M \to N$ also induces a map

$$f^* : T^*_{f(p)}N \to T^*_p M. \tag{5.35}$$

Note that $f_*$ goes in the same direction as $f$, while $f^*$ goes backward, hence the name **pullback**, see section 2.2. If we take $V \in T_pM$ and $\omega \in T^*_{f(p)}N$, the pullback of $\omega$ by $f^*$ is defined by

$$\langle f^*\omega, V\rangle = \langle \omega, f_*V\rangle. \tag{5.36}$$

The pullback $f^*$ naturally extends to tensors of type $(0, r)$, $f^* : \mathcal{T}^0_{r,f(p)}(N) \to \mathcal{T}^0_{r,p}(M)$. The component expression of $f^*$ is given by the Jacobian matrix $(\partial y^\alpha/\partial x^\mu)$, see exercise 5.4.

*Exercise 5.4.* Let $f : M \to N$ be a smooth map. Show that for $\omega = \omega_\alpha \mathrm{d}y^\alpha \in T^*_{f(p)}N$, the induced one-form $f^*\omega = \xi_\mu\, \mathrm{d}x^\mu \in T^*_pM$ has components

$$\xi_\mu = \omega_\alpha \frac{\partial y^\alpha}{\partial x^\mu}. \tag{5.37}$$

*Exercise 5.5.* Let $f$ and $g$ be as in exercise 5.3. Show that the pullback of the composite map $g \circ f$ is

$$(g \circ f)^* = f^* \circ g^*. \tag{5.38}$$

There is no natural extension of the induced map for a tensor of mixed type. The extension is only possible if $f : M \to N$ is a diffeomorphism, where the Jacobian of $f^{-1}$ is also defined.

*Exercise 5.6.* Let

$$T^\mu{}_\nu \frac{\partial}{\partial x^\mu} \otimes \mathrm{d}x^\nu$$

be a tensor field of type $(1, 1)$ on $M$ and let $f : M \to N$ be a diffeomorphism. Show that the induced tensor on $N$ is

$$f_*\left(T^\mu{}_\nu \frac{\partial}{\partial x^\mu} \otimes \mathrm{d}x^\nu\right) = T^\mu{}_\nu \left(\frac{\partial y^\alpha}{\partial x^\mu}\right)\left(\frac{\partial x^\nu}{\partial y^\beta}\right)\frac{\partial}{\partial y^\alpha} \otimes \mathrm{d}y^\beta$$

where $x^\mu$ and $y^\alpha$ are local coordinates in $M$ and $N$, respectively.

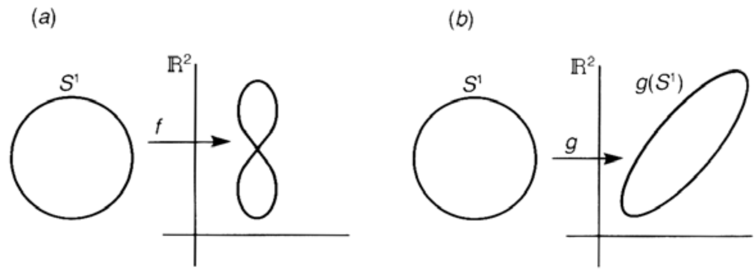


**Figure 5.11.** (*a*) An immersion $f$ which is not an embedding. (*b*) An embedding $g$ and the submanifold $g(S^1)$.

### 5.2.7 Submanifolds

Before we close this section, we define a submanifold of a manifold. The meaning of embedding is also clarified here.

*Definition 5.3.* (**Immersion, submanifold, embedding**) Let $f : M \to N$ be a smooth map and let $\dim M \le \dim N$.

(a) The map $f$ is called an **immersion** of $M$ into $N$ if $f_* : T_pM \to T_{f(p)}N$ is an injection (one to one), that is rank $f_* = \dim M$.
(b) The map $f$ is called an **embedding** if $f$ is an injection and an immersion. The image $f(M)$ is called a **submanifold** of $N$. [In practice, $f(M)$ thus defined is diffeomorphic to $M$.]

If $f$ is an immersion, $f^*$ maps $T_pM$ isomorphically to an $m$-dimensional vector subspace of $T_{f(p)}N$ since rank$f_* = \dim M$. From theorem 2.1, we also find $\ker f_* = \{0\}$. If $f$ is an embedding, $M$ is diffeomorphic to $f(M)$. Examples will clarify these rather technical points. Consider a map $f : S^1 \to \mathbb{R}^2$ in figure 5.11($a$). It is an immersion since a one-dimensional tangent space of $S^1$ is mapped by $f_*$ to a subspace of $T_{f(p)}\mathbb{R}^2$. The image $f(S^1)$ is not a submanifold of $\mathbb{R}^2$ since $f$ is not an injection. The map $g : S^1 \to \mathbb{R}^2$ in figure 5.11($b$) is an embedding and $g(S^1)$ is a submanifold of $\mathbb{R}^2$. Clearly, an embedding is an immersion although the converse is not necessarily true. In the previous section, we occasionally mentioned the embedding of $S^n$ into $\mathbb{R}^{n+1}$. Now this meaning is clear; if $S^n$ is embedded by $f : S^n \to \mathbb{R}^{n+1}$ then $S^n$ is diffeomorphic to $f(S^n)$.

## 5.3 Flows and Lie derivatives

Let $X$ be a vector field in $M$. An integral curve $x(t)$ of $X$ is a curve in $M$, whose tangent vector at $x(t)$ is $X|_x$. Given a chart $(U, \varphi)$, this means

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = X^\mu(x(t)) \tag{5.39}$$

where $x^\mu(t)$ is the $\mu$th component of $\varphi(x(t))$ and $X = X^\mu \partial/\partial x^\mu$. Note the abuse of the notation: $x$ is used to denote a point in $M$ as well as its coordinates. [For later convenience we assume the point $x(0)$ is included in $U$.] Put in another way, finding the integral curve of a vector field $X$ is equivalent to solving the autonomous system of ordinary differential equations (ODEs) (5.39). The initial condition $x_0^\mu = x^\mu(0)$ corresponds to the coordinates of an integral curve at $t = 0$. The existence and uniqueness theorem of ODEs guarantees that there is a unique solution to (5.39), at least locally, with the initial data $x_0^\mu$. It may happen that the integral curve is defined only on a subset of $\mathbb{R}$, in which case we have to pay attention so that the parameter $t$ does not exceed the given interval. In the following we assume that $t$ is maximally extended. It is known that if $M$ is a compact manifold, the integral curve exists for all $t \in \mathbb{R}$.

Let $\sigma(t, x_0)$ be an integral curve of $X$ which passes a point $x_0$ at $t = 0$ and denote the coordinate by $\sigma^\mu(t, x_0)$. Equation (5.39) then becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\sigma^\mu(t, x_0) = X^\mu(\sigma(t, x_0)) \tag{5.40a}$$

with the initial condition

$$\sigma^\mu(0, x_0) = x_0^\mu. \tag{5.40b}$$

The map $\sigma : \mathbb{R} \times M \to M$ is called a **flow** generated by $X \in \mathcal{X}(M)$. A flow satisfies the rule

$$\sigma(t, \sigma^\mu(s, x_0)) = \sigma(t + s, x_0) \tag{5.41}$$

for any $s, t \in \mathbb{R}$ such that both sides of (5.41) make sense. This can be seen from the uniqueness of ODEs. In fact, we note that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\sigma^\mu(t, \sigma^\mu(s, x_0)) &= X^\mu(\sigma(t, \sigma^\mu(s, x_0))) \\ \sigma(0, \sigma(s, x_0)) &= \sigma(s, x_0) \end{aligned}$$

and

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\sigma^\mu(t + s, x_0) &= \frac{\mathrm{d}}{\mathrm{d}(t + s)}\sigma^\mu(t + s, x_0) = X^\mu(\sigma(t + s, x_0)) \\ \sigma(0 + s, x_0) &= \sigma(s, x_0). \end{aligned}$$

Thus, both sides of (5.41) satisfy the same ODE and the same initial condition. From the uniqueness of the solution, they should be the same. We have obtained the following theorem.

*Theorem 5.1.* For any point $x \in M$, there exists a differentiable map $\sigma : \mathbb{R} \times M \to M$ such that

(i) $\sigma(0, x) = x$;
(ii) $t \mapsto \sigma(t, x)$ is a solution of (5.40a) and (5.40b); and

(iii) $\sigma(t, \sigma^{\mu}(s, x)) = \sigma(t+s, x)$.

[*Note:* We denote the initial point by $x$ instead of $x_0$ to emphasize that $\sigma$ is a map $\mathbb{R} \times M \to M$.]

We may imagine a flow as a (steady) stream flow. If a particle is observed at a point $x$ at $t = 0$, it will be found at $\sigma(t, x)$ at later time $t$.

*Example 5.9.* Let $M = \mathbb{R}^2$ and let $X((x, y)) = -y\partial/\partial x + x\partial/\partial y$ be a vector field in $M$. It is easy to verify that

$$\sigma(t, (x, y)) = (x\cos t - y\sin t, x\sin t + y\cos t)$$

is a flow generated by $X$. The flow through $(x, y)$ is a circle whose centre is at the origin. Clearly, $\sigma(t, (x, y)) = (x, y)$ if $t = 2n\pi, n \in \mathbb{Z}$. If $(x, y) = (0, 0)$, the flow stays at $(0, 0)$.

*Exercise 5.7.* Let $M = \mathbb{R}^2$, and let $X = y\partial/\partial x + x\partial/\partial y$ be a vector field in $M$. Find the flow generated by $X$.

### 5.3.1 One-parameter group of transformations

For fixed $t \in \mathbb{R}$, a flow $\sigma(t, x)$ is a diffeomorphism from $M$ to $M$, denoted by $\sigma_t : M \to M$. It is important to note that $\sigma_t$ is made into a *commutative group* by the following rules.

(i) $\sigma_t(\sigma_s(x)) = \sigma_{t+s}(x)$, that is, $\sigma_t \circ \sigma_s = \sigma_{t+s}$;
(ii) $\sigma_0 =$ the identity map (= unit element); and
(iii) $\sigma_{-t} = (\sigma_t)^{-1}$.

This group is called the **one-parameter group of transformations**. The group locally looks like the additive group $\mathbb{R}$, although it may not be isomorphic to $\mathbb{R}$ globally. In fact, in example 5.9, $\sigma_{2\pi n+t}$ was the same map as $\sigma_t$ and we find that the one-parameter group is isomorphic to SO(2), the multiplicative group of $2 \times 2$ real matrices of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

or U(1), the multiplicative group of complex numbers of unit modulus $\mathrm{e}^{\mathrm{i}\theta}$.

Under the action of $\sigma_\varepsilon$, with an infinitesimal $\varepsilon$, we find from (5.40a) and (5.40b) that a point $x$ whose coordinate is $x^\mu$ is mapped to

$$\sigma^{\mu}_{\varepsilon}(x) = \sigma^{\mu}(\varepsilon, x) = x^{\mu} + \varepsilon X^{\mu}(x). \tag{5.42}$$

The vector field $X$ is called, in this context, the **infinitesimal generator** of the transformation $\sigma_t$.

Given a vector field $X$, the corresponding flow $\sigma$ is often referred to as the **exponentiation** of $X$ and is denoted by

$$\sigma^{\mu}(t, x) = \exp(tX)x^{\mu}. \tag{5.43}$$

The name 'exponentiation' is justified as we shall see now. Let us take a parameter $t$ and evaluate the coordinate of a point which is separated from the initial point $x = \sigma(0, x)$ by the parameter distance $t$ along the flow $\sigma$. The coordinate corresponding to the point $\sigma(t, x)$ is

$$\begin{aligned}\sigma^{\mu}(t, x) &= x^{\mu} + t\frac{\mathrm{d}}{\mathrm{d}s}\sigma^{\mu}(s, x)\bigg|_{s=0} + \frac{t^2}{2!}\left(\frac{\mathrm{d}}{\mathrm{d}s}\right)^2 \sigma^{\mu}(s, x)\bigg|_{s=0} + \cdots \\ &= \left[1 + t\frac{\mathrm{d}}{\mathrm{d}s} + \frac{t^2}{2!}\left(\frac{\mathrm{d}}{\mathrm{d}s}\right)^2 + \cdots\right]\sigma^{\mu}(s, x)\bigg|_{s=0} \\ &\equiv \exp\left(t\frac{\mathrm{d}}{\mathrm{d}s}\right)\sigma^{\mu}(s, x)\bigg|_{s=0}.\end{aligned} \tag{5.44}$$

The last expression can also be written as $\sigma^{\mu}(t, x) = \exp(tX)x^{\mu}$, as in (5.43). The flow $\sigma$ satisfies the following exponential properties.

$$\text{(i)} \qquad \sigma(0, x) = x = \exp(0X)x \tag{5.45a}$$

$$\text{(ii)} \qquad \frac{\mathrm{d}\sigma(t, x)}{\mathrm{d}t} = X\exp(tX)x = \frac{\mathrm{d}}{\mathrm{d}t}[\exp(tX)x] \tag{5.45b}$$

$$\begin{aligned}\text{(iii)} \qquad \sigma(t, \sigma(s, x)) &= \sigma(t, \exp(sX)x) = \exp(tX)\exp(sX)x \\ &= \exp\{(t + s)X\}x = \sigma(t + s, x).\end{aligned} \tag{5.45c}$$

### 5.3.2 Lie derivatives

Let $\sigma(t, x)$ and $\tau(t, x)$ be two flows generated by the vector fields $X$ and $Y$,

$$\frac{\mathrm{d}\sigma^{\mu}(s, x)}{\mathrm{d}s} = X^{\mu}(\sigma(s, x)) \tag{5.46a}$$

$$\frac{\mathrm{d}\tau^{\mu}(t, x)}{\mathrm{d}t} = Y^{\mu}(\tau(t, x)). \tag{5.46b}$$

Let us evaluate the change of the vector field $Y$ along $\sigma(s, x)$. To do this, we have to compare the vector $Y$ at a point $x$ with that at a nearby point $x' = \sigma_{\varepsilon}(x)$, see figure 5.12. However, we cannot simply take the difference between the components of $Y$ at two points since they belong to different tangent spaces $T_pM$ and $T_{\sigma_{\varepsilon}(x)}M$; the naive difference between vectors at different points is ill defined. To define a sensible derivative, we first map $Y|_{\sigma_{\varepsilon}(x)}$ to $T_xM$ by $(\sigma_{-\varepsilon})_* : T_{\sigma_{\varepsilon}(x)}M \to T_xM$, after which we take a difference between two vectors $(\sigma_{-\varepsilon})_* Y|_{\sigma_{\varepsilon}(x)}$ and $Y|_x$, both of which are vectors in $T_xM$. The **Lie derivative** of a vector field $Y$ along the flow $\sigma$ of $X$ is defined by

$$\mathcal{L}_X Y = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}[(\sigma_{-\varepsilon})_* Y|_{\sigma_{\varepsilon}(x)} - Y|_x]. \tag{5.47}$$

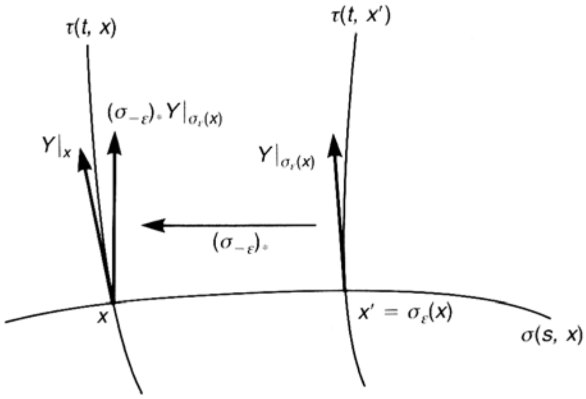


**Figure 5.12.** To compare a vector $Y|_x$ with $Y|_{\sigma_\varepsilon(x)}$, the latter must be transported back to $x$ by the differential map $(\sigma_{-\varepsilon})_*$.

