wherever defined. Note that this local group action has no global counterpart on $\mathbb{R}^2$; indeed $|\Psi(\varepsilon, (x, y))| \to \infty$ as $\varepsilon \to 1/x$ for $x \neq 0$. The orbits of the action consist of the straight rays emanating from the origin, and the origin itself. The action is regular on the punctured plane $\mathbb{R}^2 \backslash \{0\}$.

(d) The "*irrational flow*" on the torus: Let $G = \mathbb{R}$ and M be the two-dimensional torus T^2. Let ω be a fixed real number. Using the angular coordinates (θ, ρ) on T^2 we define a global group action

$$\Psi(\varepsilon, (\theta, \rho)) = (\theta + \varepsilon, \rho + \omega\varepsilon) \mod 2\pi.$$

The orbits of G are easily seen to all be one-dimensional submanifolds of T^2, so G acts semi-regularly in all cases. If ω is a rational number, the orbits are closed curves, and the action is regular. On the other hand, if ω is irrational, each orbit is a dense submanifold of T^2. This is the simplest example of a semi-regular group action which is not regular.

1.3. Vector Fields

The main tool in the theory of Lie groups and transformation groups is the "infinitesimal transformation". In order to present this, we need first to develop the concept of a vector field on a manifold. We begin with a discussion of tangent vectors. Suppose C is a smooth curve on a manifold M, parametrized by $\phi: I \to M$, where I is a subinterval of $\mathbb{R}$. In local coordinates $x = (x^1, \ldots, x^m)$, C is given by m smooth functions $\phi(\varepsilon) = (\phi^1(\varepsilon), \ldots, \phi^m(\varepsilon))$ of the real variable ε. At each point $x = \phi(\varepsilon)$ of C the curve has a *tangent vector*, namely the derivative $\dot\phi(\varepsilon) = d\phi/d\varepsilon = (\dot\phi^1(\varepsilon), \ldots, \dot\phi^m(\varepsilon))$. In order to distinguish between tangent vectors and local coordinate expressions for points on the manifold, we adopt the notation

$$\mathbf{v}|_x = \dot\phi(\varepsilon) = \dot\phi^1(\varepsilon)\frac{\partial}{\partial x^1} + \dot\phi^2(\varepsilon)\frac{\partial}{\partial x^2} + \cdots + \dot\phi^m(\varepsilon)\frac{\partial}{\partial x^m} \tag{1.4}$$

for the tangent vector to C at $x = \phi(\varepsilon)$. On first encounter, this notation may look rather strange, but its usefulness and naturalness will be amply demonstrated throughout this book. For the moment, the reader can view the symbols $\partial/\partial x^i$ just as "place holders" for the components $\dot\phi^i(\varepsilon)$ of the tangent vector $\mathbf{v}|_x$, or, equivalently, as a special "basis" of tangent vectors corresponding to the coordinate curves whose local coordinate expressions are $x + \varepsilon e_i$, e_i being the i-th basis vector of $\mathbb{R}^m$. Later we will see how each $\partial/\partial x^i$ does indeed correspond to a partial differential operator.

For example, the helix

$$\phi(\varepsilon) = (\cos \varepsilon, \sin \varepsilon, \varepsilon)$$

in $\mathbb{R}^3$, with coordinates (x, y, z), has tangent vector

$$\dot\phi(\varepsilon) = -\sin \varepsilon\frac{\partial}{\partial x} + \cos \varepsilon\frac{\partial}{\partial y} + \frac{\partial}{\partial z} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} + \frac{\partial}{\partial z}$$

at the point $(x, y, z) = \phi(\varepsilon) = (\cos \varepsilon, \sin \varepsilon, \varepsilon)$.

Two curves $C = \{\phi(\varepsilon)\}$ and $\tilde{C} = \{\tilde{\phi}(\theta)\}$ passing through the same point

$$x = \phi(\varepsilon^*) = \tilde{\phi}(\theta^*)$$

for some ε^*, θ^*, have the same tangent vector if and only if their derivatives agree at the point:

$$\frac{d\phi}{d\varepsilon}(\varepsilon^*) = \frac{d\tilde{\phi}}{d\theta}(\theta^*). \tag{1.5}$$

