CHAPTER 3

Differentiable Manifolds, Tangent Spaces, and Vector Fields

This chapter touches mostly on the topics that are relevant to the later applications in this monograph. For other important topics in differential geometry, for instance fibre bundles, connections, Riemann metric, curvature, etc., the reader is referred to the literature in this field; see, e.g., Bishop and Crittenden (1964), or Greup, Halperin, and Vanstone (1972). For applications of differential geometry to statistical parameter spaces see Amari, Barndorff-Nielsen, Kass, Lauritzen, and Rao (1987).

3.1. Manifolds. The spaces and groups encountered in this monograph have more structure than merely being topological: they are manifolds. Loosely speaking, a manifold is a space that is locally Euclidean at each point. A trivial example is a Euclidean space itself. More interesting examples are curved subsets of Euclidean spaces. For instance, the parabola $x_2 = x_1^2$ is a one-dimensional manifold embedded in R^2, and the sphere $x_1^2 + x_2^2 + x_3^3 = 1$ is a two-dimensional manifold embedded in R^3. But the subset $\{(x_1, x_2) : x_1 x_2 = 0\}$ of R^2 is not a manifold because the point $(0,0)$ does not have a Euclidean neighborhood.

Formally, a d-dimensional manifold is a Hausdorff space M together with an assignment at every $p \in M$ of a neighborhood U_p of p and a function ϕ_p mapping U_p homeomorphically onto an open subset of R^d. It follows that if two neighborhoods U_p and U_q have nonempty intersection, then the function $\phi_{pq} \equiv \phi_q \circ \phi_p^{-1}$ is a homeomorphism of $\phi_p(U_p \cap U_q)$ onto $\phi_q(U_p \cap U_q)$.

For $p \in M$, the function ϕ_p assigns a point of R^d, given by its d coordinates, to every point $q \in U_p$. The choice of ϕ_p is sometimes called a **parametrization** of U_p, or at p. The pair (U_p, ϕ_p) is often called a **chart** at p, or a coordinate neighborhood of p (and a family of charts, one at each $p \in M$, an **atlas**). Whenever two charts overlap, the intersection receives two parametrizations which are continuous functions of each other. If for $q \in U_p$ the coordinates of $\phi_p(q)$ are $x_1, \ldots, x_d$, then we shall call these often **local coordinates** on a neighborhood of p, and if $x = (x_1, \ldots, x_d)$, then we shall often write $x(q)$ instead of $\phi_p(q)$. Although the greatest interest lies in manifolds of dimension $d \geq 1$, occasionally we have to deal with manifolds of dimension 0. These are spaces with the discrete topology (Section 2.2); for instance, a finite point set.

A richer theory of manifolds results from imposing more smoothness on the functions ϕ_{pq} and ϕ_{pq}^{-1} than mere continuity. If they are continuously differentiable of order k, $1 \leq k \leq \infty$, then M will be called a **differentiable manifold** of class C^k, or simply a C^k manifold. For our purpose the case $k = 1$ will suffice most of the time. Sometimes k has to be > 1. For instance, the notion of a bracket of two vector fields (Section 3.5) is not even defined unless $k \geq 2$. It is sometimes convenient to take $k = \infty$. Then some statements become simpler, for instance the definition of smoothness class of a vector field or of a differential form. Therefore, whenever convenient we shall feel free to assume $k = \infty$ while realizing that some finite value of k might suffice. This liberty will be taken in Sections 3.5, 3.6, and in Chapter 4. Actually, all our applications will be to C^∞ manifolds. It is possible to impose even more regularity than C^∞ differentiability and require the functions ϕ_{pq} and ϕ_{pq}^{-1} to be **analytic** for every $p, q \in M$;

i.e., they can be developed in convergent power series. Then M is called an **analytic manifold**. An analytic manifold is also C^∞, but the converse is false in general. In all applications in later chapters the spaces and groups will in fact be analytic, but this will not be used explicitly everywhere.