*Exercise 5.8.* Show that $\mathcal{L}_X Y$ is also written as

$$\begin{aligned}\mathcal{L}_X Y &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[Y|_x - (\sigma_\varepsilon)_* Y|_{\sigma_{-\varepsilon}(x)}]\\ &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[Y|_{\sigma_\varepsilon(x)} - (\sigma_\varepsilon)_* Y|_x].\end{aligned}$$

Let $(U, \varphi)$ be a chart with the coordinates $x$ and let $X = X^\mu \partial/\partial x^\mu$ and $Y = Y^\mu \partial/\partial x^\mu$ be vector fields defined on $U$. Then $\sigma_\varepsilon(x)$ has the coordinates $x^\mu + \varepsilon X^\mu(x)$ and

$$\begin{aligned}Y|_{\sigma_\varepsilon(x)} &= Y^\mu(x^\nu + \varepsilon X^\nu(x)) e_\mu|_{x+\varepsilon X}\\ &\simeq [Y^\mu(x) + \varepsilon X^\mu(x)\partial_\nu Y^\mu(x)] e_\mu|_{x+\varepsilon X}\end{aligned}$$

where $\{e_\mu\} = \{\partial/\partial x^\mu\}$ is the coordinate basis and $\partial_\nu \equiv \partial/\partial x^\nu$. If we map this vector defined at $\sigma_\varepsilon(x)$ to $x$ by $(\sigma_{-\varepsilon})_*$, we obtain

$$\begin{aligned}&[Y^\mu(x) + \varepsilon X^\lambda(x)\partial_\lambda Y^\mu(x)]\partial_\mu[x^\nu - \varepsilon X^\nu(x)] e_\nu|_x\\ &\quad= [Y^\mu(x) + \varepsilon X^\lambda(x)\partial_\lambda Y^\mu(x)][\delta_\mu^\nu - \varepsilon\partial_\mu X^\nu(x)] e_\nu|_x\\ &\quad= Y^\mu(x) e_\mu|_x + \varepsilon[X^\mu(x)\partial_\mu Y^\nu(x) - Y^\mu(x)\partial_\mu X^\nu(x)] e_\nu|_x + O(\varepsilon^2).\end{aligned} \tag{5.48}$$

From (5.47) and (5.48), we find that

$$\mathcal{L}_X Y = (X^\mu \partial_\mu Y^\nu - Y^\mu \partial_\mu X^\nu) e_\nu. \tag{5.49a}$$

*Exercise 5.9.* Let $X = X^\mu \partial/\partial x^\mu$ and $Y = Y^\mu \partial/\partial x^\mu$ be vector fields in $M$. Define the **Lie bracket** $[X, Y]$ by

$$[X, Y]f = X[Y[f]] - Y[X[f]] \tag{5.50}$$

where $f \in \mathcal{F}(M)$. Show that $[X, Y]$ is a vector field given by

$$(X^\mu \partial_\mu Y^\nu - Y^\mu \partial_\mu X^\nu)e_\nu.$$

This exercise shows that the Lie derivative of $Y$ along $X$ is

$$\mathcal{L}_X Y = [X, Y]. \tag{5.49b}$$

[*Remarks:* Note that neither $XY$ nor $YX$ is a vector field since they are second-order derivatives. The combination $[X, Y]$ is, however, a first-order derivative and indeed a vector field.]

*Exercise 5.10.* Show that the Lie bracket satisfies

(a) bilinearity

$$\begin{aligned}[X, c_1Y_1 + c_2Y_2] &= c_1[X, Y_1] + c_2[X, Y_2]\\ [c_1X_1 + c_2X_2, Y] &= c_1[X_1, Y] + c_2[X_2, Y]\end{aligned}$$

for any constants $c_1$ and $c_2$,

(b) skew-symmetry

$$[X, Y] = -[YX]$$

(c) the Jacobi identity

$$[[X, Y], Z] + [[Z, X], Y] + [[Y, Z], X] = 0.$$

*Exercise 5.11.* (a) Let $X, Y \in \mathcal{X}(M)$ and $f \in \mathcal{F}(M)$. Show that

$$\mathcal{L}_{fX} Y = f[X, Y] - Y[f]X \tag{5.51a}$$

$$\mathcal{L}_X(fY) = f[X, Y] + X[f]Y. \tag{5.51b}$$

(b) Let $X, Y \in \mathcal{X}(M)$ and $f : M \to N$. Show that

$$f_*[X, Y] = [f_*X, f_*Y]. \tag{5.52}$$

Geometrically, the Lie bracket shows the non-commutativity of two flows. This is easily observed from the following consideration. Let $\sigma(s, x)$ and $\tau(t, x)$ be two flows generated by vector fields $X$ and $Y$, as before, see figure 5.13. If we move by a small parameter distance $\varepsilon$ along the flow $\sigma$ first, then by $\delta$ along $\tau$, we shall be at the point whose coordinates are

$$\begin{aligned}\tau^\mu(\delta, \sigma(\varepsilon, x)) &\simeq \tau^\mu(\delta, x^\nu + \varepsilon X^\nu(x))\\ &\simeq x^\mu + \varepsilon X^\mu(x) + \delta Y^\mu(x^\nu + \varepsilon X^\nu(x))\\ &\simeq x^\mu + \varepsilon X^\mu(x) + \delta Y^\mu(x) + \varepsilon\delta X^\nu(x)\partial_\nu Y^\nu(x).\end{aligned}$$

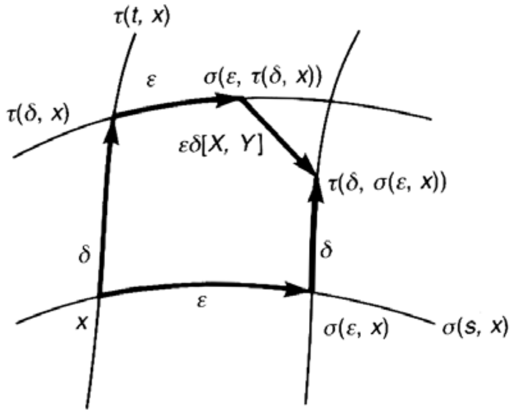


**Figure 5.13.** A Lie bracket $[X, Y]$ measures the failure of the closure of the parallelogram.

If, however, we move by $\delta$ along $\tau$ first, then by $\varepsilon$ along $\sigma$, we will be at the point

$$\begin{aligned}\sigma^\mu(\varepsilon, \tau(\delta, x)) &\simeq \sigma^\mu(\varepsilon, x^\nu + \delta Y^\nu(x)) \\ &\simeq x^\mu + \delta Y^\mu(x) + \varepsilon X^\mu(x^\nu + \delta Y^\nu(x)) \\ &\simeq x^\mu + \delta Y^\mu(x) + \varepsilon X^\mu(x) + \varepsilon\delta Y^\nu(x)\partial_\nu X^\mu(x).\end{aligned}$$

The difference between the coordinates of these two points is proportional to the Lie bracket,

$$\tau^\mu(\delta, \sigma(\varepsilon, x)) - \sigma^\mu(\varepsilon, \tau(\delta, x)) = \varepsilon\delta[X, Y]^\mu.$$

The Lie bracket of $X$ and $Y$ measures the failure of the closure of the parallelogram in figure 5.13. It is easy to see $\mathcal{L}_X Y = [X, Y] = 0$ if and only if

$$\sigma(s, \tau(t, x)) = \tau(t, \sigma(s, x)). \tag{5.53}$$

We may also define the Lie derivative of a one-form $\omega \in \Omega^1(M)$ along $X \in \mathcal{X}(M)$ by

$$\mathcal{L}_X\omega \equiv \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}[(\sigma_\varepsilon)^*\omega|_{\sigma_\varepsilon(x)} - \omega|_x] \tag{5.54}$$

where $\omega|_x \in T_x^*M$ is $\omega$ at $x$. Put $\omega = \omega_\mu \mathrm{d}x^\mu$. Repeating a similar analysis as before, we obtain

$$(\sigma_\varepsilon)^*\omega|_{\sigma_\varepsilon(x)} = \omega_\mu(x)\,\mathrm{d}x^\mu + \varepsilon[X^\nu(x)\partial_\nu\omega_\mu(x) + \partial_\mu X^\nu(x)\omega_\nu(x)]\,\mathrm{d}x^\mu$$

which leads to

$$\mathcal{L}_X\omega = (X^\nu\partial_\nu\omega_\mu + \partial_\mu X^\nu\omega_\nu)\,\mathrm{d}x^\mu. \tag{5.55}$$

Clearly $\mathcal{L}_X\omega \in T_x^*(M)$, since it is a difference of two one-forms at the same point $x$.

The Lie derivative of $f \in \mathcal{F}(M)$ along a flow $\sigma_s$ generated by a vector field $X$ is

$$\begin{aligned}\mathcal{L}_X f &\equiv \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[f(\sigma_\varepsilon(x)) - f(x)]\\ &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[f(x^\mu + \varepsilon X^\mu(x)) - f(x^\mu)]\\ &= X^\mu(x)\frac{\partial f}{\partial x^\mu} = X[f]\end{aligned} \tag{5.56}$$

which is the usual directional derivative of $f$ along $X$.

The Lie derivative of a general tensor is obtained from the following proposition.

*Proposition 5.1.* The Lie derivative satisfies

$$\mathcal{L}_X(t_1 + t_2) = \mathcal{L}_X t_1 + \mathcal{L}_X t_2 \tag{5.57a}$$

where $t_1$ and $t_2$ are tensor fields of the same type and

$$\mathcal{L}_X(t_1 \otimes t_2) = (\mathcal{L}_X t_1) \otimes t_2 + t_1 \otimes (\mathcal{L}_X t_2) \tag{5.57b}$$

where $t_1$ and $t_2$ are tensor fields of arbitrary types.

*Proof.* (a) is obvious. Rather than giving the general proof of (b), which is full of indices, we give an example whose extension to more general cases is trivial. Take $Y \in \mathcal{X}(M)$ and $\omega \in \Omega^1(M)$ and construct the tensor product $Y \otimes \omega$. Then $(Y \otimes \omega)|_{\sigma_\varepsilon(x)}$ is mapped onto a tensor at $x$ by the action of $(\sigma_{-\varepsilon})_* \otimes (\sigma_\varepsilon)^*$:

$$[(\sigma_{-\varepsilon})_* \otimes (\sigma_\varepsilon)^*](Y \otimes \omega)|_{\sigma_\varepsilon(x)} = [(\sigma_{-\varepsilon})_* Y \otimes (\sigma_\varepsilon)^*\omega]|_x.$$

Then there follows (the Leibnitz rule):

$$\begin{aligned}\mathcal{L}_X(Y \otimes \omega) &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[\{(\sigma_{-\varepsilon})_* Y \otimes (\sigma_\varepsilon)^*\omega\}|_x - (Y \otimes \omega)|_x]\\ &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[(\sigma_{-\varepsilon})_* Y \otimes \{(\sigma_\varepsilon)^*\omega - \omega\} + \{(\sigma_{-\varepsilon})_* Y - Y\} \otimes \omega]\\ &= Y \otimes (\mathcal{L}_X\omega) + (\mathcal{L}_X Y) \otimes \omega.\end{aligned}$$

Extensions to more general cases are obvious. □

This proposition enables us to calculate the Lie derivative of a general tensor field. For example, let $t = t_\mu{}^\nu\, \mathrm{d}x^\mu \otimes e_\nu \in \mathcal{T}^1_1(M)$. Proposition 5.1 gives

$$\mathcal{L}_X t = X[t_\mu{}^\nu]\, \mathrm{d}x^\mu \otimes e_\nu + t_\mu{}^\nu(\mathcal{L}_X\, \mathrm{d}x^\mu) \otimes e_\nu + t_\mu{}^\nu\, \mathrm{d}x^\mu \otimes (\mathcal{L}_X e_\nu).$$

*Exercise 5.12.* Let $t$ be a tensor field. Show that

$$\mathcal{L}_{[X,Y]}t = \mathcal{L}_X\mathcal{L}_Y t - \mathcal{L}_Y\mathcal{L}_X t. \tag{5.58}$$

## 5.4 Differential forms

Before we define differential forms, we examine the symmetry property of tensors. The symmetry operation on a tensor $\omega \in \mathcal{T}^0_{r,p}(M)$ is defined by

$$P\omega(V_1, \ldots, V_r) \equiv \omega(V_{P(1)}, \ldots, V_{P(r)}) \tag{5.59}$$

where $V_i \in T_pM$ and $P$ is an element of $S_r$, the **symmetric group** of order $r$. Take the coordinate basis $\{e_\mu\} = \{\partial/\partial x^\mu\}$. The component of $\omega$ in this basis is

$$\omega(e_{\mu_1}, e_{\mu_2}, \ldots, e_{\mu_r}) = \omega_{\mu_1\mu_2\ldots\mu_r}.$$

The component of $P\omega$ is obtained from (5.59) as

$$P\omega(e_{\mu_1}, e_{\mu_2}, \ldots, e_{\mu_r}) = \omega_{\mu_{P(1)}\mu_{P(2)}\ldots\mu_{P(r)}}.$$

For a general tensor of type $(q, r)$, the symmetry operations are defined for $q$ indices and $r$ indices separately.

For $\omega \in \mathcal{T}^0_{r,p}(M)$, the **symmetrizer** $\mathcal{S}$ is defined by

$$\mathcal{S}\omega = \frac{1}{r!}\sum_{P\in S_r} P\omega \tag{5.60}$$

while the **anti-symmetrizer** $\mathcal{A}$ is

$$\mathcal{A}\omega = \frac{1}{r!}\sum_{P\in S_r} \mathrm{sgn}(P)P\omega \tag{5.61}$$

where $\mathrm{sgn}(P) = +1$ for even permutations and $-1$ for odd permutations. $\mathcal{S}\omega$ is *totally symmetric* (that is, $P\mathcal{S}\omega = \mathcal{S}\omega$ for any $P \in S_r$) and $\mathcal{A}\omega$ is *totally anti-symmetric* ($P\mathcal{A}\omega = \mathrm{sgn}(P)\mathcal{A}\omega$).

### 5.4.1 Definitions

*Definition 5.4.* A **differential form** of order $r$ or an ***r*-form** is a totally anti-symmetric tensor of type $(0, r)$.

Let us define the **wedge product** $\wedge$ of $r$ one-forms by the totally anti-symmetric tensor product

$$\mathrm{d}x^{\mu_1} \wedge \mathrm{d}x^{\mu_2} \wedge \ldots \wedge \mathrm{d}x^{\mu_r} = \sum_{P\in S_r} \mathrm{sgn}(P)\, \mathrm{d}x^{\mu_{P(1)}} \wedge \mathrm{d}x^{\mu_{P(2)}} \wedge \ldots \wedge \mathrm{d}x^{\mu_{P(r)}}. \tag{5.62}$$

For example,

$$\begin{aligned}
\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu &= \mathrm{d}x^\mu \otimes \mathrm{d}x^\nu - \mathrm{d}x^\nu \otimes \mathrm{d}x^\mu \\
\mathrm{d}x^\lambda \wedge \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu &= \mathrm{d}x^\lambda \otimes \mathrm{d}x^\mu \otimes \mathrm{d}x^\nu + \mathrm{d}x^\nu \otimes \mathrm{d}x^\lambda \otimes \mathrm{d}x^\mu \\
&\quad + \mathrm{d}x^\mu \otimes \mathrm{d}x^\nu \otimes \mathrm{d}x^\lambda - \mathrm{d}x^\lambda \otimes \mathrm{d}x^\nu \otimes \mathrm{d}x^\mu \\
&\quad - \mathrm{d}x^\nu \otimes \mathrm{d}x^\mu \otimes \mathrm{d}x^\lambda - \mathrm{d}x^\mu \otimes \mathrm{d}x^\lambda \otimes \mathrm{d}x^\nu.
\end{aligned}$$

It is readily verified that the wedge product satisfies the following.

(i) $\mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r} = 0$ if some index $\mu$ appears at least twice.
(ii) $\mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r} = \mathrm{sgn}(P)\, \mathrm{d}x^{\mu_{P(1)}} \wedge \ldots \wedge \mathrm{d}x^{\mu_{P(r)}}$.
(iii) $\mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r}$ is linear in each $\mathrm{d}x^{\mu}$.