It is not difficult to see that this concept is independent of the local coordinate system used near x. Indeed, if $x = \phi(\varepsilon) = (\phi^1(\varepsilon), \ldots, \phi^m(\varepsilon))$ is the local coordinate expression in terms of $x = (x^1, \ldots, x^m)$ and $y = \psi(x)$ is any diffeomorphism, then $y = \psi(\phi(\varepsilon))$ is the local coordinate formula for the curve in terms of the y-coordinates. The tangent vector $\mathbf{v}|_x = \dot{\phi}(\varepsilon)$, which has the formula (1.4) in the x-coordinates, takes the form

$$\mathbf{v}|_{y=\psi(x)} = \sum_{j=1}^{m} \frac{d}{d\varepsilon}\psi^j(\phi(\varepsilon))\frac{\partial}{\partial y^j} = \sum_{j=1}^{m}\sum_{k=1}^{m} \frac{\partial\psi^j}{\partial x^k}(\phi(\varepsilon))\frac{d\phi^k}{d\varepsilon}\frac{\partial}{\partial y^j} \tag{1.6}$$

in the y-coordinates. Since the Jacobian matrix $\partial\psi^j/\partial x^k$ is invertible at each point, (1.5) holds if and only if

$$\frac{d}{d\varepsilon}\psi(\phi(\varepsilon^*)) = \frac{d}{d\theta}\psi(\tilde{\phi}(\theta^*)),$$

which proves the claim. Note that (1.6) tells how a tangent vector (1.4) behaves under the given change of coordinates $y = \psi(x)$.

The collection of all tangent vectors to all possible curves passing through a given point x in M is called the *tangent space* to M at x, and is denoted by $TM|_x$. If M is an m-dimensional manifold, then $TM|_x$ is an m-dimensional vector space, with $\{\partial/\partial x^1, \ldots, \partial/\partial x^m\}$ providing a basis for $TM|_x$ in the given local coordinates. The collection of all tangent spaces corresponding to all points x in M is called the *tangent bundle* of M, denoted by

$$TM = \bigcup_{x \in M} TM|_x.$$

These tangent spaces are "glued" together in an obvious smooth fashion, so that if $\phi(\varepsilon)$ is any smooth curve then the tangent vectors $\dot{\phi}(\varepsilon) \in TM|_{\phi(\varepsilon)}$ will vary smoothly from point to point. This makes the tangent bundle TM into a smooth manifold of dimension $2m$.

For example, if $M = \mathbb{R}^m$, then we can identify the tangent space $T\mathbb{R}^m|_x$ at any $x \in \mathbb{R}^m$ with $\mathbb{R}^m$ itself. This stems from the fact that the tangent vector $\dot{\phi}(\varepsilon)$ to a smooth curve $\phi(\varepsilon)$ can be realized as an actual vector in $\mathbb{R}^m$, namely $(\dot{\phi}^1(\varepsilon), \ldots, \dot{\phi}^m(\varepsilon))$. Another way of looking at this identification is that we are identifying the basis vector $\partial/\partial x^i$ of $T\mathbb{R}^m|_x$ with the standard basis vector e_i of $\mathbb{R}^m$. The tangent bundle of $\mathbb{R}^m$ is thus a Cartesian product $T\mathbb{R}^m \simeq \mathbb{R}^m \times \mathbb{R}^m$. If S is a surface in $\mathbb{R}^3$, then the tangent space $TS|_x$ can be identified with the usual geometric tangent plane to S at each point $x \in S$. This again uses the identification $T\mathbb{R}^3|_x \simeq \mathbb{R}^3$, so $TS|_x \subset T\mathbb{R}^3|_x$ is a plane in $\mathbb{R}^3$.

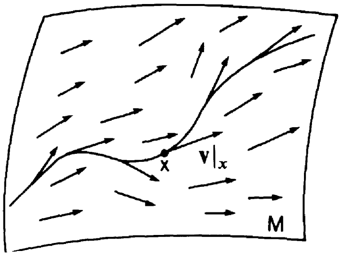

Figure 4. Vector field and integral curve on a manifold.

A *vector field* $\mathbf{v}$ on M assigns a tangent vector $\mathbf{v}|_x \in TM|_x$ to each point $x \in M$, with $\mathbf{v}|_x$ varying smoothly from point to point. In local coordinates $(x^1, \ldots, x^m)$, a vector field has the form

$$\mathbf{v}|_x = \xi^1(x)\frac{\partial}{\partial x^1} + \xi^2(x)\frac{\partial}{\partial x^2} + \cdots + \xi^m(x)\frac{\partial}{\partial x^m},$$

where each $\xi^i(x)$ is a smooth function of x. (Technically, we should put the symbol $|_x$ on each $\partial/\partial x^i$ to indicate in which tangent space $TM|_x$ it lies, but this should be clear from the context.) A good physical example of a vector field is the velocity field of a steady fluid flow in some open subset $M \subset \mathbb{R}^3$. At each point $(x, y, z) \in M$, the vector $\mathbf{v}|_{(x,y,z)}$ would be the velocity of the fluid particles passing through the point (x, y, z).