It is sometimes possible to put a single chart on the whole of M. In that case the functions ϕ_{pq} are the identity functions so that M together with the chosen chart is trivially an analytic manifold. The parabola $x_2 = x_1^2$ is of that nature (with global coordinate x_1). However, it is impossible to do this with the sphere $x_1^2 + x_2^2 + x_3^3 = 1$ (or with the circle $x_1^2 + x_2^2 = 1$) unless one point is removed.

So far a C^k manifold has been defined as a Hausdorff space together with a family of charts satisfying the requirement that the functions ϕ_{pq} are C^k. This is not quite right since the parametrizations furnished by the functions ϕ_p may be changed in a C^k way without changing M as a differentiable manifold. Thus, more precisely, M is defined as a differentiable manifold by an equivalence class of parametrizations, where two parametrizations are called equivalent if they are in C^k relation to each other. The same is true for an analytic manifold, with "C^k" replaced by "analytic." Thus, the same manifold (C^k or analytic) can always be parametrized in many different ways. For instance, on the real line R one can put the usual chart that assigns to the point x the coordinate x, or another chart that assigns to the point x the coordinate $\tanh x$. Since the function $y = \tanh x$ is analytic in both directions, these two parametrizations define the same analytic manifold. However, the chart that assigns to the point x the coordinate x^3 turns R into a different analytic manifold since the inverse of the function $y = x^3$ is not analytic (not even C^1). It will always be assumed in the following without special mention that if a chart is chosen at a point of M, then it is a chart belonging to the equivalence class of charts that defines M. Such a chart is called **admissible**. Any equivalence class of parametrizations is called a **differentiable structure** (or **analytic structure** as the case may be). The above example shows that the same set may receive different differentiable structures, producing different differentiable manifolds.

Let M and N be two C^k manifolds and f a function $M \to N$. Let $p \in M$, then we shall say that f is of class C^k at p if there is a chart (U_p, ϕ_p) at p and a chart (V_q, ψ_q) at $q = f(p)$ such that the function $\psi_q \circ f \circ \phi_p^{-1}$ on $\phi_p(U_p)$ into $\psi_q(V_q)$ is of class C^k. Expressed in words, in terms of local coordinates the function is C^k on a neighborhood of p. Clearly, this does not depend on the choice of admissible charts. We shall say that f is of class C^k if f is of class C^k at every point $p \in M$. An analogous definition holds with "C^k" replaced by "analytic." Important special cases are $f : M \to R$ and $f : R \to M$. A **curve** in the C^k manifold M is a C^k function γ on an interval of R (possibly the whole of R) into M. Similarly with "C^k" replaced by "analytic."

Jacobians and diffeomorphisms. Let M and N be two d-dimensional C^1 manifolds and f a C^1 function $M \to N$. If $x = (x_1, \dots, x_d)$ are local coordinates on a neighborhood U of $p \in M$ and similarly $y = (y_1, \dots, y_d)$ on $f(U)$ in N, then the Jacobian $\frac{\partial(y)}{\partial(x)}$ on U will be defined as

$$\frac{\partial(y)}{\partial(x)} = \operatorname{abs}\det\left(\left(\frac{\partial y_i}{\partial x_j}\right)\right), \tag{3.1.1}$$

i.e., the absolute value of the determinant of the matrix whose (i, j) element is $\frac{\partial y_i}{\partial x_j}$. The following inverse function theorem is a special case of the implicit function theorem.

3.1.1. THEOREM. *For $1 \le k \le \infty$ let M and N be C^k manifolds of the same dimension and let $f : M \to N$ be C^k. If at $p \in M$ f has a positive Jacobian, then there exists a neighborhood U of p such that f is 1-1 on U and $f^{-1} : f(U) \to U$ is C^k. If M, N, and f are analytic, then so is f^{-1}.*

PROOF. Dieudonné (1960), Theorem 10.2.5. □

If the C^k manifolds are of the same dimension and $f : M \to N$ a bijection, then f is called a C^k **diffeomorphism** (or simply a diffeomorphism) if f and f^{-1} are C^k. An **analytic diffeomorphism** is defined similarly. By Theorem 3.1.1, if f is bijective and C^k (resp. analytic), then M and N are diffeomorphic (resp. analytically diffeomorphic) if f has a positive Jacobian everywhere.