If we denote the vector space of $r$-forms at $p \in M$ by $\Omega^r_p(M)$, the set of $r$-forms (5.62) forms a basis of $\Omega^r_p(M)$ and an element $\omega \in \Omega^r_p(M)$ is expanded as

$$\omega = \frac{1}{r!}\omega_{\mu_1\mu_2\ldots\mu_r}\, \mathrm{d}x^{\mu_1} \wedge \mathrm{d}x^{\mu_2} \wedge \ldots \wedge \mathrm{d}x^{\mu_r} \tag{5.63}$$

where $\omega_{\mu_1\mu_2\ldots\mu_r}$ are taken *totally anti-symmetric*, reflecting the anti-symmetry of the basis. For example, the components of any second-rank tensor $\omega_{\mu\nu}$ are decomposed into the symmetric part $\sigma_{\mu\nu}$ and the anti-symmetric part $\alpha_{\mu\nu}$:

$$\sigma_{\mu\nu} = \omega_{(\mu\nu)} \equiv \tfrac{1}{2}(\omega_{\mu\nu} + \omega_{\nu\mu}) \tag{5.64a}$$

$$\alpha_{\mu\nu} = \omega_{[\mu\nu]} \equiv \tfrac{1}{2}(\omega_{\mu\nu} - \omega_{\nu\mu}). \tag{5.64b}$$

Observe that $\sigma_{\mu\nu}\, \mathrm{d}x^{\mu} \wedge \mathrm{d}x^{\nu} = 0$, while $\alpha_{\mu\nu}\, \mathrm{d}x^{\mu} \wedge \mathrm{d}x^{\nu} = \omega_{\mu\nu}\, \mathrm{d}x^{\mu} \wedge \mathrm{d}x^{\nu}$.

Since there are $\binom{m}{r}$ choices of the set $(\mu_1, \mu_2, \ldots, \mu_r)$ out of $(1, 2, \ldots, m)$ in (5.62), the dimension of the vector space $\Omega^r_p(M)$ is

$$\binom{m}{r} = \frac{m!}{(m-r)!r!}.$$

For later convenience we define $\Omega^0_p(M) = \mathbb{R}$. Clearly $\Omega^1_p(M) = T^*_pM$. If $r$ in (5.62) exceeds $m$, it vanishes identically since some index appears at least twice in the anti-symmetrized summation. The equality $\binom{m}{r} = \binom{m}{m-r}$ implies $\dim \Omega^r_p(M) = \dim \Omega^{m-r}_p(M)$. Since $\Omega^r_p(M)$ is a vector space, $\Omega^r_p(M)$ is isomorphic to $\Omega^{m-r}_p(M)$ (see section 2.2).

Define the **exterior product** of a $q$-form and an $r$-form $\wedge : \Omega^q_p(M) \times \Omega^r_p(M) \to \Omega^{q+r}_p(M)$ by a trivial extension. Let $\omega \in \Omega^q_p(M)$ and $\xi \in \Omega^r_p(M)$, for example. The action of the $(q+r)$-form $\omega \wedge \xi$ on $q+r$ vectors is defined by

$$(\omega \wedge \xi)(V_1, \ldots, V_{q+r}) = \frac{1}{q!r!}\sum_{P \in S_{q+r}} \mathrm{sgn}(P)\omega(V_{P(1)}, \ldots, V_{P(q)})\xi(V_{P(q+1)}, \ldots, V_{P(q+r)}) \tag{5.65}$$

where $V_i \in T_pM$. If $q + r > m$, $\omega \wedge \xi$ vanishes identically. With this product, we define an algebra

$$\Omega^*_p(M) \equiv \Omega^0_p(M) \oplus \Omega^1_p(M) \oplus \ldots \oplus \Omega^m_p(M). \tag{5.66}$$

**Table 5.1.**

| $r$-forms | Basis | Dimension |
|---|---|---|
| $\Omega^0(M) = \mathcal{F}(M)$ | $\{1\}$ | 1 |
| $\Omega^1(M) = T^*M$ | $\{dx^\mu\}$ | $m$ |
| $\Omega^2(M)$ | $\{dx^{\mu_1} \wedge dx^{\mu_2}\}$ | $m(m-1)/2$ |
| $\Omega^3(M)$ | $\{dx^{\mu_1} \wedge dx^{\mu_2} \wedge dx^{\mu_3}\}$ | $m(m-1)(m-2)/6$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $\Omega^m(M)$ | $\{dx^1 \wedge dx^2 \wedge \ldots dx^m\}$ | 1 |

$\Omega_p^*(M)$ is the space of all differential forms at $p$ and is closed under the exterior product.

*Exercise 5.13.* Take the Cartesian coordinates $(x, y)$ in $\mathbb{R}^2$. The two-form $dx \wedge dy$ is the oriented area element (the vector product in elementary vector algebra). Show that, in polar coordinates, this becomes $r dr \wedge d\theta$.

*Exercise 5.14.* Let $\xi \in \Omega_p^q(M)$, $\eta \in \Omega_p^r(M)$ and $\omega \in \Omega_p^s(M)$. Show that

$$\xi \wedge \xi = 0 \qquad \text{if } q \text{ is odd} \tag{5.67a}$$

$$\xi \wedge \eta = (-1)^{qr} \eta \wedge \xi \tag{5.67b}$$

$$(\xi \wedge \eta) \wedge \omega = \xi \wedge (\eta \wedge \omega). \tag{5.67c}$$

We may assign an $r$-form smoothly at each point on a manifold $M$. We denote the space of smooth $r$-forms on $M$ by $\Omega^r(M)$. We also define $\Omega^0(M)$ to be the algebra of smooth functions, $\mathcal{F}(M)$. In summary we have table 5.1.

### 5.4.2 Exterior derivatives

*Definition 5.5.* The exterior derivative $d_r$ is a map $\Omega^r(M) \to \Omega^{r+1}(M)$ whose action on an $r$-form

$$\omega = \frac{1}{r!}\omega_{\mu_1 \ldots \mu_r}\, dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_r}$$

is defined by

$$d_r\omega = \frac{1}{r!}\left(\frac{\partial}{\partial x^\nu}\omega_{\mu_1 \ldots \mu_r}\right) dx^\nu \wedge dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_r}. \tag{5.68}$$

It is common to drop the subscript $r$ and write simply d. The wedge product automatically anti-symmetrizes the coefficient.

*Example 5.10.* The $r$-forms in three-dimensional space are:

(i) $\omega_0 = f(x, y, z)$,
(ii) $\omega_1 = \omega_x(x, y, z)\,\mathrm{d}x + \omega_y(x, y, z)\,\mathrm{d}y + \omega_z(x, y, z)\,\mathrm{d}z$,
(iii) $\omega_2 = \omega_{xy}(x, y, z)\,\mathrm{d}x \wedge \mathrm{d}y + \omega_{yz}(x, y, z)\,\mathrm{d}y \wedge \mathrm{d}z + \omega_{zx}(x, y, z)\,\mathrm{d}z \wedge \mathrm{d}x$ and
(iv) $\omega_3 = \omega_{xyz}(x, y, z)\,\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z$.

If we define an *axial vector* $\alpha^\mu$ by $\varepsilon^{\mu\nu\lambda}\omega_{\nu\lambda}$, a two-form may be regarded as a 'vector'. The **Levi-Civita symbol** $\varepsilon^{\mu\nu\lambda}$ is defined by $\varepsilon^{P(1)P(2)P(3)} = \mathrm{sgn}(P)$ and provides the isomorphism between $\mathcal{X}(M)$ and $\Omega^2(M)$. [Note that both of these are of dimension three.]

The action of d is

(i) $\mathrm{d}\omega_0 = \dfrac{\partial f}{\partial x}\,\mathrm{d}x + \dfrac{\partial f}{\partial y}\,\mathrm{d}y + \dfrac{\partial f}{\partial z}\,\mathrm{d}z$,
(ii) $\mathrm{d}\omega_1 = \left(\dfrac{\partial \omega_y}{\partial x} - \dfrac{\partial \omega_x}{\partial y}\right)\mathrm{d}x \wedge \mathrm{d}y + \left(\dfrac{\partial \omega_z}{\partial y} - \dfrac{\partial \omega_y}{\partial z}\right)\mathrm{d}y \wedge \mathrm{d}z + \left(\dfrac{\partial \omega_x}{\partial z} - \dfrac{\partial \omega_z}{\partial x}\right)\mathrm{d}z \wedge \mathrm{d}x$,
(iii) $\mathrm{d}\omega_2 = \left(\dfrac{\partial \omega_{yz}}{\partial x} + \dfrac{\partial \omega_{zx}}{\partial y} + \dfrac{\partial \omega_{xy}}{\partial z}\right)\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z$ and
(iv) $\mathrm{d}\omega_3 = 0$.

Hence, the action of d on $\omega_0$ is identified with 'grad', on $\omega_1$ with 'rot' and on $\omega_2$ with 'div' in the usual vector calculus.

*Exercise 5.15.* Let $\xi \in \Omega^q(M)$ and $\omega \in \Omega^r(M)$. Show that

$$\mathrm{d}(\xi \wedge \omega) = \mathrm{d}\xi \wedge \omega + (-1)^q \xi \wedge \mathrm{d}\omega. \tag{5.69}$$

A useful expression for the exterior derivative is obtained as follows. Let us take $X = X^\mu \partial/\partial x^\mu$, $Y = Y^\nu \partial/\partial x^\nu \in \mathcal{X}(M)$ and $\omega = \omega_\mu\,\mathrm{d}x^\mu \in \Omega^1(M)$. It is easy to see that the combination

$$X[\omega(Y)] - Y[\omega(X)] - \omega([X, Y]) = \frac{\partial \omega_\mu}{\partial x^\nu}(X^\nu Y^\mu - X^\mu Y^\nu)$$

is equal to $\mathrm{d}\omega(X, Y)$, and we have the coordinate-free expression

$$\mathrm{d}\omega\,(X, Y) = X[\omega(Y)] - Y[\omega(X)] - \omega([X, Y]). \tag{5.70}$$

For an $r$-form $\omega \in \Omega^r(M)$, this becomes

$$\begin{aligned} &\mathrm{d}\omega\,(X_1, \ldots, X_{r+1}) \\ &\quad = \sum_{i=1}^{r}(-1)^{i+1} X_i \omega(X_1, \ldots, \hat{X}_i, \ldots, X_{r+1}) \\ &\qquad + \sum_{i<j}(-1)^{i+j}\omega([X_i, X_j], X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_{r+1}) \end{aligned} \tag{5.71}$$

where the entry below $\hat{}$ has been omitted. As an exercise, the reader should verify (5.71) explicitly for $r = 2$.

We now prove an important formula:

$$\mathrm{d}^2 = 0 \qquad (\text{or } \mathrm{d}_{r+1}\,\mathrm{d}_r = 0). \tag{5.72}$$

Take

$$\omega = \frac{1}{r!}\omega_{\mu_1\ldots\mu_r}\,\mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r} \in \Omega^r(M).$$

The action of $\mathrm{d}^2$ on $\omega$ is

$$\mathrm{d}^2\omega = \frac{1}{r!}\frac{\partial^2\omega_{\mu_1\ldots\mu_r}}{\partial x^\lambda \partial x^\nu}\,\mathrm{d}x^\lambda \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r}.$$

This vanishes identically since $\partial^2\omega_{\mu_1\ldots\mu_r}/\partial x^\lambda \partial x^\nu$ is symmetric with respect to $\lambda$ and $\nu$ while $\mathrm{d}x^\lambda \wedge \mathrm{d}x^\nu$ is anti-symmetric.

*Example 5.11.* It is known that the electromagnetic potential $A = (\phi, A)$ is a one-form, $A = A_\mu \mathrm{d}x^\mu$ (see chapter 10). The electromagnetic tensor is defined by $F = \mathrm{d}A$ and has the components

$$\begin{pmatrix} 0 & -E_x & -E_y & -E_x \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix} \tag{5.73}$$

where

$$\mathbf{E} = -\nabla\phi - \frac{\partial}{\partial x^0}\mathbf{A} \qquad \text{and} \qquad \mathbf{B} = \nabla \times \mathbf{A}$$

as usual. Two Maxwell equations, $\nabla \cdot \mathbf{B} = 0$ and $\partial\mathbf{B}/\partial t = -\nabla \times \mathbf{E}$ follow from the identity $\mathrm{d}F = \mathrm{d}(\mathrm{d}A) = 0$, which is known as the **Bianchi identity**, while the other set is the equation of motion derived from the Lagrangian (1.245).

A map $f : M \to N$ induces the pullback $f^* : T^*_{f(p)}N \to T^*_pM$ and $f^*$ is naturally extended to tensors of type $(0, r)$; see section 5.2. Since an $r$-form is a tensor of type $(0, r)$, this applies as well. Let $\omega \in \Omega^r(N)$ and let $f$ be a map $M \to N$. At each point $f(p) \in N$, $f$ induces the pullback $f^* : \Omega^r_{f(p)}N \to \Omega^r_pM$ by

$$(f^*\omega)(X_1, \ldots, X_r) \equiv \omega(f_*X_1, \ldots, f_*X_r) \tag{5.74}$$

where $X_i \in T_pM$ and $f_*$ is the differential map $T_pM \to T_{f(p)}N$.

*Exercise 5.16.* Let $\xi, \omega \in \Omega^r(N)$ and let $f : M \to N$. Show that

$$\mathrm{d}\,(f^*\omega) = f^*(d\omega) \tag{5.75}$$

$$f^*(\xi \wedge \omega) = (f^*\xi) \wedge (f^*\omega). \tag{5.76}$$

The exterior derivative $\mathrm{d}_r$ induces the sequence

$$0 \xrightarrow{i} \Omega^0(M) \xrightarrow{\mathrm{d}_0} \Omega^1(M) \xrightarrow{\mathrm{d}_1} \cdots \xrightarrow{d_{m-2}} \Omega^{m-1}(M) \xrightarrow{\mathrm{d}_{m-1}} \Omega^m(M) \xrightarrow{\mathrm{d}_m} 0 \quad (5.77)$$

where $i$ is the inclusion map $0 \hookrightarrow \Omega^0(M)$. This sequence is called the **de Rham complex**. Since $\mathrm{d}^2 = 0$, we have $\mathrm{im}\,\mathrm{d}_r \subset \ker \mathrm{d}_{r+1}$. [Take $\omega \in \Omega^r(M)$. Then $\mathrm{d}_r\omega \in \mathrm{im}\,\mathrm{d}_r$ and $\mathrm{d}_{r+1}(\mathrm{d}_r\omega) = 0$ imply $\mathrm{d}_r\omega \in \ker \mathrm{d}_{r+1}$.] An element of $\ker \mathrm{d}_r$ is called a **closed *r*-form**, while an element of $\mathrm{im}\,\mathrm{d}_{r-1}$ is called an **exact *r*-form**. Namely, $\omega \in \Omega^r(M)$ is closed if $\mathrm{d}\omega = 0$ and exact if there exists an $(r-1)$-form $\psi$ such that $\omega = \mathrm{d}\psi$. The quotient space $\ker \mathrm{d}_r / \mathrm{im}\, d_{r-1}$ is called the $r$th **de Rham cohomology group** which is made into the dual space of the homology group; see chapter 6.

### 5.4.3 Interior product and Lie derivative of forms

Another important operation is the **interior product** $\mathrm{i}_X : \Omega^r(M) \to \Omega^{r-1}(M)$, where $X \in \mathcal{X}(M)$. For $\omega \in \Omega^r(M)$, we define

$$\mathrm{i}_X\omega(X_1, \ldots, X_{r-1}) \equiv \omega(X, X_1, \ldots, X_{r-1}). \quad (5.78)$$

For $X = X^\mu \partial/\partial x^\mu$ and $\omega = (1/r!)\omega_{\mu_1\ldots\mu_r}\, \mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r}$ we have

$$\begin{aligned}
\mathrm{i}_X\omega &= \frac{1}{(r-1)!} X^\nu \omega_{\nu\mu_2\ldots\mu_r}\, \mathrm{d}x^{\mu_2} \wedge \ldots \wedge \mathrm{d}x^{\mu_r} \\
&= \frac{1}{r!}\sum_{s=1}^{r} X^{\mu_s}\omega_{\mu_1\ldots\mu_s\ldots\mu_r}(-1)^{s-1}\, \mathrm{d}x^{\mu_1} \wedge \ldots \wedge \widehat{\mathrm{d}x^{\mu_s}} \wedge \ldots \wedge \mathrm{d}x^{\mu_r}
\end{aligned} \quad (5.79)$$

where the entry below $\hat{}$ has been omitted. For example, let $(x, y, z)$ be the coordinates of $\mathbb{R}^3$. Then

$$\mathrm{i}_{e_x}(\mathrm{d}x \wedge \mathrm{d}y) = \mathrm{d}y, \qquad \mathrm{i}_{e_x}(\mathrm{d}y \wedge \mathrm{d}z) = 0, \qquad \mathrm{i}_{e_x}(\mathrm{d}z \wedge \mathrm{d}x) = -\mathrm{d}z.$$

The Lie derivative of a form is most neatly written with the interior product. Let $\omega = \omega_\mu \mathrm{d}x^\mu$ be a one-form. Consider the combination

$$\begin{aligned}
(\mathrm{d}\mathrm{i}_X + \mathrm{i}_X\mathrm{d})\omega &= \mathrm{d}\,(X^\mu\omega_\mu) + \mathrm{i}_X[\tfrac{1}{2}(\partial_\mu\omega_\nu - \partial_\nu\omega_\mu)\, \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu] \\
&= (\omega_\mu \partial_\nu X^\mu + X^\mu \partial_\nu \omega_\mu)\, \mathrm{d}x^\nu + X^\mu(\partial_\mu\omega_\nu - \partial_\nu\omega_\mu)\, \mathrm{d}x^\nu \\
&= (\omega_\mu \partial_\nu X^\mu + X^\mu \partial_\mu \omega_\nu)\, \mathrm{d}x^\nu.
\end{aligned}$$

Comparing this with (5.55), we find that

$$\mathcal{L}_X\omega = (\mathrm{d}\mathrm{i}_X + \mathrm{i}_X\mathrm{d})\omega. \quad (5.80)$$

For a general $r$-form $\omega = (1/r!)\omega_{\mu_1\ldots\mu_r}\,\mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_r}$, we have

$$\begin{aligned}\mathcal{L}_X\omega &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}((\sigma_\varepsilon)^*\omega|_{\sigma_\varepsilon(x)} - \omega|_x)\\ &= X^\nu \frac{1}{r!}\partial_\nu\omega_{\mu_1\ldots\mu_r}\mathrm{d}x^{\mu_1}\wedge\ldots\wedge\mathrm{d}x^{\mu_r}\\ &\quad + \sum_{s=1}^{r}\partial_{\mu_s}X^\nu\frac{1}{r!}\omega_{\mu_1\ldots\overset{s}{\overset{\downarrow}{\nu}}\ldots\mu_r}\mathrm{d}x^{\mu_1}\wedge\ldots\wedge\mathrm{d}x^{\mu_r}. \end{aligned} \tag{5.81}$$

We also have

$$\begin{aligned}&(\mathrm{di}_X + \mathrm{i}_X\mathrm{d})\omega\\ &\quad= \frac{1}{r!}\sum_{s=1}^{r}[\partial_\nu X^{\mu_s}\omega_{\mu_1\ldots\mu_s\ldots\mu_r} + X^{\mu_s}\partial_\nu\omega_{\mu_1\ldots\mu_s\ldots\mu_r}]\\ &\qquad\times(-1)^{s-1}\,\mathrm{d}x^\nu\wedge\mathrm{d}x^{\mu_1}\wedge\ldots\wedge\widehat{\mathrm{d}x^{\mu_s}}\wedge\mathrm{d}x^{\mu_r}\\ &\qquad+\frac{1}{r!}[X^\nu\partial_\nu\omega_{\mu_1\ldots\mu_r}\mathrm{d}x^{\mu_1}\wedge\ldots\wedge\,\mathrm{d}x^{\mu_r}\\ &\qquad+\sum_{s=1}^{r}X^{\mu_s}\omega_{\mu_1\ldots\mu_s\ldots\mu_r}(-1)^s\,\mathrm{d}x^\nu\wedge\mathrm{d}x^{\mu_1}\wedge\ldots\wedge\widehat{\mathrm{d}x^{\mu_s}}\wedge\ldots\wedge\mathrm{d}x^{\mu_r}]\\ &\quad=\frac{1}{r!}\sum_{s=1}^{r}[\partial_\nu X^{\mu_s}\omega_{\mu_1\ldots\mu_s\ldots\mu_r}(-1)^{s-1}\,\mathrm{d}x^\nu\wedge\mathrm{d}x^{\mu_1}\wedge\ldots\wedge\widehat{\mathrm{d}x^{\mu_s}}\wedge\ldots\wedge\mathrm{d}x^{\mu_r}\\ &\qquad+\frac{1}{r!}X^\nu\partial_\nu\omega_{\mu_1\ldots\mu_r}\,\mathrm{d}x^{\mu_1}\wedge\ldots\wedge\mathrm{d}x^{\mu_r}.\end{aligned}$$

If we interchange the roles of $\mu_s$ and $\nu$ in the first term of the last expression and compare it with (5.81), we verify that

$$(\mathrm{di}_X + \mathrm{i}_X\,\mathrm{d})\omega = \mathcal{L}_X\omega \tag{5.82}$$

for any $r$-form $\omega$.

*Exercise 5.17.* Let $X, Y \in \mathcal{X}(M)$ and $\omega \in \Omega^r(M)$. Show that

$$\mathrm{i}_{[X,Y]}\omega = X(\mathrm{i}_Y\omega) - Y(\mathrm{i}_X\omega). \tag{5.83}$$

Show also that $\mathrm{i}_X$ is an anti-derivation,

$$\mathrm{i}_X(\omega\wedge\eta) = \mathrm{i}_X\omega\wedge\eta + (-1)^r\omega\wedge\mathrm{i}_X\eta \tag{5.84}$$

and nilpotent,

$$\mathrm{i}_X^2 = 0. \tag{5.85}$$

Use the nilpotency to prove

$$\mathcal{L}_X\mathrm{i}_X\omega = \mathrm{i}_X\mathcal{L}_X\omega. \tag{5.86}$$

*Exercise 5.18.* Let $t \in \mathcal{T}^n_m(M)$. Show that

$$(\mathcal{L}_X t)^{\mu_1\ldots\mu_n}_{\nu_1\ldots\nu_m} = X^\lambda \partial_\lambda t^{\mu_1\ldots\mu_n}_{\nu_1\ldots\nu_m} + \sum_{s=1}^{n} \partial_{\nu_s} X^\lambda t^{\mu_1\ldots\mu_n}_{\nu_1\ldots\lambda\ldots\nu_m} - \sum_{s=1}^{n} \partial_\lambda X^{\mu_s} t^{\mu_1\ldots\lambda\ldots\mu_n}_{\nu_1\ldots\nu_m}. \quad (5.87)$$

*Example 5.12.* Let us reformulate Hamiltonian mechanics (section 1.1) in terms of differential forms. Let $H$ be a Hamiltonian and $(q^\mu, p_\mu)$ be its phase space. Define a two-form

$$\omega = \mathrm{d}p_\mu \wedge \mathrm{d}q^\mu \quad (5.88)$$

called the **symplectic two-form**. If we introduce a one-form

$$\theta = q^\mu \, \mathrm{d}p_\mu, \quad (5.89)$$

the symplectic two-form is expressed as

$$\omega = \mathrm{d}\theta. \quad (5.90)$$

Given a function $f(q, p)$ in the phase space, one can define the **Hamiltonian vector field**

$$X_f = \frac{\partial f}{\partial p_\mu}\frac{\partial}{\partial q^\mu} - \frac{\partial f}{\partial q^\mu}\frac{\partial}{\partial p_\mu}. \quad (5.91)$$

Then it is easy to verify that

$$\mathrm{i}_{X_f}\omega = -\frac{\partial f}{\partial p_\mu}\,\mathrm{d}p^\mu - \frac{\partial f}{\partial q^\mu}\,\mathrm{d}q^\mu = -\mathrm{d}f.$$

Consider a vector field generated by the Hamiltonian

$$X_H = \frac{\partial H}{\partial p_\mu}\frac{\partial}{\partial q^\mu} - \frac{\partial H}{\partial q^\mu}\frac{\partial}{\partial p_\mu}. \quad (5.92)$$

For the solution $(q^\mu, p_\mu)$ to Hamilton's equation of motion

$$\frac{\mathrm{d}q^\mu}{\mathrm{d}t} = \frac{\partial H}{\partial p_\mu} \qquad \frac{\mathrm{d}p_\mu}{\mathrm{d}t} = -\frac{\partial H}{\partial q^\mu}, \quad (5.93)$$

we also obtain

$$X_H = \frac{\mathrm{d}p_\mu}{\mathrm{d}t}\frac{\partial}{\partial p_\mu}\frac{\mathrm{d}q^\mu}{\mathrm{d}t}\frac{\partial}{\partial q^\mu} = \frac{\mathrm{d}}{\mathrm{d}t}. \quad (5.94)$$

The symplectic two-form $\omega$ is left invariant along the flow generated by $X_H$,

$$\begin{aligned}\mathcal{L}_{X_H}\omega &= \mathrm{d}(\mathrm{i}_{X_H}\omega) + \mathrm{i}_{X_H}(\mathrm{d}\omega) \\ &= \mathrm{d}(\mathrm{i}_{X_H}\omega) = -\mathrm{d}^2 H = 0\end{aligned} \quad (5.95)$$

where use has been made of (5.82). Conversely, if $X$ satisifes $\mathcal{L}_X\omega = 0$, there exists a Hamiltonian $H$ such that Hamilton's equation of motion is satisfied along the flow generated by $X$. This follows from the previous observation that $\mathcal{L}_X\omega = \mathrm{d}(\mathrm{i}_X\omega) = 0$ and hence by Poincaré's lemma, there exists a function $H(q, p)$ such that

$$\mathrm{i}_X\omega = -\mathrm{d}H.$$

The Poisson bracket is cast into a form independent of the special coordinates chosen with the help of the Hamiltonian vector fields. In fact,

$$\mathrm{i}_{X_f}(\mathrm{i}_{X_g}\omega) = -\mathrm{i}_{X_f}(\mathrm{d}g) = \frac{\partial f}{\partial q^\mu}\frac{\partial g}{\partial p_\mu} - \frac{\partial f}{\partial q^\mu}\frac{\partial g}{\partial p_\mu} = [f, g]_{\mathrm{PB}}. \tag{5.96}$$

## 5.5 Integration of differential forms

### 5.5.1 Orientation

An integration of a differential form over a manifold $M$ is defined only when $M$ is 'orientable'. So we first define an **orientation** of a manifold. Let $M$ be a connected $m$-dimensional differentiable manifold. At a point $p \in M$, the tangent space $T_pM$ is spanned by the basis $\{e_\mu\} = \{\partial/\partial x^\mu\}$, where $x^\mu$ is the local coordinate on the chart $U_i$ to which $p$ belongs. Let $U_j$ be another chart such that $U_i \cap U_j \neq \emptyset$ with the local coordinates $y^\alpha$. If $p \in U_i \cap U_j$, $T_pM$ is spanned by either $\{e_\mu\}$ or $\{\widetilde{e}_\alpha\} = \{\partial/\partial y^\alpha\}$. The basis changes as

$$\widetilde{e}_\alpha = \left(\frac{\partial x^\mu}{\partial y^\alpha}\right) e_\mu. \tag{5.97}$$

If $J = \det(\partial x^\mu/\partial y^\alpha) > 0$ on $U_i \cap U_j$, $\{e_\mu\}$ and $\{\widetilde{e}_\alpha\}$ are said to define the *same orientation* on $U_i \cap U_j$ and if $J < 0$, they define the *opposite orientation*.

*Definition 5.6.* Let $M$ be a connected manifold covered by $\{U_i\}$. The manifold $M$ is **orientable** if, for any overlapping charts $U_i$ and $U_j$, there exist local coordinates $\{x^\mu\}$ for $U_i$ and $\{y^\alpha\}$ for $U_j$ such that $J = \det(\partial x^\mu/\partial y^\alpha) > 0$.

If $M$ is non-orientable, $J$ cannot be positive in all intersections of charts. For example, the Möbius strip in figure 5.14($a$) is non-orientable since we have to choose $J$ to be negative in the intersection B.

If an $m$-dimensional manifold $M$ is orientable, there exists an $m$-form $\omega$ which vanishes nowhere. This $m$-form $\omega$ is called a **volume element**, which plays the role of a measure when we integrate a function $f \in \mathcal{F}(M)$ over $M$. Two volume elements $\omega$ and $\omega'$ are said to be *equivalent* if there exists a strictly positive function $h \in \mathcal{F}(M)$ such that $\omega = h\omega'$. A negative-definite function $h' \in \mathcal{F}(M)$ gives an inequivalent orientation to $M$. Thus, any orientable manifold admits *two* inequivalent orientations, one of which is called **right handed**, the other **left handed**. Take an $m$-form

$$\omega = h(p)\,\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^m \tag{5.98}$$

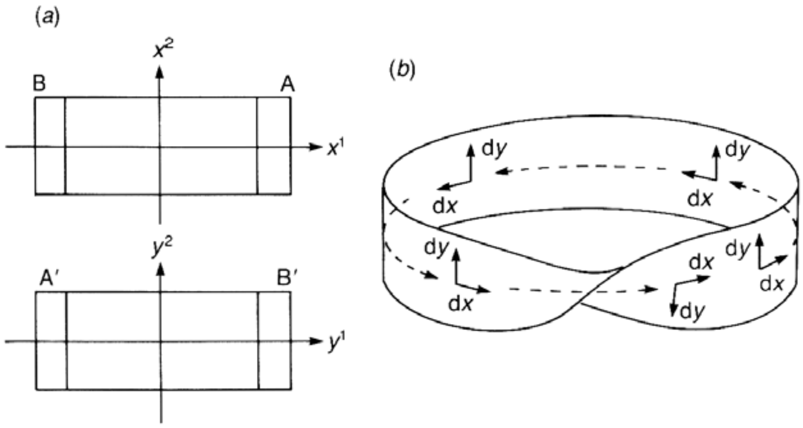


**Figure 5.14.** (*a*) The Möbius strip is obtained by twisting the part B′ of the second strip by $\pi$ before pasting A with A′ and B with B′. The coordinate change on B is $y^1 = x^1, y^2 = -x^2$ and the Jacobian is $-1$. (*b*) Basis frames on the Möbius strip.

with a positive-definite $h(p)$ on a chart $(U, \varphi)$ whose coordinate is $x = \varphi(p)$. If $M$ is orientable, we may extend $\omega$ throughout $M$ such that the component $h$ is positive definite on any chart $U_i$. If $M$ is orientable, this $\omega$ is a volume element. Note that this positivity of $h$ is independent of the choice of coordinates. In fact, let $p \in U_i \cap U_j \neq \emptyset$ and let $x^\mu$ and $y^\alpha$ be the coordinates of $U_i$ and $U_j$, respectively. Then (5.98) becomes

$$\omega = h(p)\frac{\partial x^1}{\partial y^{\mu_1}}\,\mathrm{d}y^{\mu_1} \wedge \ldots \wedge \frac{\partial x^m}{\partial y^{\mu_m}}\,\mathrm{d}y^{\mu_m} = h(p)\det\left(\frac{\partial x^\mu}{\partial y^\nu}\right)\mathrm{d}y^1 \wedge \ldots \wedge \mathrm{d}y^m. \tag{5.99}$$

The determinant in (5.99) is the Jacobian of the coordinate transformation and must be positive by assumed orientability. If $M$ is non-orientable, $\omega$ with a positive-definite component cannot be defined on $M$. Let us look at figure 5.14 again. If we circumnavigate the strip along the direction shown in the figure, $\omega = \mathrm{d}x \wedge \mathrm{d}y$ changes the signature $\mathrm{d}x \wedge \mathrm{d}y \to -\mathrm{d}x \wedge \mathrm{d}y$ when we come back to the starting point. Hence, $\omega$ cannot be defined uniquely on $M$.

### 5.5.2 Integration of forms

Now we are ready to define an integration of a function $f : M \to \mathbb{R}$ over an orientable manifold $M$. Take a volume element $\omega$. In a coordinate neighbourhood $U_i$ with the coordinate $x$, we define the integration of an $m$-form $f\omega$ by

$$\int_{U_i} f\omega \equiv \int_{\varphi(U_i)} f(\varphi_i^{-1}(x))h(\varphi_i^{-1}(x))\,\mathrm{d}x^1 \ldots \mathrm{d}x^m. \tag{5.100}$$

The RHS is an ordinary multiple integration of a function of $m$ variables. Once the integral of $f$ over $U_i$ is defined, the integral of $f$ over the whole of $M$ is given with the help of the 'partition of unity' defined now.