As a particular case take $N = M$. Then by Theorem 3.1.1 two admissible charts at a point $p \in M$ are related by a positive Jacobian, and, conversely, if the Jacobian is positive then one chart is admissible if and only if the other one is. In this form it is stated by Chevalley (1946), Chapter III, §1, Proposition 1.

3.2. Tangent vectors and spaces. First an example. Let the points of R^3 be denoted (x, y, z) and consider the sphere M whose equation is $x^2 + y^2 + (z - 1)^2 = 1$. The point $p = (0, 0, 0)$ lies on M, and for any real numbers a and b, not both 0, the line $\{(au, bu, 0) : -\infty < u < \infty\}$ is tangent to M at p. We also say that the vector $(a, b, 0)$ is a tangent vector of M at p. However, this elementary analytic-geometric notion does not extend very well to arbitrary differentiable manifolds. Instead, tangent vectors will be defined as directional derivatives. In the above example the "lower" half of the sphere is a neighborhood U_p of p that can be parametrized by the first two coordinates (x, y) of its points. A C^1 real valued function f on U_p can then be expressed as a C^1 function $f(x, y)$. Now let $\gamma(u) = (au, bu)$ be a curve in M with $|u| < u_0$, where u_0 is sufficiently small so that $\gamma(u)$ lies entirely in U_p. (Geometrically, the curve $\{\gamma(u) : -u_0 < u < u_0\}$ is part of a great circle through p.) The composition of γ and f is a real valued function on $(-u_0, u_0)$. Its derivative at $u = 0$ is

$$\frac{d}{du} f(\gamma(u))\bigg|_{u=0} = \frac{d}{du} f(au, bu)\bigg|_{u=0} = \left(a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y}\right) f(x, y)\bigg|_{x=y=0}.$$

The expression

$$t = a\,\frac{\partial}{\partial x} + b\frac{\partial}{\partial y}\bigg|_{x=y=0} \tag{3.2.1}$$

is called a tangent vector of M at p in the direction (a, b). This example motivates the formal definition of tangent vector given below.

Let M be a C^1 manifold. Given any $p \in M$ let $\mathcal{F}_p(M)$ be the family of real valued functions on M that are of class C^1 at p. If there is no danger of confusion we shall write $\mathcal{F}_p$ instead of $\mathcal{F}_p(M)$.

3.2.1. DEFINITION. *A function* $t : \mathcal{F}_p \to R$ *is called a* **tangent vector** *at* p *if*

(i) t *is linear:* $t(af + bg) = at(f) + bt(g)$ *for* $f, g \in \mathcal{F}_p$, $a, b \in R$;
(ii) t *is a derivation:* $t(fg) = f(p)t(g) + g(p)t(f)$ *for* $f, g \in \mathcal{F}_p$.

Any particular tangent vector t may be represented by a linear combination of partial derivatives, as in (3.2.1), by choosing a chart at p with local coordinates $x_1, \dots, x_d$, say. Then t is of the form

$$t = \sum_{i=1}^{d} a_i \frac{\partial}{\partial x_i}\bigg|_{x(p)}, \tag{3.2.2}$$

where the partial derivatives are to be evaluated at $x(p)$. The constants a_i will depend on the chosen chart. If $y_1, \dots, y_d$ are other local coordinates such that the x_i and y_j are C^1 function of each other, then t can also be expressed in the form

$$t = \sum_{i=1}^{d} b_j \frac{\partial}{\partial y_j}\bigg|_{y(p)} \tag{3.2.3}$$

and the b_j are function of the a_i, given by

$$b_j = \sum_{i=1}^{d} a_i \frac{\partial y_j}{\partial x_i}\bigg|_{x(p)}, \quad j = 1, \dots, d. \tag{3.2.4}$$

However, for any $f \in \mathcal{F}_p$, the value $t(f)$ does not depend on the choice of chart.

The sum of two tangent vectors at p, say t_1 and t_2, is defined in the obvious way: $(t_1 + t_2)(f) = t_1(f) + t_2(f)$ and is easily seen to satisfy Definition 3.2.1. Similarly a scalar multiple of a tangent vector. The **tangent space** at p, denoted M_p, is the vector space of all tangent vectors at p. In terms of a chosen chart at p with local coordinates $x_1, \dots, x_d$, a basis of M_p is

$$\left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_d}\right)\bigg|_{x(p)} . \tag{3.2.5}$$

Therefore, $\dim M_p = d$; i.e., M and M_p have the same dimension.

3.3. Differential of a mapping. Let M and N be C^1 manifolds, not necessarily of the same dimension, and f a C^1 function $M \to N$. Let $p \in M$, $q = f(p)$, and let M_p, N_q be the tangent spaces at p, q respectively. To each $t \in M_p$ there corresponds a tangent vector $u \in N_q$ as follows. For $g \in \mathcal{F}_q(N)$ the function $g \circ f = g^*$, say, is $\in \mathcal{F}_p(M)$. Then define $u(g) = t(g^*)$. This defines a function $M_p \to N_q$ which is easily seen to be linear and which is called the **differential** of f, denoted df. We can also express this definition by the formula

$$((df)(t))(g) = t(g \circ f), \quad t \in M_p,\ g \in \mathcal{F}_q(N). \tag{3.3.1}$$

In terms of local coordinates $x_1, \dots, x_d$ at $p \in M$ and $y_1, \dots, y_e$ at $q = f(p) \in N$ (where $e = \dim N$), and if t is given by (3.2.2), then

$$(df)(t) = \sum_{j=1}^{e} b_j \frac{\partial}{\partial y_j}\bigg|_{y(q)} , \tag{3.3.2}$$

in which the b_j are given by

$$b_j = \sum_{i=1}^{d} a_i \frac{\partial f_j}{\partial x_i}\bigg|_{x(p)} , \quad j = 1, \dots, e. \tag{3.3.3}$$

Now with the above parametrization let t be represented by the column vector a, $(df)(t)$ by b then it follows from (3.3.3) that $b = Aa$ in which the (i,j) element of the matrix A is $\partial f_i/\partial x_j|_{x(p)} = \partial y_i/\partial x_j|_{x(p)}$, where we have substituted y_i for $f_i(x)$. Now take the case where

$e = d$, then A is the Jacobian matrix on the right-hand side of (3.1.1) evaluated at $x(p)$. Therefore, A is invertible if and only if the Jacobian (3.1.1) evaluated at $x(p)$ is positive. On the other hand, the matrix A represents the linear map $df : M_p \to N_{f(p)}$ and is therefore invertible if and only if df is bijective, i.e., is a linear isomorphism. Thus, we have

3.3.1. THEOREM. *Theorem 3.1.1 is valid if the expression "f has a positive Jacobian" is replaced by "df is a linear isomorphism of M_p and $N_{f(p)}$."*

It follows from Theorems 3.1.1 and 3.3.1 that in order to show that a C^k function $f : M \to N$ is a C^k diffeomorphism (or an analytic diffeomorphism in the case of analytic M, N, f) it suffices to show that f is a bijection and that df is a linear isomorphism $M_p \to N_{f(p)}$ at each $p \in M$.

The concept of the differential df of a mapping f is so basic and useful that it may be worthwhile to express it in an informal way in order to get a better "feel" for it. Let $p \in M$ and take $p_1 \in M$ very close to p. Then define a functional t on functions $g : M \to R$ by $t(g) = \delta g \equiv g(p_1) - g(p)$. This t almost satisfies Definition 3.2.1 (with f, g there replaced by g, h, for notational reasons): t satisfies (i) and it satisfies (ii) approximately by neglecting the second order term $\delta g \delta h$. Within this approximation there is then a correspondence between points on M close to p and "small" tangent vectors at p. The same is true on N at $q = f(p)$. Then if t corresponds to p_1 close to p, its image under df is the small tangent vector at q that corresponds to $f(p_1)$ close to q. Extend to all tangent vectors by linearity. It may be of further help in the visualization process by thinking of M as a manifold embedded in some Euclidean space and picturing a point p_1 near p as a little arrow, say $\overrightarrow{pp}_1$, that runs from p to p_1; similarly $\overrightarrow{qq}_1$ for points $q_1 \in N$ close to q. Then df maps $\overrightarrow{pp}_1$ into $\overrightarrow{qq}_1$, where $q_1 = f(p_1)$.