*Definition 5.7.* Take an open covering $\{U_i\}$ of $M$ such that each point of $M$ is covered with a finite number of $U_i$. [If this is always possible, $M$ is called **paracompact**, which we assume to be the case.] If a family of differentiable functions $\varepsilon_i(p)$ satisfies

(i) $0 \le \varepsilon_i(p) \le 1$
(ii) $\varepsilon_i(p) = 0$ if $p \notin U_i$ and
(iii) $\varepsilon_1(p) + \varepsilon_2(p) + \ldots = 1$ for any point $p \in M$

the family $\{\varepsilon(p)\}$ is called a **partition of unity** subordinate to the covering $\{U_i\}$.

From condition (iii), it follows that

$$f(p) = \sum_i f(p)\varepsilon_i(p) = \sum_i f_i(p) \tag{5.101}$$

where $f_i(p) \equiv f(p)\varepsilon_i(p)$ vanishes outside $U_i$ by (ii). Hence, given a point $p \in M$, assumed paracompactness ensures that there are only finite terms in the summation over $i$ in (5.101). For each $f_i(p)$, we may define the integral over $U_i$ according to (5.100). Finally the integral of $f$ on $M$ is given by

$$\int_M f\omega \equiv \sum_i \int_{U_i} f_i\omega. \tag{5.102}$$

Although a different atlas $\{(V_i, \psi_i)\}$ gives different coordinates and a different partition of unity, the integral defined by (5.102) remains the same.

*Example 5.13.* Let us take the atlas of $S^1$ defined in example 5.2. Let $U_1 = S^1 - \{(1,0)\}$, $U_2 = S^1 - \{(-1,0)\}$, $\varepsilon_1(\theta) = \sin^2(\theta/2)$ and $\varepsilon_2(\theta) = \cos^2(\theta/2)$. The reader should verify that $\{\varepsilon_i(\theta)\}$ is a partition of unity subordinate to $\{U_i\}$. Let us integrate a function $f = \cos^2\theta$, for example. [Of course we know

$$\int_0^{2\pi} \mathrm{d}\theta\ \cos^2\theta = \pi$$

but let us use the partition of unity.] We have

$$\begin{aligned}\int_{S^1} \mathrm{d}\theta\ \cos^2\theta &= \int_0^{2\pi} \mathrm{d}\theta\ \ \sin^2\frac{\theta}{2}\cos^2\theta + \int_{-\pi}^{\pi} \mathrm{d}\theta\ \ \cos^2\frac{\theta}{2}\cos^2\theta \\ &= \tfrac{1}{2}\pi + \tfrac{1}{2}\pi = \pi.\end{aligned}$$

So far, we have left $h$ arbitrary provided it is strictly positive. The reader might be tempted to choose $h$ to he unity. However, as we found in (5.99), $h$ is multiplied by the Jacobian under the change of coordinates and there is no canonical way to single out the component $h$; unity in one coordinate might not be unity in the other. The situation changes if the manifold is endowed with a metric, as we will see in chapter 7.

## 5.6 Lie groups and Lie algebras

A Lie group is a manifold on which the group manipulations, *product* and *inverse*, are defined. Lie groups play an extremely important role in the theory of fibre bundles and also find vast applications in physics. Here we will work out the geometrical aspects of Lie groups and Lie algebras.

### 5.6.1 Lie groups

*Definition 5.8.* A Lie group $G$ is a differentiable manifold which is endowed with a group structure such that the group operations

(i) $\cdot : G \times G \to G,\ (g_1, g_2) \mapsto g_1 \cdot g_2$
(ii) $^{-1} : G \to G,\ g \mapsto g^{-1}$

are differentiable. [*Remark:* It can be shown that $G$ has a unique analytic structure with which the product and the inverse operations are written as convergent power series.]

The unit element of a Lie group is written as $e$. The dimension of a Lie group $G$ is defined to be the dimension of $G$ as a manifold. The product symbol may be omitted and $g_1 \cdot g_2$ is usually written as $g_1 g_2$. For example, let $\mathbb{R}^* \equiv \mathbb{R} - \{0\}$. Take three elements $x, y, z \in \mathbb{R}^*$ such that $xy = z$. Obviously if we multiply a number close to $x$ by a number close to $y$, we have a number close to $z$. Similarly, an inverse of a number close to $x$ is close to $1/x$. In fact, we can differentiate these maps with respect to the relevant arguments and $\mathbb{R}^*$ is made into a Lie group with these group operations. If the product is commutative, namely $g_1 g_2 = g_2 g_1$, we often use the additive symbol $+$ instead of the product symbol.

*Exercise 5.19.*

(a) Show that $\mathbb{R}^+ = \{x \in \mathbb{R} | x > 0\}$ is a Lie group with respect to multiplication.
(b) Show that $\mathbb{R}$ is a Lie group with respect to addition.
(c) Show that $\mathbb{R}^2$ is a Lie group with respect to addition defined by $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$.

*Example 5.14.* Let $S^1$ be the unit circle on the complex plane,

$$S^1 = \{e^{i\theta} | \theta \in \mathbb{R} \quad (\text{mod } 2\pi)\}.$$

The group operations defined by $e^{i\theta} e^{i\varphi} = e^{i(\theta+\varphi)}$ and $(e^{i\theta})^{-1} = e^{-i\theta}$ are differentiable and $S^1$ is made into a Lie group, which we call U(1). It is easy to see that the group operations are the same as those in exercise 5.19(*b*) modulo $2\pi$.

Of particular interest in physical applications are the matrix groups which are subgroups of general linear groups $\mathrm{GL}(n, \mathbb{R})$ or $\mathrm{GL}(n, \mathbb{C})$. The product of elements is simply the matrix multiplication and the inverse is given by the matrix inverse. The coordinates of $\mathrm{GL}(n, \mathbb{R})$ are given by $n^2$ entries of $M = \{x_{ij}\}$. $\mathrm{GL}(n, \mathbb{R})$ is a non-compact manifold of real dimension $n^2$.

Interesting subgroups of $\mathrm{GL}(n, \mathbb{R})$ are the **orthogonal group** $\mathrm{O}(n)$, the **special linear group** $\mathrm{SL}(n, \mathbb{R})$ and the **special orthogonal group** $\mathrm{SO}(n)$:

$$\mathrm{O}(n) = \{M \in \mathrm{GL}(n, \mathbb{R}) | M M^{\mathrm{t}} = M^{\mathrm{t}} M = I_n\} \tag{5.103}$$

$$\mathrm{SL}(n, \mathbb{R}) = \{M \in \mathrm{GL}(n, \mathbb{R}) | \det M = 1\} \tag{5.104}$$

$$\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}(n, \mathbb{R}) \tag{5.105}$$

where $^{\mathrm{t}}$ denotes the transpose of a matrix. In special relativity, we are familiar with the **Lorentz group**

$$\mathrm{O}(1, 3) = \{M \in \mathrm{GL}(4, \mathbb{R}) | M \eta M^t = \eta\}$$

where $\eta$ is the Minkowski metric, $\eta = \mathrm{diag}(-1, 1, 1, 1)$. Extension to higher-dimensional spacetime is trivial.

*Exercise 5.20.* Show that the group $\mathrm{O}(1, 3)$ is non-compact and has four connected components according to the sign of the determinant and the sign of the $(0, 0)$ entry. The component that contains the unit matrix is denoted by $\mathrm{O}_+^{\uparrow}(1, 3)$.

The group $\mathrm{GL}(n, \mathbb{C})$ is the set of non-singular linear transformations in $\mathbb{C}^n$, which are represented by $n \times n$ non-singular matrices with complex entries. The **unitary group** $\mathrm{U}(n)$, the **special linear group** $\mathrm{SL}(n, \mathbb{C})$ and the **special unitary group** $\mathrm{SU}(n)$ are defined by

$$\mathrm{U}(n) = \{M \in \mathrm{GL}(n, \mathbb{C}) | M M^\dagger = M^\dagger M = \mathbf{1}\} \tag{5.106}$$

$$\mathrm{SL}(n, \mathbb{C}) = \{M \in \mathrm{GL}(n, \mathbb{C}) | \det M = 1\} \tag{5.107}$$

$$\mathrm{SU}(n) = \mathrm{U}(n) \cap \mathrm{SL}(n, \mathbb{C}) \tag{5.108}$$

where $^\dagger$ is the Hermitian conjugate.

So far we have just mentioned that the matrix groups are subgroups of a Lie group $\mathrm{GL}(n, \mathbb{R})$ (or $\mathrm{GL}(n, \mathbb{C})$). The following theorem guarantees that they are Lie subgroups, that is, these subgroups are Lie groups by themselves. We accept this important (and difficult to prove) theorem without proof.

*Theorem 5.2.* Every closed subgroup $H$ of a Lie group $G$ is a Lie subgroup.

For example, $\mathrm{O}(n)$, $\mathrm{SL}(n, \mathbb{R})$ and $\mathrm{SO}(n)$ are Lie subgroups of $\mathrm{GL}(n, \mathbb{R})$. To see why $\mathrm{SL}(n, \mathbb{R})$ is a closed subgroup, consider a map $f : \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}$ defined by $A \mapsto \det A$. Obviously $f$ is a continuous map and $f^{-1}(1) = \mathrm{SL}(n, \mathbb{R})$. A point $\{1\}$ is a closed subset of $\mathbb{R}$, hence $f^{-1}(1)$ is closed in $\mathrm{GL}(n, \mathbb{R})$. Then theorem 5.2 states that $\mathrm{SL}(n, \mathbb{R})$ is a Lie subgroup. The reader should verify that $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ are also Lie subgroups of $\mathrm{GL}(n, \mathbb{R})$.

Let $G$ be a Lie group and $H$ a Lie subgroup of $G$. Define an equivalence relation $\sim$ by $g \sim g'$ if there exists an element $h \in H$ such that $g' = gh$. An equivalence class $[g]$ is a set $\{gh | h \in H\}$. The coset space $G/H$ is a manifold (not necessarily a Lie group) with $\dim G/H = \dim G - \dim H$. $G/H$ is a Lie group if $H$ is a normal subgroup of $G$, that is, if $ghg^{-1} \in H$ for any $g \in G$ and $h \in H$. In fact, take equivalence classes $[g], [g'] \in G/H$ and construct the product $[g][g']$. If the group structure is well defined in $G/H$, the product must be independent of the choice of the representatives. Let $gh$ and $g'h'$ be the representatives of $[g]$ and $[g']$ respectively. Then $ghg'h' = gg'h''h' \in [gg']$ where the equality follows since there exists $h'' \in H$ such that $hg' = g'h''$. It is left as an exercise to the reader to show that $[g]^{-1}$ is also a well defined operation and $[g]^{-1} = [g^{-1}]$.

### 5.6.2 Lie algebras

*Definition 5.9.* Let $a$ and $g$ be elements of a Lie group $G$. The **right-translation** $R_a : G \to G$ and the **left-translation** $L_a : G \to G$ of $g$ by $a$ are defined by

$$R_a g = ga \tag{5.109a}$$

$$L_a g = ag. \tag{5.109b}$$

By definition, $R_a$ and $L_a$ are diffeomorphisms from $G$ to $G$. Hence, the maps $L_a : G \to G$ and $R_a : G \to G$ induce $L_{a*} : T_g G \to T_{ag} G$ and $R_{a*} : T_g G \to T_{ga} G$; see section 5.2. Since these translations give equivalent theories, we are concerned mainly with the left-translation in the following. The analysis based on the right-translation can be carried out in a similar manner.

Given a Lie group $G$, there exists a special class of vector fields characterized by an invariance under group action. [On the usual manifold there is no canonical way of discriminating some vector fields from the others.]

*Definition 5.10.* Let $X$ be a vector field on a Lie group $G$. $X$ is said to be a **left-invariant vector field** if $L_{a*} X|_g = X|_{ag}$.

*Exercise 5.21.* Verify that a left-invariant vector field $X$ satisfies

$$L_{a*} X|_g = X^\mu(g) \frac{\partial x^\nu(ag)}{\partial x^\mu(g)} \left.\frac{\partial}{\partial x^\nu}\right|_{ag} = X^\nu(ag) \left.\frac{\partial}{\partial x^\nu}\right|_{ag} \tag{5.110}$$

where $x^\mu(g)$ and $x^\mu(ag)$ are coordinates of $g$ and $ag$, respectively.

A vector $V \in T_e G$ defines a unique left-invariant vector field $X_V$ throughout $G$ by

$$X_V|_g = L_{g*} V \qquad g \in G. \tag{5.111}$$

In fact, we verify from (5.34) that $X_V|_{ag} = L_{ag*} V = (L_a L_g)_* V = L_{a*} L_{g*} V = L_{a*} X_V|_g$. Conversely, a left-invariant vector field $X$ defines a unique vector $V = X|_e \in T_e G$. Let us denote the set of left-invariant vector fields on $G$ by

$\mathfrak{g}$. The map $T_eG \to \mathfrak{g}$ defined by $V \mapsto X_V$ is an isomorphism and it follows that the set of left-invariant vector fields is a vector space isomorphic to $T_eG$. In particular, $\dim \mathfrak{g} = \dim G$.

Since $\mathfrak{g}$ is a set of vector fields, it is a subset of $\mathfrak{X}(G)$ and the Lie bracket defined in section 5.3 is also defined on $\mathfrak{g}$. We show that $\mathfrak{g}$ is closed under the Lie bracket. Take two points $g$ and $ag = L_a g$ in $G$. If we apply $L_{a*}$ to the Lie bracket $[X, Y]$ of $X, Y \in \mathfrak{g}$, we have

$$L_{a*}[X,Y]|_g = [L_{a*}X|_g, L_{a*}Y|_g] = [X,Y]|_{ag} \tag{5.112}$$

where the left-invariances of $X$ and $Y$ and (5.52) have been used. Thus, $[X, Y] \in \mathfrak{g}$, that is $\mathfrak{g}$ is closed under the Lie bracket.

It is instructive to work out the left-invariant vector field of $\mathrm{GL}(n, \mathbb{R})$. The coordinates of $\mathrm{GL}(n, \mathbb{R})$ are given by $n^2$ entries $x^{ij}$ of the matrix. The unit element is $e = I_n = (\delta^{ij})$. Let $g = \{x^{ij}(g)\}$ and $a = \{x^{ij}(a)\}$ be elements of $\mathrm{GL}(n, \mathbb{R})$. The left-translation is

$$L_a g = ag = \sum x^{ik}(a)x^{kj}(g).$$

Take a vector $V = \sum V^{ij} \partial/\partial x^{ij}|_e \in T_eG$ where the $V^{ij}$ are the entries of $V$. The left-invariant vector field generated by $V$ is

$$\begin{aligned}
X_V|_g = L_{g*}V &= \sum_{ijklm} V^{ij} \left.\frac{\partial}{\partial x^{ij}}\right|_e x^{kl}(g)x^{lm}(e) \left.\frac{\partial}{\partial x^{km}}\right|_g \\
&= \sum V^{ij} x^{kl}(g)\delta^l_i \delta^m_j \left.\frac{\partial}{\partial x^{km}}\right|_g \\
&= \sum x^{ki}(g)V^{ij} \left.\frac{\partial}{\partial x^{kj}}\right|_g = \sum (gV)^{kj} \left.\frac{\partial}{\partial x^{kj}}\right|_g
\end{aligned} \tag{5.113}$$

where $gV$ is the usual matrix multiplication of $g$ and $V$. The vector $X_V|_g$ is often abbreviated as $gV$ since it gives the components of the vector.

The Lie bracket of $X_V$ and $X_W$ generated by $V = V^{ij}\partial/\partial x^{ij}|_e$ and $W = W^{ij}\partial/\partial x^{ij}|_e$ is

$$\begin{aligned}
[X_V, X_W]|_g &= \sum x^{ki}(g)V^{ij} \left.\frac{\partial}{\partial x^{kj}}\right|_g x^{ca}(g)W^{ab} \left.\frac{\partial}{\partial x^{cb}}\right|_g - (V \leftrightarrow W) \\
&= \sum x^{ij}(g)[V^{jk}W^{kl} - W^{jk}V^{kl}] \left.\frac{\partial}{\partial x^{il}}\right|_g \\
&= \sum (g[V,W])^{ij} \left.\frac{\partial}{\partial x^{ij}}\right|_g .
\end{aligned} \tag{5.114}$$

Clearly, (5.113) and (5.114) remain true for any matrix group and we establish that

$$L_{g*}V = gV \tag{5.115}$$

$$[X_V, X_W]|_g = L_{g*}[V,W] = g[V,W]. \tag{5.116}$$

Now a Lie algebra is defined as the set of left-invariant vector fields $\mathfrak{g}$ with the Lie bracket.

*Definition 5.11.* The set of left-invariant vector fields $\mathfrak{g}$ with the Lie bracket $[\quad,\quad]:\mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ is called the **Lie algebra** of a Lie group $G$.

We denote the Lie algebra of a Lie group by the corresponding lower-case German gothic letter. For example $\mathfrak{so}(n)$ is the Lie algebra of $\mathrm{SO}(n)$.

*Example 5.15.*

(a) Take $G=\mathbb{R}$ as in exercise 5.19(b). If we define the left translation $L_a$ by $x\mapsto x+a$, the left-invariant vector field is given by $X=\partial/\partial x$. In fact,

$$L_{a*}X\Big|_x=\frac{\partial(a+x)}{\partial x}\frac{\partial}{\partial(a+x)}=\frac{\partial}{\partial(x+a)}=X\Big|_{x+a}.$$

Clearly this is the only left-invariant vector field on $\mathbb{R}$. We also find that $X=\partial/\partial\theta$ is the unique left-invariant vector field on $G=\mathrm{SO}(2)=\{\mathrm{e}^{\mathrm{i}\theta}|0\le\theta\le 2\pi\}$. Thus, the Lie groups $\mathbb{R}$ and SO(2) share the common Lie algebra.

(b) Let $\mathfrak{gl}(n,\mathbb{R})$ be the Lie algebra of $\mathrm{GL}(n,\mathbb{R})$ and $c:(-\varepsilon,\varepsilon)\to\mathrm{GL}(n,\mathbb{R})$ be a curve with $c(0)=I_n$. The curve is approximated by $c(s)=I_n+sA+O(s^2)$ near $s=0$, where $A$ is an $n\times n$ matrix of real entries. Note that for small enough $s$, $\det c(s)$ cannot vanish and $c(s)$ is, indeed, in $\mathrm{GL}(n,\mathbb{R})$. The tangent vector to $c(s)$ at $I_n$ is $c'(s)\big|_{s=0}=A$. This shows that $\mathfrak{gl}(n,\mathbb{R})$ is the set of $n\times n$ matrices. Clearly $\dim\mathfrak{gl}(n,\mathbb{R})=n^2=\dim\mathrm{GL}(n,\mathbb{R})$. Subgroups of $\mathrm{GL}(n,\mathbb{R})$ are more interesting.

(c) Let us find the Lie algebra $\mathfrak{sl}(n,\mathbb{R})$ of $\mathrm{SL}(n,\mathbb{R})$. Following this prescription, we approximate a curve through $I_n$ by $c(s)=I_n+sA+O(s^2)$. The tangent vector to $c(s)$ at $I_n$ is $c'(s)\big|_{s=0}=A$. Now, for the curve $c(s)$ to be in $\mathrm{SL}(n,\mathbb{R})$, $c(s)$ has to satisfy $\det c(s)=1+s\mathrm{tr}A=1$, namely $\mathrm{tr}\,A=0$. Thus, $\mathfrak{sl}(n,\mathbb{R})$ is the set of $n\times n$ traceless matrices and $\dim\mathfrak{sl}(n,\mathbb{R})=n^2-1$.

(d) Let $c(s)=I_n+sA+O(s^2)$ be a curve in $\mathrm{SO}(n)$ through $I_n$. Since $c(s)$ is a curve in $\mathrm{SO}(n)$, it satisfies $c(s)^{\mathrm{t}}c(s)=I_n$. Differentiating this identity, we obtain $c'(s)^{\mathrm{t}}c(s)+c(s)^{\mathrm{t}}c'(s)=0$. At $s=0$, this becomes $A^{\mathrm{t}}+A=0$. Hence, $\mathfrak{so}(n)$ is the set of skew-symmetric matrices. Since we are interested only in the vicinity of the unit element, the Lie algebra of $\mathrm{O}(n)$ is the same as that of $\mathrm{SO}(n)$: $\mathfrak{o}(n)=\mathfrak{so}(n)$. It is easy to see that $\dim\mathfrak{o}(n)=\dim\mathfrak{so}(n)=n(n-1)/2$.

(e) A similar analysis can be carried out for matrix groups of $\mathrm{GL}(n,\mathbb{C})$. $\mathfrak{gl}(n,\mathbb{C})$ is the set of $n\times n$ matrices with complex entries and $\dim\mathfrak{gl}(n,\mathbb{C})=2n^2$ (the dimension here is a real dimension). $\mathfrak{sl}(n,\mathbb{C})$ is the set of traceless matrices with real dimension $2(n^2-1)$. To find $\mathfrak{u}(n)$, we consider a curve $c(s)=I_n+sA+O(s^2)$ in $\mathrm{U}(n)$. Since $c(s)^\dagger c(s)=I_n$, we have $c'(s)^\dagger c(s)+c(s)^\dagger c'(s)=0$. At $s=0$, we have $A^\dagger+A=0$.

Hence, $\mathfrak{u}(n)$ is the set of skew-Hermitian matrices with $\dim \mathfrak{u}(n) = n^2$. $\mathfrak{su}(n) = \mathfrak{u}(n) \cap \mathfrak{sl}(n)$ is the set of traceless skew-Hermitian matrices with $\dim \mathfrak{su}(n) = n^2 - 1$.

*Exercise 5.22.* Let

$$c(s) = \begin{pmatrix} \cos s & -\sin s & 0 \\ \sin s & \cos s & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

be a curve in SO(3). Find the tangent vector to this curve at $I_3$.

### 5.6.3 The one-parameter subgroup

A vector field $X \in \mathcal{X}(M)$ generates a flow in $M$ (section 5.3). Here we are interested in the flow generated by a left-invariant vector field.

*Definition 5.12.* A curve $\phi : \mathbb{R} \to G$ is called a **one-parameter subgroup** of $G$ if it satisfies the condition

$$\phi(t)\phi(s) = \phi(t+s). \tag{5.117}$$

It is easy to see that $\phi(0) = e$ and $\phi^{-1}(t) = \phi(-t)$. Note that the curve $\phi$ thus defined is a homomorphism from $\mathbb{R}$ to $G$. Although $G$ may be non-Abelian, a one-parameter subgroup is an Abelian subgroup: $\phi(t)\phi(s) = \phi(t+s) = \phi(s+t) = \phi(s)\phi(t)$.

Given a one-parameter subgroup $\phi : \mathbb{R} \to G$, there exists a vector field $X$, such that

$$\frac{\mathrm{d}\phi^\mu(t)}{\mathrm{d}t} = X^\mu(\phi(t)). \tag{5.118}$$

We now show that the vector field $X$ is left-invariant. First note that the vector field $\mathrm{d}/\mathrm{d}t$ is left-invariant on $\mathbb{R}$, see example 5.15(a). Thus, we have

$$(L_t)_* \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_0 = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_t . \tag{5.119}$$

Next, we apply the induced map $\phi_* : T_t\mathbb{R} \to T_{\phi(t)}G$ on the vectors $\mathrm{d}/\mathrm{d}t|_0$ and $\mathrm{d}/\mathrm{d}t|_t$,

$$\phi_* \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_0 = \left.\frac{\mathrm{d}\phi^\mu(t)}{\mathrm{d}t}\right|_0 \left.\frac{\partial}{\partial g^\mu}\right|_e = X|_e \tag{5.120a}$$

$$\phi_* \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_t = \left.\frac{\mathrm{d}\phi^\mu(t)}{\mathrm{d}t}\right|_t \left.\frac{\partial}{\partial g^\mu}\right|_g = X|_g \tag{5.120b}$$

where we put $\phi(t) = g$. From (5.119) and (5.120b), we have

$$(\phi L_t)_* \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_0 = \phi_* L_{t*} \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_0 = X|_g. \tag{5.121a}$$

It follows from the commutativity $\phi L_t = L_g \phi$ that $\phi_* L_{t*} = L_{g*}\phi_*$. Then (5.121a) becomes

$$\phi_* L_{t*} \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_0 = L_{g*}\phi_* \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_0 = L_{g*} X|_e. \tag{5.121b}$$

From (5.121), we conclude that

$$L_{g*} X|_e = X|_g. \tag{5.122}$$

Thus, given a flow $\phi(t)$, there exists an associated left-invariant vector field $X \in \mathfrak{g}$.

Conversely, a left-invariant vector field $X$ defines a one-parameter group of transformations $\sigma(t, g)$ such that $\mathrm{d}\sigma(t, g)/\mathrm{d}t = X$ and $\sigma(0, g) = g$. If we define $\phi : \mathbb{R} \to G$ by $\phi(t) \equiv \sigma(t, e)$, the curve $\phi(t)$ becomes a one-parameter subgroup of $G$. To prove this, we have to show $\phi(s + t) = \phi(s)\phi(t)$. By definition, $\sigma$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\sigma(t, \sigma(s, e)) = X(\sigma(t, \sigma(s, e))). \tag{5.123}$$

[We have omitted the coordinate indices for notational simplicity. If readers feel uneasy, they may supplement the indices as in (5.118).] If the parameter $s$ is fixed, $\bar{\sigma}(t, \phi(s)) \equiv \phi(s)\phi(t)$ is a curve $\mathbb{R} \to G$ at $\phi(s)\phi(0) = \phi(s)$. Clearly $\sigma$ and $\bar{\sigma}$ satisfy the same initial condition,

$$\sigma(0, \sigma(s, e)) = \bar{\sigma}(0, \phi(s)) = \phi(s). \tag{5.124}$$

$\bar{\sigma}$ also satisfies the same differential equation as $\sigma$:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\bar{\sigma}(t, \phi(t)) &= \frac{\mathrm{d}}{\mathrm{d}t}\phi(s)\phi(t) = (L_{\phi(s)})_* \frac{\mathrm{d}}{\mathrm{d}t}\phi(t) \\
&= (L_{\phi(s)})_* X(\phi(t)) \\
&= X(\phi(s)\phi(t)) \qquad \text{(left-invariance)} \\
&= X(\bar{\sigma}(t, \phi(s))).
\end{aligned} \tag{5.125}$$

From the uniqueness theorem of ODEs, we conclude that

$$\phi(s + t) = \phi(s)\phi(t). \tag{5.126}$$

We have found that there is a one-to-one correspondence between a one-parameter subgroup of $G$ and a left-invariant vector field. This correspondence becomes manifest if we define the exponential map as follows.

*Definition 5.13.* Let $G$ be a Lie group and $V \in T_eG$. The exponential map $\exp : T_eG \to G$ is defined by

$$\exp V \equiv \phi_V(1) \tag{5.127}$$

where $\phi_V$ is a one-parameter subgroup of $G$ generated by the left-invariant vector field $X_V|_g = L_{g*}V$.

*Proposition 5.2.* Let $V \in T_e G$ and let $t \in \mathbb{R}$. Then

$$\exp(tV) = \phi_V(t) \tag{5.128}$$

where $\phi_V(t)$ is a one-parameter subgroup generated by $X_V|_g = L_{g*}V$.

*Proof.* Let $a \neq 0$ be a constant. Then $\phi_V(at)$ satisfies

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\phi_V(at)\right|_{t=0} = a\left.\frac{\mathrm{d}}{\mathrm{d}t}\phi_V(t)\right|_{t=0} = aV$$

which shows that $\phi_V(at)$ is a one-parameter subgroup generated by $L_{g*}aV$. The left-invariant vector field $L_{g*}aV$ also generates $\phi_{aV}(t)$ and, from the uniqueness of the solution, we find that $\phi_V(at) = \phi_{aV}(t)$. From definition 5.13, we have

$$\exp(aV) = \phi_{aV}(1) = \phi_V(a).$$

The proof is completed if $a$ is replaced by $t$. □

For a matrix group, the exponential map is given by the exponential of a matrix. Take $G = \mathrm{GL}(n, \mathbb{R})$ and $A \in \mathfrak{gl}(n, \mathbb{R})$. Let us define a one-parameter subgroup $\phi_A : \mathbb{R} \to \mathrm{GL}(n, \mathbb{R})$ by

$$\phi_A(t) = \exp(tA) = I_n + tA + \frac{t^2}{2!}A^2 + \cdots + \frac{t^n}{n!}A^n + \cdots. \tag{5.129}$$

In fact, $\phi_A(t) \in \mathrm{GL}(n, \mathbb{R})$ since $[\phi_A(t)]^{-1} = \phi_A(-t)$ exists. It is also easy to see $\phi_A(t)\phi_A(s) = \phi(t+s)$. Now the exponential map is given by

$$\phi_A(1) = \exp(A) = I_n + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{n!}A^n + \cdots. \tag{5.130}$$

The curve $g\exp(tA)$ is a flow through $g \in G$. We find that

$$\left.\frac{\mathrm{d}}{\mathrm{d}t} g\exp(tA)\right|_{t=0} = L_{g*}A = X_A|_g$$

where $X_A$ is a left-invariant vector field generated by $A$. From (5.115), we find, for a matrix group $G$, that

$$L_{g*}A = X_A|_g = gA. \tag{5.131}$$

The curve $g\exp(tA)$ defines a map $\sigma_t : G \to G$ by $\sigma_t(g) \equiv g\exp(tA)$ which is also expressed as a right-translation,

$$\sigma_t = R_{\exp(tA)}. \tag{5.132}$$

### 5.6.4 Frames and structure equation

Let the set of $n$ vectors $\{V_1, V_2, \ldots, V_n\}$ be a basis of $T_eG$ where $n = \dim G$. [We assume throughout this book that $n$ is finite.] The basis defines the set of $n$ linearly independent left-invariant vector fields $\{X_1, X_2, \ldots, X_n\}$ at each point $g$ in $G$ by $X_\mu\big|_g = L_{g*}V_\mu$. Note that the set $\{X_\mu\}$ is a frame of a basis defined throughout $G$. Since $[X_\mu, X_\nu]|_g$ is again an element of $\mathfrak{g}$ at $g$, it can be expanded in terms of $\{X_\mu\}$ as

$$[X_\mu, X_\nu] = c_{\mu\nu}{}^\lambda X_\lambda \tag{5.133}$$

where $c_{\mu\nu}{}^\lambda$ are called the **structure constants** of the Lie group $G$. If $G$ is a matrix group, the LHS of (5.133) at $g = e$ is precisely the commutator of matrices $V_\mu$ and $V_\nu$; see (5.116). We show that the $c_{\mu\nu}{}^\lambda$ are, indeed, constants independent of $g$. Let $c_{\mu\nu}{}^\lambda(e)$ be the structure constants at the unit element. If $L_{g*}$ is applied to the Lie bracket, we have

$$[X_\mu, X_\nu]|_g = c_{\mu\nu}{}^\lambda(e) X_\lambda|_g$$

which shows the $g$-independence of the structure constants. In a sense, the structure constants determine a Lie group completely (Lie's theorem).

*Exercise 5.23.* Show that the structure constants satisfy

(a) *skew-symmetry*

$$c_{\mu\nu}{}^\lambda = -c_{\nu\mu}{}^\lambda \tag{5.134}$$

(b) *Jacobi identity*

$$c_{\mu\nu}{}^\tau c_{\tau\rho}{}^\lambda + c_{\rho\mu}{}^\tau c_{\tau\nu}{}^\lambda + c_{\nu\rho}{}^\tau c_{\tau\mu}{}^\lambda = 0. \tag{5.135}$$

Let us introduce a dual basis to $\{X_\mu\}$ and denote it by $\{\theta^\mu\}$; $\langle \theta^\mu, X_\nu \rangle = \delta^\mu_\nu$. $\{\theta^\mu\}$ is a basis for the left-invariant one-forms. We will show that the dual basis satisfies **Maurer–Cartan's structure equation**,

$$\mathrm{d}\theta^\mu = -\tfrac{1}{2} c_{\nu\lambda}{}^\mu \theta^\nu \wedge \theta^\lambda. \tag{5.136}$$

This can be seen by making use of (5.70):

$$\begin{aligned}\mathrm{d}\theta^\mu(X_\nu, X_\lambda) &= X_\nu[\theta^\mu(X_\lambda)] - X_\lambda[\theta^\mu(X_\nu)] - \theta^\mu([X_\nu, X_\lambda]) \\ &= X_\nu[\delta^\mu_\lambda] - X_\lambda[\delta^\mu_\nu] - \theta^\mu(c_{\nu\lambda}{}^\kappa X_\kappa) = -c_{\nu\lambda}{}^\mu\end{aligned}$$

which proves (5.136).