Differential of a composition. Let L, M, N be C^1 manifolds and $f : L \to M$, $g : M \to N$ C^1 mappings. Let p be an arbitrary point of L and $q = g(f(p)) \in N$, then $d(g \circ f)$ is a linear map of L_p into N_q. From the definitions it follows immediately that $d(g \circ f) = dg \circ df$.

Differential of a real valued function. This turns out to be of special interest since it has two possible interpretations. There is on the real line R a single chart with coordinate y, say, and if q is an arbitrary point of R, then the tangent space R_q at q is a copy of R and is spanned by a single vector for which we may take $d/dy|_q$. Let M be a C^1 d-dimensional manifold and let $f : M \to R$ be C^1. Let $p \in M$, $f(p) = q$, and $t \in M_p$. Since $df(t) \in R_q$, we must have $df(t) = a(t)d/dy|_q$, with some constant $a(t)$ depending on t. It is easy to get an explicit expression for $a(t)$ by taking in (3.3.1) $g(y) = y$. Then the left-handed side of (3.3.1) equals $a(t)(d/dy)y = a(t)$ and the right-hand side is $t(f)$. Therefore, for real valued C^1 f we have

$$df(t) = t(f) \left.\frac{d}{dy}\right|_{f(p)}, \quad t \in M_p. \tag{3.3.4}$$

By (3.3.4), df associates to each $t \in M_p$ the real number $t(f)$ and this association is clearly linear. Thus, df may be regarded as a linear real valued function on M_p, i.e., a linear functional, according to the formula

$$df(t) = t(f), \quad t \in M_p. \tag{3.3.5}$$

We have now two interpretations of df evaluated at p: first, a linear function $M_p \to R_{f(p)}$; second, a linear functional on M_p. The latter interpretation can be applied, in particular, to the coordinate functions $x_1, \dots, x_d$ of a chart at $p \in M$. Then a basis of M_p can be chosen as (3.2.5). Each dx_i can be considered a linear functional on M_p. Taking in (3.3.5) $f = x_i$ and $t = \partial/\partial x_j|_{x(p)}$, we get

$$dx_i \left(\left.\frac{\partial}{\partial x_j}\right|_{x(p)} \right) = \delta_{ij}, \quad i, j = 1, \dots, d, \tag{3.3.6}$$

where $\delta_{ij} = 1$ or 0 according as $i = j$ or $i \neq j$ (Kronecker delta). For arbitrary C^1 real valued f we can then write

$$df = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i} dx_i \tag{3.3.7}$$

(where the partial derivatives are to be evaluated at $x(p)$) since by (3.3.5) and (3.3.6) the values at $t = \partial/\partial x_j$ of both sides equals $\partial f/\partial x_j$ evaluated at $x(p)$. If $y_1, \ldots, y_d$ is another admissible coordinate system at p, then by (3.3.7) we have

$$dy_i = \sum_{j=1}^{d} \frac{\partial y_i}{\partial x_j} dx_j, \quad i = 1, \ldots, d, \tag{3.3.8}$$

where the partial derivatives are to be evaluated at $x(p)$.

Frequently, the differential of a product of two or more real valued functions is needed. From (3.3.5) and Definition 3.2.1(ii), or from 3.3.7, it follows immediately that

$$d(fg) = f\,dg + g\,df. \tag{3.3.9}$$

The dual vector space to M_p, say M_p^*, is the space of all linear functionals on M_p. It follows from (3.3.6) that not only is $(dx_1, \ldots, dx_d)$ a basis of M_p^*, but it is the basis dual to (3.2.5).

3.4. Immersion, imbedding, submanifold. Let N and M be C^1 manifolds and $f : N \to M$ a C^1 mapping. Then f is called an **immersion** if df is 1-1 at every point of N (*note*: this does not imply that f is 1-1). For example, let $N = R$, $M = R^2$, and $f(u) = (x, y) = (\cos u, \sin u)$ for $u \in R$. Then $df(d/du) = -(\sin u)\partial/\partial x + (\cos u)\partial/\partial y$ which is never 0 so that df is 1-1 at every point. However, f maps R into the unit circle in R^2 and is not 1-1. But an immersion is locally 1-1. i.e., at each $p \in N$ there is a neighborhood U such that f is 1-1 on U (and f is approximately linear if U is small). Let d, e be the dimensions of M, N, respectively, and $x = (x_1, \ldots, x_e)$ a chart at $p \in N$, $y = (y_1, \ldots, y_d)$ a chart at $q = f(p) \in M$. Then f is an immersion if and only if the matrix $((\partial f_i/\partial x_j))$ is of rank e, and then we have necessarily $e \leq d$. It can be shown that if f is an immersion, then for any chart $y = (y_1, \ldots, y_d)$ there is a subset $y_{i_1}, \ldots, y_{i_e}$ such that $x = (x_1, \ldots, x_e)$ with $x_i = y_i \circ f$ $(i = 1, \ldots, e)$ forms a chart at p (Chevalley, 1946, III §IV, Proposition 1).

If f is 1-1 and an immersion, then f is called an **imbedding**. Thus, in the example above f is not an imbedding because f is many-to-one. An example of an imbedding of R into R^2 is $f(u) = (x, y) = (u, u^2)$, $-\infty < u < \infty$ (the parabola). A special case of an imbedding is $N \subset M$ and $f = i$, where $i : N \to M$ is the **inclusion map** $i(p) = p$, provided that i is an immersion. This is called a **submanifold**. Thus, N is a submanifold of M if $N \subset M$ and di is 1-1 everywhere. A rather trivial example is an open subset N of M if N inherits its differentiable structure from M. But an open subset with a different differentiable structure is no longer a submanifold. For instance, take $N = M = R$ in which N is parametrized by the variable x, M by y, and $i(p) = p$ is represented by $y = x^3$. Then $di(d/dx) = 2x^2(d/dy)$ so that $di = 0$ at $x = 0$ and therefore di is not 1-1 at $x = 0$.

A familiar example of a submanifold N of lower dimension than M is the one-dimensional straight line $ax + by = c$ in R^2 supplied with the usual differentiable structure. It is assumed here that not both a and b are 0. There are three cases to be distinguished: (i) $a = 0$, $b \neq 0$; (ii) $a \neq 0$, $b = 0$; and (iii) $a \neq 0$, $b \neq 0$. In case (i) N can be parametrized by x, in (ii) by y, and in (iii) by either. In all cases di is not $= 0$ anywhere which implies that $di(N_p)$ has dimension 1 for every $p \in N$; i.e., di is 1-1 everywhere. For instance, in case (iii) with parametrization x, $di(d/dx) = \partial/\partial x - (a/b)(\partial/\partial y)$. A similar situation prevails if N is the circle $x^2 + y^2 = 1$. In the points of N where $x = 0$ there is a chart with local coordinate x; similarly, in the neighborhoods of the points where $y = 0$ the parametrizations can be furnished by y; in all other points either x or y will do. In general, if N is a submanifold of M, with $\dim N = e \leq d = \dim M$, and at $p \in N$ (therefore $p \in M$) there is a chart in M with local coordinates

$x_1, \ldots, x_d$, then since i is an immersion it is possible to choose a subset $x_{i_1}, \ldots, x_{i_e}$ that form the local coordinates of an admissible chart in N. This is an equivalent criterion for $N \subset M$ to be a submanifold of M (Cohn, 1957, Section 1.9).