We define a Lie-algebra-valued one-form $\theta : T_gG \to T_eG$ by

$$\theta : X \mapsto (L_{g^{-1}})_* X = (L_g)_*^{-1} X \qquad X \in T_gG. \tag{5.137}$$

$\theta$ is called the **canonical one-form** or **Maurer–Cartan form** on $G$.

*Theorem 5.3.* (a) The canonical one-form $\theta$ is expanded as

$$\theta = V_\mu \otimes \theta^\mu \tag{5.138}$$

where $\{V_\mu\}$ is the basis of $T_eG$ and $\{\theta^\mu\}$ the dual basis of $T_e^*G$.
(b) The canonical one-form $\theta$ satisfies

$$\mathrm{d}\theta + \tfrac{1}{2}[\theta \wedge \theta] = 0 \tag{5.139}$$

where $\mathrm{d}\theta \equiv V_\mu \otimes \mathrm{d}\theta^\mu$ and

$$[\theta \wedge \theta] \equiv [V_\mu, V_\nu] \otimes \theta^\mu \wedge \theta^\nu. \tag{5.140}$$

*Proof.*

(a) Take any vector $Y = Y^\mu X_\mu \in T_gG$, where $\{X_\mu\}$ is the set of frame vectors generated by $\{V_\mu\}$; $X_\mu|_g = L_{g*}V_\mu$. From (5.137), we find

$$\theta(Y) = Y^\mu \theta(X_\mu) = Y^\mu (L_{g*})^{-1}[L_{g*}V_\mu] = Y^\mu V_\mu.$$

However,

$$(V_\mu \otimes \theta^\mu)(Y) = Y^\nu V_\mu \theta^\mu(X_\nu) = Y^\nu V_\mu \delta^\mu_\nu = Y^\mu V_\mu.$$

Since $Y$ is arbitrary, we have $\theta = V_\mu \otimes \theta^\mu$.
(b) We use the Maurer–Cartan structure equation (5.136):

$$\mathrm{d}\theta + \tfrac{1}{2}[\theta \wedge \theta] = -\tfrac{1}{2}V_\mu \otimes c_{\nu\lambda}{}^\mu \theta^\nu \wedge \theta^\lambda + \tfrac{1}{2}c_{\nu\lambda}{}^\mu V_\mu \otimes \theta^\nu \wedge \theta^\lambda = 0$$

where the $c_{\nu\lambda}{}^\mu$ are the structure constants of $G$. □

## 5.7 The action of Lie groups on manifolds

In physics, a Lie group often appears as the set of transformations acting on a manifold. For example, SO(3) is the group of rotations in $\mathbb{R}^3$, while the Poincaré group is the set of transformations acting on the Minkowski spacetime. To study more general cases, we abstract the action of a Lie group $G$ on a manifold $M$. We have already encountered this interaction between a group and geometry. In section 5.3 we defined a flow in a manifold $M$ as a map $\sigma : \mathbb{R} \times M \to M$, in which $\mathbb{R}$ acts as an additive group. We abstract this idea as follows.

### 5.7.1 Definitions

*Definition 5.14.* Let $G$ be a Lie group and $M$ be a manifold. The **action** of $G$ on $M$ is a differentiable map $\sigma : G \times M \to M$ which satisfies the conditions

$$\text{(i)} \qquad \sigma(e, p) = p \qquad \text{for any } p \in M \tag{5.141a}$$

$$\text{(ii)} \qquad \sigma(g_1, \sigma(g_2, p)) = \sigma(g_1 g_2, p). \tag{5.141b}$$

[*Remark:* We often use the notation $gp$ instead of $\sigma(g, p)$. The second condition in this notation is $g_1(g_2 p) = (g_1 g_2)p$.]

*Example 5.16.* (a) A flow is an action of $\mathbb{R}$ on a manifold $M$. If a flow is periodic with a period $T$, it may be regarded as an action of U(1) or SO(2) on $M$. Given a periodic flow $\sigma(t, x)$ with period $T$, we construct a new action $\bar{\sigma}(\exp(2\pi \mathrm{i}t/T), x) \equiv \sigma(t, x)$ whose group $G$ is U(1).

(b) Let $M \in \mathrm{GL}(n, \mathbb{R})$ and let $x \in \mathbb{R}^n$. The action of $\mathrm{GL}(n, \mathbb{R})$ on $\mathbb{R}^n$ is defined by the usual matrix action on a vector:

$$\sigma(M, x) = M \cdot x. \tag{5.142}$$

The action of the subgroups of $\mathrm{GL}(n, \mathbb{R})$ is defined similarly. They may also act on a smaller space. For example, $\mathrm{O}(n)$ acts on $S^{n-1}(r)$, an $(n-1)$-sphere of radius $r$,

$$\sigma : \mathrm{O}(n) \times S^{n-1}(r) \to S^{n-1}(r). \tag{5.143}$$

(c) It is known that $\mathrm{SL}(2, \mathbb{C})$ acts on a four-dimensional Minkowski space $M_4$ in a special manner. For $x = (x^0, x^1, x^2, x^3) \in M_4$, define a Hermitian matrix,

$$X(x) \equiv x^\mu \sigma_\mu = \begin{pmatrix} x^0 + x^3 & x^1 - \mathrm{i}x^2 \\ x^1 + \mathrm{i}x^2 & x^0 - x^3 \end{pmatrix} \tag{5.144}$$

where $\sigma_\mu = (I_2, \sigma_1, \sigma_2, \sigma_3)$, $\sigma_i$ $(i = 1, 2, 3)$ being the Pauli matrices. Conversely, given a Hermitian matrix $X$, a unique vector $(x^\mu) \in M_4$ is defined as

$$x^\mu = \tfrac{1}{2}\,\mathrm{tr}(\sigma_\mu X) \tag{5.130}$$

where tr is over the $2 \times 2$ matrix indices. Thus, there is an isomorphism between $M_4$ and the set of $2 \times 2$ Hermitian matrices. It is interesting to note that $\det X(x) = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = -X^{\mathrm{t}}\eta X = -(\text{Minkowski norm})^2$. Accordingly

$$\begin{aligned} \det X(x) &> 0 \qquad \text{if } x \text{ is a timelike vector} \\ &= 0 \qquad \text{if } x \text{ is on the light cone} \\ &< 0 \qquad \text{if } x \text{ is a spacelike vector.} \end{aligned}$$

Take $A \in \mathrm{SL}(2, \mathbb{C})$ and define an action of $\mathrm{SL}(2, \mathbb{C})$ on $M_4$ by

$$\sigma(A, x) \equiv AX(x)A^\dagger. \tag{5.145}$$

The reader should verify that this action, in fact, satisfies the axioms of definition 5.14. The action of $\mathrm{SL}(2, \mathbb{C})$ on $M_4$ represents the Lorentz transformation $\mathrm{O}(1, 3)$. First we note that the action preserves the Minkowski norm,

$$\det \sigma(A, x) = \det[AX(x)A^\dagger] = \det X(x)$$

since $\det A = \det A^\dagger = 1$. Moreover, there is a homomorphism $\varphi : \mathrm{SL}(2, \mathbb{C}) \to \mathrm{O}(1, 3)$ since

$$A(BXB^\dagger)A^\dagger = (AB)X(AB)^\dagger.$$

However, this homomorphism cannot be one to one, since $A \in \mathrm{SL}(2, \mathbb{C})$ and $-A$ give the same element of O(1, 3); see (5.145). We verify (exercise 5.24) that the following matrix is an explicit form of a rotation about the unit vector $\hat{\boldsymbol{n}}$ by an angle $\theta$,

$$A = \exp\left[-\mathrm{i}\frac{\theta}{2}(\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma})\right] = \cos\frac{\theta}{2} I_2 - \mathrm{i}(\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}) \sin\frac{\theta}{2}. \tag{5.146a}$$

The appearance of $\theta/2$ ensures that the homomorphism between $\mathrm{SL}(2, \mathbb{C})$ and the O(3) subgroup of O(1, 3) is indeed two to one. In fact, rotations about $\hat{\boldsymbol{n}}$ by $\theta$ and by $2\pi + \theta$ should be the same O(3) rotation, but $A(2\pi + \theta) = -A(\theta)$ in $\mathrm{SL}(2, \mathbb{C})$. This leads to the existence of spinors. [See Misner *et al* (1973) and Wald (1984).] A boost along the direction $\hat{\boldsymbol{n}}$ with the velocity $v = \tanh\alpha$ is given by

$$A = \exp\left[\frac{\alpha}{2}(\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma})\right] = \cosh\frac{\alpha}{2} I_2 + (\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}) \sinh\frac{\alpha}{2}. \tag{5.146b}$$

We show that $\varphi$ maps $\mathrm{SL}(2, \mathbb{C})$ onto the proper orthochronous Lorentz group $\mathrm{O}_+^\uparrow(1, 3) = \{\Lambda \in \mathrm{O}(1, 3) | \det\Lambda = +1, \Lambda_{00} > 0\}$. Take any

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C})$$

and suppose $x^\mu = (1, 0, 0, 0)$ is mapped to $x'^\mu$. If we write $\varphi(A) = \Lambda$, we have

$$\begin{aligned} x'^0 &= \frac{1}{2}\operatorname{tr}(AXA^\dagger) = \frac{1}{2}\operatorname{tr}\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix}\right] \\ &= \frac{1}{2}(|a|^2 + |b|^2 + |c|^2 + |d|^2) > 0 \end{aligned}$$

hence $\Lambda_{00} > 0$. To show $\det A = +1$, we note that any element of $\mathrm{SL}(2, \mathbb{C})$ may be written as

$$\begin{aligned} A &= \begin{pmatrix} \mathrm{e}^{\mathrm{i}\alpha} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\alpha} \end{pmatrix}\begin{pmatrix} \cos\beta & \sin\beta\,\mathrm{e}^{\mathrm{i}\gamma} \\ -\sin\beta\,\mathrm{e}^{-\mathrm{i}\gamma} & \cos\beta \end{pmatrix} B \\ &= \begin{pmatrix} \mathrm{e}^{\mathrm{i}\alpha/2} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\alpha/2} \end{pmatrix}^2 \begin{pmatrix} \cos(\beta/2) & \sin(\beta/2)\mathrm{e}^{\mathrm{i}\gamma} \\ -\sin(\beta/2)\mathrm{e}^{-\mathrm{i}\gamma} & \cos(\beta/2) \end{pmatrix}^2 B \\ &\equiv M^2 N^2 B_0^2 \end{aligned}$$

where $B \equiv B_0^2$ is a positive-definite matrix. This shows that $\varphi(A)$ is positive definite:

$$\det\varphi(A) = (\det\varphi(M))^2(\det\varphi(N))^2(\det\varphi(B_0))^2 > 0.$$

Now we have established that $\varphi(\mathrm{SL}(2,\mathbb{C})) \subset \mathrm{O}_+^\uparrow(1,3)$. Equations (5.146a) and (5.146b) show that for any element of $\mathrm{O}_+^\uparrow(1,3)$, there is a corresponding matrix $A \in \mathrm{SL}(2,\mathbb{C})$, hence $\varphi$ is onto. Thus, we have established that

$$\varphi(\mathrm{SL}(2,\mathbb{C})) = \mathrm{O}_+^\uparrow(1,3). \tag{5.147}$$

It can be shown that $\mathrm{SL}(2,\mathbb{C})$ is simply connected and is the universal covering group $\mathrm{SPIN}(1,3)$ of $\mathrm{O}_+^\uparrow(1,3)$, see section 4.6.

*Exercise 5.24.* Verify by explicit calculations that

(a)

$$A = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\theta/2} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\theta/2} \end{pmatrix}$$

represents a rotation about the $z$-axis by $\theta$;

(b)

$$A = \begin{pmatrix} \cosh(\alpha/2) + \sinh(\alpha/2) & 0 \\ 0 & \cosh(\alpha/2) - \sinh(\alpha/2) \end{pmatrix}$$

represents a boost along the $z$-axis with the velocity $v = \tanh \alpha$.

*Definition 5.15.* Let $G$ be a Lie group that acts on a manifold $M$ by $\sigma : G \times M \to M$. The action $\sigma$ is said to be

(a) **transitive** if, for any $p_1, p_2 \in M$, there exists an element $g \in G$ such that $\sigma(g, p_1) = p_2$;

(b) **free** if every non-trivial element $g \neq e$ of $G$ has no fixed points in $M$, that is, if there exists an element $p \in M$ such that $\sigma(g, p) = p$, then $g$ must be the unit element $e$; and

(c) **effective** if the unit element $e \in G$ is the unique element that defines the trivial action on $M$, i.e. if $\sigma(g, p) = p$ for all $p \in M$, then $g$ must be the unit element $e$.

*Exercise 5.25.* Show that the right translation $R : (a, g) \mapsto R_a g$ and left translation $L : (a, g) \mapsto L_a g$ of a Lie group are free and transitive.

### 5.7.2 Orbits and isotropy groups

Given a point $p \in M$, the action of $G$ on $p$ takes $p$ to various points in $M$. The **orbit** of $p$ under the action $\sigma$ is the subset of $M$ defined by

$$Gp = \{\sigma(g, p) | g \in G\}. \tag{5.148}$$

If the action of $G$ on $M$ is transitive, the orbit of any $p \in M$ is $M$ itself. Clearly the action of $G$ on any orbit $Gp$ is transitive.

*Example 5.17.* (a) A flow $\sigma$ generated by a vector field $X = -y\partial/\partial x + x\partial/\partial y$ is periodic with period $2\pi$, see example 5.9. The action $\sigma : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ defined by $(t, (x, y)) \to \sigma(t, (x, y))$ is not effective since $\sigma(2\pi n, (x, y)) = (x, y)$ for all $(x, y) \in \mathbb{R}^2$. For the same reason, this flow is not free either. The orbit through $(x, y) \neq (0, 0)$ is a circle $S^1$ centred at the origin.

(b) The action of O($n$) on $\mathbb{R}^n$ is not transitive since if $|x| \neq |x'|$, no element of O($n$) takes $x$ to $x'$. However, the action of O($n$) on $S^{n-1}$ is obviously transitive. The orbit through $x$ is the sphere $S^{n-1}$ of radius $|x|$. Accordingly, given an action $\sigma : \mathrm{O}(n) \times \mathbb{R}^n \to \mathbb{R}^n$, the orbits divide $\mathbb{R}^n$ into mutually disjoint spheres of different radii. Introduce a relation by $x \sim y$ if $y = \sigma(g, x)$ for some $g \in G$. It is easily verified that $\sim$ is an equivalence relation. The equivalence class $[x]$ is an orbit through $x$. The coset space $\mathbb{R}^n/\mathrm{O}(n)$ is $[0, \infty)$ since each equivalence class is parametrized by the radius.

*Definition 5.16.* Let $G$ be a Lie group that acts on a manifold $M$. The **isotropy group** of $p \in M$ is a subgroup of $G$ defined by

$$H(p) = \{g \in G | \sigma(g, p) = p\}. \tag{5.149}$$

$H(p)$ is also called the **little group** or **stabilizer** of $p$.

It is easy to see that $H(p)$ is indeed a subgroup. Let $g_1, g_2 \in H(p)$, then $g_1 g_2 \in H(p)$ since $\sigma(g_1 g_2, p) = \sigma(g_1, \sigma(g_2, p)) = \sigma(g_1, p) = p$. Clearly $e \in H(p)$ since $\sigma(e, p) = p$ by definition. If $g \in H(p)$, then $g^{-1} \in H(p)$ since $p = \sigma(e, p) = \sigma(g^{-1}g, p) = \sigma(g^{-1}, \sigma(g, p)) = \sigma(g^{-1}, p)$.

*Exercise 5.26.* Suppose a Lie group $G$ acts on a manifold $M$ freely. Show that $H(p) = \{e\}$ for any $p \in M$.

*Theorem 5.4.* Let $G$ be a Lie group which acts on a manifold $M$. Then the isotropy group $H(p)$ for any $p \in M$ is a Lie subgroup.

*Proof.* For fixed $p \in M$, we define a map $\varphi_p : G \to M$ by $\varphi_p(g) \equiv gp$. Then $H(p)$ is the inverse image $\varphi_p^{-1}(p)$ of a *point* $p$, and hence a closed set. The group properties have been shown already. It follows from theorem 5.2 that $H(p)$ is a Lie subgroup. □

For example, let $M = \mathbb{R}^3$ and $G = \mathrm{SO}(3)$ and take a point $p = (0, 0, 1) \in \mathbb{R}^3$. The isotropy group $H(p)$ is the set of rotations about the $z$-axis, which is isomorphic to SO(2).

Let $G$ be a Lie group and $H$ any subgroup of $G$. The coset space $G/H$ admits a differentiable structure and $G/H$ becomes a manifold, called a **homogeneous space**. Note that $\dim G/H = \dim G - \dim H$. Let $G$ be a Lie group which acts on a manifold $M$ transitively and let $H(p)$ be an isotropy group of $p \in M$. $H(p)$ is a Lie subgroup and the coset space $G/H(p)$ is a homogeneous space.

In fact, if $G$, $H(p)$ and $M$ satisfy certain technical requirements (for example, $G/H(p)$ compact) is, it can be shown that $G/H(p)$ is homeomorphic to $M$, see example 5.18.

*Example 5.18.* (a) Let $G = \mathrm{SO}(3)$ be a group acting on $\mathbb{R}^3$ and $H = \mathrm{SO}(2)$ be the isotropy group of $x \in \mathbb{R}^3$. The group SO(3) acts on $S^2$ transitively and we have $\mathrm{SO}(3)/\mathrm{SO}(2) \cong S^2$. What is the geometrical picture of this? Let $g' = gh$ where $g, g' \in G$ and $h \in H$. Since $H$ is the set of rotations in a plane, $g$ and $g'$ must be rotations about the common axis. Then the equivalence class $[g]$ is specified by the polar angles $(\theta, \phi)$. Thus, we again find that $G/H = S^2$. Since SO(2) is not a normal subgroup of SO(3), $S^2$ does not admit a group structure.

It is easy to generalize this result to higher-dimensional rotation groups and we have the useful result

$$\mathrm{SO}(n+1)/\mathrm{SO}(n) = S^n. \tag{5.150}$$

$\mathrm{O}(n+1)$ also acts on $S^n$ transitively and we have

$$\mathrm{O}(n+1)/\mathrm{O}(n) = S^n. \tag{5.151}$$

Similar relations hold for $\mathrm{U}(n)$ and $\mathrm{SU}(n)$:

$$\mathrm{U}(n+1)/\mathrm{U}(n) = \mathrm{SU}(n+1)/\mathrm{SU}(n) = S^{2n+1}. \tag{5.152}$$

(b) The group $\mathrm{O}(n+1)$ acts on $\mathbb{R}P^n$ transitively from the left. Note, first, that $\mathrm{O}(n+1)$ acts on $\mathbb{R}^{n+1}$ in the usual manner and preserves the equivalence relation employed to define $\mathbb{R}P^n$ (see example 5.12). In fact, take $x, x' \in \mathbb{R}^{n+1}$ and $g \in \mathrm{O}(n+1)$. If $x \sim x'$ (that is if $x' = ax$ for some $a \in \mathbb{R} - \{0\}$), then it follows that $gx \sim gx'$ ($gx' = agx$). Accordingly, this action of $\mathrm{O}(n+1)$ on $\mathbb{R}^{n+1}$ induces the natural action of $\mathrm{O}(n+1)$ on $\mathbb{R}P^n$. Clearly this action is transitive on $\mathbb{R}P^n$. (Look at two representatives with the same norm.) If we take a point $p$ in $\mathbb{R}P^n$, which corresponds to a point $(1, 0, \ldots, 0) \in \mathbb{R}^{n+1}$, the isotropy group $H(p)$ is

$$H(p) = \begin{pmatrix} \pm 1 & 0 & 0 & \ldots & 0 \\ 0 & & & & \\ 0 & & & & \\ \vdots & & \mathrm{O}(n) & & \\ 0 & & & & \end{pmatrix} = \mathrm{O}(1) \times \mathrm{O}(n) \tag{5.153}$$

where $\mathrm{O}(1)$ is the set $\{-1, +1\} = \mathbb{Z}_2$. Now we find that

$$\mathrm{O}(n+1)/[\mathrm{O}(1) \times \mathrm{O}(n)] \cong S^n/\mathbb{Z}_2 \cong \mathbb{R}P^n. \tag{5.154}$$

(c) This result is easily generalized to the Grassmann manifolds: $G_{k,n}(\mathbb{R}) = \mathrm{O}(n)/[\mathrm{O}(k) \times \mathrm{O}(n-k)]$. We first show that $\mathrm{O}(n)$ acts on $G_{k,n}(\mathbb{R})$ transitively.

Let $A$ be an element of $G_{k,n}(\mathbb{R})$, then $A$ is a $k$-dimensional plane in $\mathbb{R}^n$. Define an $n \times n$ matrix $P_A$ which projects a vector $v \in \mathbb{R}^n$ to the plane $A$. Let us introduce an orthonormal basis $\{e_1, \ldots, e_n\}$ in $\mathbb{R}^n$ and another orthonormal basis $\{f_1, \ldots, f_k\}$ in the plane $A$, where the orthonormality is defined with respect to the Euclidean metric in $\mathbb{R}^n$. In terms of $\{e_i\}$, $f_a$ is expanded as $f_a = \sum_i f_{ai} e_i$ and the projected vector is

$$\begin{aligned} P_A v &= (vf_1)f_1 + \cdots + (vf_k)f_k \\ &= \sum_{i,j} (v_i f_{1i} f_{1j} + \cdots + v_i f_{ki} f_{kj}) e_j = \sum_{i,a,j} v_i f_{ai} f_{aj} e_j. \end{aligned}$$

Thus, $P_A$ is represented by a matrix

$$(P_A)_{ij} = \sum f_{ai} f_{aj}. \tag{5.155}$$

Note that $P_A^2 = P_A$, $P_A^{\mathrm{t}} = P_A$ and $\operatorname{tr} P_A = k$. [The last relation holds since it is always possible to choose a coordinate system such that

$$P_A = \operatorname{diag}(\underbrace{1, 1, \ldots, 1}_{k}, \underbrace{0, \ldots, 0}_{n-k}).$$

This guarantees that $A$ is, indeed, a $k$-dimensional plane.] Conversely any matrix $P$ that satisfies these three conditions determines a unique $k$-dimensional plane in $\mathbb{R}^n$, that is a unique element of $G_{k,n}(\mathbb{R})$.

We now show that $\mathrm{O}(n)$ acts on $G_{k,n}(\mathbb{R})$ transitively. Take $A \in G_{k,n}(\mathbb{R})$ and $g \in \mathrm{O}(n)$ and construct $P_B \equiv g P_A g^{-1}$. The matrix $P_B$ determines an element $B \in G_{k,n}(\mathbb{R})$ since $P_B^2 = P_B$, $P_B^{\mathrm{t}} = P_B$ and $\operatorname{tr} P_B = k$. Let us denote this action by $B = \sigma(g, A)$. Clearly this action is transitive since given a standard $k$-dimensional basis of $A$, $\{f_1, \ldots, f_k\}$ for example, any $k$-dimensional basis $\{\widetilde{f}_1, \ldots, \widetilde{f}_k\}$ can be reached by an action of $\mathrm{O}(n)$ on this basis.

Let us take a special plane $C_0$ which is spanned by the standard basis $\{f_1, \ldots, f_k\}$. Then an element of the isotropy group $H(C_0)$ is of the form

$$M = \begin{array}{cc} \begin{array}{cc} k & n-k \end{array} & \\ \begin{pmatrix} g_1 & 0 \\ 0 & g_2 \end{pmatrix} & \begin{array}{c} k \\ n-k \end{array} \end{array} \tag{5.156}$$

where $g_1 \in \mathrm{O}(k)$. Since $M \in \mathrm{O}(n)$, an $(n-k) \times (n-k)$ matrix $g_2$ must be an element of $\mathrm{O}(n-k)$. Thus, the isotropy group is isomorphic to $\mathrm{O}(k) \times \mathrm{O}(n-k)$. Finally we verified that

$$G_{k,n}(\mathbb{R}) \cong \mathrm{O}(n)/[\mathrm{O}(k) \times \mathrm{O}(n-k)]. \tag{5.157}$$

The dimension of $G_{k,n}(\mathbb{R})$ is obtained from the general formula as

$$\begin{aligned} \dim G_{k,n}(\mathbb{R}) &= \dim \mathrm{O}(n) - \dim[\mathrm{O}(k) \times \mathrm{O}(n-k)] \\ &= \tfrac{1}{2}n(n-1) - [\tfrac{1}{2}k(k-1) + \tfrac{1}{2}(n-k)(n-k-1)] \\ &= k(n-k) \end{aligned} \tag{5.158}$$

in agreement with the result of example 5.5. Equation (5.157) also shows that the Grassmann manifold is compact.

### 5.7.3 Induced vector fields

Let $G$ be a Lie group which acts on $M$ as $(g, x) \mapsto gx$. A left-invariant vector field $X_V$ generated by $V \in T_eG$ naturally induces a vector field in $M$. Define a flow in $M$ by

$$\sigma(t, x) = \exp(tV)x, \tag{5.159}$$

$\sigma(t, x)$ is a one-parameter group of transformations, and define a vector field called the **induced vector field** denoted by $V^\sharp$,

$$V^\sharp|_x = \left.\frac{\mathrm{d}}{\mathrm{d}t}\exp(tV)x\right|_{t=0}. \tag{5.160}$$

Thus, we have obtained a map $\sharp : T_eG \to \mathfrak{X}(M)$ defined by $V \mapsto V^\sharp$.

*Exercise 5.27.* The Lie group SO(2) acts on $M = \mathbb{R}^2$ in the usual way. Let

$$V = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

be an element of $\mathfrak{so}(2)$.

(a) Show that

$$\exp(tV) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

and find the induced flow through

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

(b) Show that $V^\sharp|_x = -y\partial/\partial x + x\partial/\partial y$.

*Example 5.19.* Let us take $G = \mathrm{SO}(3)$ and $M = \mathbb{R}^3$. The basis vectors of $T_eG$ are generated by rotations about the $x$, $y$ and $z$ axes. We denote them by $X_x$, $X_y$ and $X_z$, respectively (see exercise 5.22),

$$X_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad X_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad X_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Repeating a similar analysis to the previous one, we obtain the corresponding induced vectors,

$$X_x^\sharp = -z\frac{\partial}{\partial y} + y\frac{\partial}{\partial z}, \qquad X_y^\sharp = -x\frac{\partial}{\partial z} + z\frac{\partial}{\partial x}, \qquad X_z^\sharp = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}.$$

### 5.7.4 The adjoint representation

A Lie group $G$ acts on $G$ itself in a special way.

*Definition 5.17.* Take any $a \in G$ and define a homomorphism $\mathrm{ad}_a : G \to G$ by the conjugation,

$$\mathrm{ad}_a : g \mapsto aga^{-1}. \tag{5.161}$$

This homomorphism is called the **adjoint representation** of $G$.

*Exercise 5.28.* Show that $\mathrm{ad}_a$ is a homomorphism. Define a map $\sigma : G \times G \to G$ by $\sigma(a, g) \equiv \mathrm{ad}_a g$. Show that $\sigma(a, g)$ is an action of $G$ on itself.

Noting that $\mathrm{ad}_a e = e$, we restrict the induced map $\mathrm{ad}_{a*} : T_g G \to T_{\mathrm{ad}_a g} G$ to $g = e$,

$$\mathrm{Ad}_a : T_e G \to T_e G \tag{5.162}$$

where $\mathrm{Ad}_a \equiv \mathrm{ad}_{a*}|_{T_e G}$. If we identify $T_e G$ with the Lie algebra $\mathfrak{g}$, we have obtained a map $\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g}$ called the **adjoint map** of $G$. Since $\mathrm{ad}_{a*}\mathrm{ad}_{b*} = \mathrm{ad}_{ab*}$, it follows that $\mathrm{Ad}_a \mathrm{Ad}_b = \mathrm{Ad}_{ab}$. Similarly, $\mathrm{Ad}_{a^{-1}} = \mathrm{Ad}_a^{-1}$ follows from $\mathrm{ad}_{a^{-1}*}\mathrm{ad}_{a*}|_{T_e G} = \mathrm{id}_{T_e G}$.

If $G$ is a matrix group, the adjoint representation becomes a simple matrix operation. Let $g \in G$ and $X_V \in \mathfrak{g}$, and let $\sigma_V(t) = \exp(tV)$ be a one-parameter subgroup generated by $V \in T_e G$. Then $\mathrm{ad}_g$ acting on $\sigma_V(t)$ yields $g \exp(tV) g^{-1} = \exp(tgVg^{-1})$. As for $\mathrm{Ad}_g$ we have $\mathrm{Ad}_g : V \mapsto gVg^{-1}$ since

$$\begin{aligned} \mathrm{Ad}_g V &= \left.\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{ad}_g \exp(tV)]\right|_{t=0} \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\exp(tgVg^{-1})\right|_{t=0} = gVg^{-1}. \end{aligned} \tag{5.163}$$

## Problems

**5.1** The Stiefel manifold $V(m, r)$ is the set of orthonormal vectors $\{\boldsymbol{e}_i\}$ $(1 \le i \le r)$ in $\mathbb{R}^m$ $(r \le m)$. We may express an element $A$ of $V(m, r)$ by an $m \times r$ matrix $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r)$. Show that $\mathrm{SO}(m)$ acts transitively on $V(m, r)$. Let

$$A_0 \equiv \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

be an element of $V(m, r)$. Show that the isotropy group of $A_0$ is $\mathrm{SO}(m-r)$. Verify that $V(m, r) = \mathrm{SO}(m)/\mathrm{SO}(m-r)$ and $\dim V(m, r) = [r(r-1)]/2 + r(m-r)$. [*Remark*: The Stiefel manifold is, in a sense, a generalization of a sphere. Observe that $V(m, 1) = S^{m-1}$.]

**5.2** Let $M$ be the Minkowski four-spacetime. Define the action of a linear operator $* : \Omega^r(M) \to \Omega^{4-r}(M)$ by

$$
\begin{array}{lll}
r = 0: & *1 = -\mathrm{d}x^0 \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3; & \\
r = 1: & *\mathrm{d}x^i = -\mathrm{d}x^j \wedge \mathrm{d}x^k \wedge \mathrm{d}x^0 & * \,\mathrm{d}x^0 = -\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3; \\
r = 2: & *\mathrm{d}x^i \wedge \mathrm{d}x^j = \mathrm{d}x^k \wedge \mathrm{d}x^0 & * \,\mathrm{d}x^i \wedge \mathrm{d}x^0 = -\mathrm{d}x^j \wedge \mathrm{d}x^k; \\
r = 3: & *\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 = -\mathrm{d}x^0 & * \,\mathrm{d}x^i \wedge \mathrm{d}x^j \wedge \mathrm{d}x^0 = -\mathrm{d}x^k; \\
r = 4: & *\mathrm{d}x^0 \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 = 1; &
\end{array}
$$

where $(i, j, k)$ is an even permutation of $(1, 2, 3)$. The vector potential $A$ and the electromagnetic tensor $F$ are defined as in example 5.11. $J = J_\mu \mathrm{d}x^\mu = \rho \mathrm{d}x^0 + j_k \,\mathrm{d}x^k$ is the current one-form.

(a) Write down the equation $\mathrm{d} * F = *J$ and verify that it reduces to two of the Maxwell equations $\nabla \cdot \boldsymbol{E} = \rho$ and $\nabla \times \boldsymbol{B} - \partial \boldsymbol{E}/\partial t = \boldsymbol{j}$.

(b) Show that the identity $0 = \mathrm{d}(\mathrm{d} * F) = \mathrm{d} * J$ reduces to the charge conservation equation

$$\partial_\mu J^\mu = \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0.$$

(c) Show that the Lorentz condition $\partial_\mu A^\mu = 0$ is expressed as $\mathrm{d} * A = 0$.