In the above examples of straight line and circle as lower dimensional submanifolds N of $R^2 = M$, the topology of N derived from its differentiable structure is the same as its relative topology as a subspace of M. Roughly speaking, points of N that are close in the topology of M are also close in the topology of N. This need not be the case in general if N is a submanifold of M. For instance, in the irrational flow on the torus M (Chapter 1) a single orbit N parametrized by a real variable is a one-dimensional submanifold of the two-dimensional M and has the topology of R as a manifold, but its relative topology as a subspace of M is quite different since the orbit keeps returning arbitrarily closely to any point of departure. A similar example can be given with $M = R^2$ and N as the union of all horizontal lines. Then N is a one-dimensional submanifold, and two points on different lines can be close together in the topology of M but are far apart in the topology of N.

Submersion. This concept will not be used in the monograph and is mentioned here only for completeness since it is closely related to immersion. If the C^1 function $f : N \to M$ is such that at every point $p \in N$, df maps N_p *onto* $M_{f(p)}$, then f is called a **submersion.** This can of course happen only if $\dim N \geq \dim M$. If f is both an immersion and a submersion, then df is a linear isomorphism at every point, so that f is locally a diffeomorphism by Theorems 3.1.1 and 3.3.1 (analytic if M, N, and f are analytic). If f is also 1-1, then f is a global C^1 (or analytic) diffeomorphism.

3.5. Vector fields, integral curves, and brackets. If M is a C^∞ manifold, then a **vector field** X is a function that assigns to each $p \in M$ an element of M_p, denoted $X(p)$. Let $f : M \to R$ be of class C^∞ and define $Xf : M \to R$ by $(Xf)(p) = X(p)f$ (henceforth we shall often omit parentheses and write, e.g., tf instead of $t(f)$ is t is a tangent vector). We shall say that X is of class C^∞ if Xf is C^∞

for every f of class C^∞. If desired, the domain of X may be restricted to an open subset of M. If the domain of X is covered by charts, then in each chart X can be expressed in the form $\sum_1^d f_i(x)\partial/\partial x_i$, where the f_i are C^∞ functions (which of course also depend on the chart). Conversely, if all these f_i are C^∞, then X is C^∞. If X is a C^∞ vector field and h a C^∞ function $M \to R$, then hX is a C^∞ vector field, where $(hX)(p) = h(p)X(p)$, $p \in M$.

Integral curve. Let X be a C^∞ vector field and γ a curve in M with domain the interval $(-a, b)$, $0 < a, b \leq \infty$, such that $\gamma(0) = p \in M$. Then γ is called an **integral curve** of X **starting at** p if $d\gamma(d/du) = X(\gamma(u))$ for every $-a < u < b$. This can also be expressed in a different way by using the definition (3.3.1) of the differential: in (3.3.1) replace f by γ and t by d/du, then for any C^∞ function $g : M \to R$ an integral curve γ satisfies

$$(3.5.1) \qquad \frac{d}{du}g(\gamma(u)) = X(\gamma(u))g, \quad -a < u < b.$$

By taking g successively the coordinate functions in a chart the equation (3.5.1) can be converted in to a set of differential equations. For instance, let there be a chart at p with local coordinates $x_1, \ldots, x_d$ and let $\gamma(u)$ (for u in a neighborhood of 0) be represented in the chart by $x(u)$ with coordinates $x_1(u), \ldots, x_d(u)$. Also, let X on the chart be represented by $X = \sum_i \alpha_i(x)\partial/\partial x_i$, with C^∞ functions α_i. Then by taking in (3.5.1) g to correspond to the coordinate function x_i we get

$$(3.5.2) \qquad \frac{d}{du}x_i(u) = \alpha_i(x(u)), \quad i = 1, \ldots, d.$$

A solution of (3.5.2) for u in a neighborhood of 0 provides an explicit expression for the integral curve locally. It follows from a theorem in ordinary differential equations that a unique C^∞ solution exists. If M and X are analytic, then so is the solution. Relevant references include: Dieudonné (1960), Theorems (10.4.5), (10.5.3); Birkhoff and Rota (1978), Chapter 6, Section 10, Corollary 2; Bieberbach (1965), §1 no. 6.

Bracket. Let X and Y be two C^∞ vector fields on M. With XY is meant the operator such that for any C^∞ real valued function f on M, $(XY)f = X(Yf)$. However, in general XY is not a vector field since, for $p \in M$, $t = (XY)(p)$ does not satisfy condition (ii) of Definition 3.2.1. (One can also see this by writing both X and Y in terms of the coordinates of a chart; then second order derivatives enter.) But $XY - YX$ does satisfy condition (ii) (in terms of local coordinates, the second order derivatives cancel). Define

$$[X,Y] = XY - YX; \tag{3.5.3}$$

this is called the **bracket** (or **commutator**) of X and Y, and is a C^∞ vector field if X and Y are. It follows immediately from the definition that $[X,Y] = -[Y,X]$, and that $[X,X] = 0$.

3.6. Transformation of vector fields under mappings. Invariant vector fields. Let M and N be C^∞ manifolds and f a C^∞ mapping $M \to N$. Let X be a C^∞ vector field on M (or on an open subset of M). Then $df\, X$ is a C^∞ vector field on a subset of N, whose value at $f(p) \in N$ is given by the definition (3.3.1) of df by taking in that formula $t = X(p)$, for every $p \in M$ where X is defined. We may rewrite (3.3.1) by replacing t by X except that then on the left-hand side we have a real valued function on N, whereas on the right-hand side the function is defined on M. This can be remedied by composing the left-hand side with f. Thus, the definition of $df\, X$ becomes

$$((df\, X)g) \circ f = X(g \circ f), \quad g \in C^\infty(N). \tag{3.6.1}$$

Now let X and Y be two C^∞ vector fields on M. We shall show that the bracket operation has the important property that it commutes with the differential df:

$$df[X,Y] = [df\, X, df\, Y]. \tag{3.6.2}$$

Put $U = df\, X$, $V = df\, Y$. Then (3.6.1) reads $(Ug) \circ f = X(g \circ f)$. Replace g by $Vg : (UVg) \circ f = X((Vg) \circ f)$. But $(Vg) \circ f = Y(g \circ f)$

by (3.6.1) with X replaced by Y. So $(UVg) \circ f = XY(g \circ f)$. Reverse the order of X and Y and subtract: $([U,V]g) \circ f = [X,Y](g \circ f)$. The right-hand side of this equation can be replaced by the left-hand side of (3.6.1) if $[X,Y]$ is substituted for X. This yields $([U,V]g) \circ f = (df[X,Y]g) \circ f$, for arbitrary C^∞ function g on N. It follows that $[U,V] = df[X,Y]$, which is (3.6.2).

An important special case of mapping arises when $M = N$ and f is a diffeomorphism of M with itself. Suppose X is defined on the whole of M, then the same is true of $df\,X$. We shall say that X is **invariant** under f if $df\,X = X$. Equation (3.6.2) shows that if X and Y are both invariant, then so is their bracket. Now suppose there is a group G acting on the left of M. The action (or left translation) of $g \in G$ on $p \in M$ was denoted $p \to gp$ in Chapter 2, but here and in Chapter 5 it is more convenient to denote left translation by L_g. Assume that $L_g : M \to M$ is a C^∞ diffeomorphism for every $g \in G$. A C^∞ vector field X on M is said to be **invariant** under G if $dL_g X = X$ for every $g \in G$. The property of being invariant under G is obviously preserved under linear operations so that the invariant vector fields (under G) form a linear space. Denote this space by $\mathfrak{m}$. Furthermore, if X and Y are invariant, then so is $[X,Y]$. Hence, $\mathfrak{m}$ is closed both under linear operations and under the formation of brackets. Let $\dim M = d$. If G is transitive over M, then an invariant vector field X is determined by its value at any given point $p_0 \in M$, for if $X(p_0) = t$, and $p = gp_0$, then $X(p) = dL_g t$. Since $t \in M_{p_0}$ and $\dim M_{p_0} = d$, it follows that $\dim \mathfrak{m} \leq d$. If G acts not only transitively but also freely, then every $t \in M_{p_0}$ generates an invariant X by the formula $X(p) = dL_g t$ if $p = gp_0$ (observe that g here is unique). It follows that then $\dim \mathfrak{m} = d$. This is the case in Chapter 5 when G is a Lie group acting on itself.