An *integral curve* of a vector field $\mathbf{v}$ is a smooth parametrized curve $x = \phi(\varepsilon)$ whose tangent vector at any point coincides with the value of $\mathbf{v}$ at the same point:

$$\dot{\phi}(\varepsilon) = \mathbf{v}|_{\phi(\varepsilon)}$$

for all ε. In local coordinates, $x = \phi(\varepsilon) = (\phi^1(\varepsilon), \ldots, \phi^m(\varepsilon))$ must be a solution to the autonomous system of ordinary differential equations

$$\frac{dx^i}{d\varepsilon} = \xi^i(x), \qquad i = 1, \ldots, m, \tag{1.7}$$

where the $\xi^i(x)$ are the coefficients of $\mathbf{v}$ at x. For $\xi^i(x)$ smooth, the standard existence and uniqueness theorems for systems of ordinary differential equations guarantee that there is a unique solution to (1.7) for each set of initial data

$$\phi(0) = x_0. \tag{1.8}$$

This in turn implies the existence of a unique *maximal* integral curve $\phi: I \to M$ passing through a given point $x_0 = \phi(0) \in M$, where "maximal" means that it is not contained in any longer integral curve; i.e. if $\tilde{\phi}: \tilde{I} \to M$ is any other integral curve with the same initial value $\tilde{\phi}(0) = x_0$, then $\tilde{I} \subset I$ and

$\tilde{\phi}(\varepsilon) = \phi(\varepsilon)$ for $\varepsilon \in \tilde{I}$. Note that if $\mathbf{v}|_{x_0} = 0$, then the integral curve through x_0 is just the point $\phi(\varepsilon) \equiv x_0$ itself, defined for all ε.

We note that if $\mathbf{v}$ is any smooth vector field on a manifold M, and $f(x)$ is any smooth, real-valued function defined for $x \in M$, then $f \cdot \mathbf{v}$ is again a smooth vector field, with $(f \cdot \mathbf{v})|_x = f(x)\mathbf{v}|_x$. In local coordinates, if $\mathbf{v} = \sum \xi^i(x)\partial/\partial x^i$, then $f \cdot \mathbf{v} = \sum f(x)\xi^i(x)\partial/\partial x^i$. If f never vanishes, the integral curves of $f \cdot \mathbf{v}$ coincide with the integral curves of $\mathbf{v}$, but the parametrizations will differ. For instance, the integral curves for $2\mathbf{v}$ will be traversed twice as fast as those of $\mathbf{v}$, but otherwise will be the same subsets of M.

Flows

If $\mathbf{v}$ is a vector field, we denote the parametrized maximal integral curve passing through x in M by $\Psi(\varepsilon, x)$ and call Ψ the *flow* generated by $\mathbf{v}$. Thus for each $x \in M$, and ε in some interval I_x containing 0, $\Psi(\varepsilon, x)$ will be a point on the integral curve passing through x in M. The flow of a vector field has the basic properties:

$$\Psi(\delta, \Psi(\varepsilon, x)) = \Psi(\delta + \varepsilon, x), \qquad x \in M, \tag{1.9}$$

for all $\delta, \varepsilon \in \mathbb{R}$ such that both sides of equation are defined,

$$\Psi(0, x) = x, \tag{1.10}$$

and

$$\frac{d}{d\varepsilon}\Psi(\varepsilon, x) = \mathbf{v}|_{\Psi(\varepsilon, x)} \tag{1.11}$$

for all ε where defined. Here (1.11) simply states that $\mathbf{v}$ is tangent to the curve $\Psi(\varepsilon, x)$ for fixed x, and (1.10) gives the initial conditions for this integral curve. The proof of (1.9) follows easily from the uniqueness of solutions to systems of ordinary differential equations; namely as functions of δ both sides of (1.9) satisfy (1.7) and have the same initial conditions at $\delta = 0$. If $\mathbf{v}$ is the velocity vector field of some steady state fluid flow, the integral curves of $\mathbf{v}$ are the stream lines followed by the fluid particles, and the flow $\Psi(\varepsilon, x)$ tells the position of a particle at time ε which started out at position x at time $\varepsilon = 0$.

Comparing the first two properties (1.9), (1.10) with (1.1), (1.2), we see that the flow generated by a vector field is the same as a local group action of the Lie group $\mathbb{R}$ on the manifold M, often called a *one-parameter group of transformations*. The vector field $\mathbf{v}$ is called the *infinitesimal generator* of the action since by Taylor's theorem, in local coordinates

$$\Psi(\varepsilon, x) = x + \varepsilon\xi(x) + O(\varepsilon^2),$$

where $\xi = (\xi^1, \ldots, \xi^m)$ are the coefficients of $\mathbf{v}$. The orbits of the one-parameter group action are the maximal integral curves of the vector field $\mathbf{v}$. Conversely, if $\Psi(\varepsilon, x)$ is any one-parameter group of transformations acting on

M, then its infinitesimal generator is obtained by specializing (1.11) at $\varepsilon = 0$:

$$\mathbf{v}|_x = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \Psi(\varepsilon, x). \tag{1.12}$$

Uniqueness of solutions to (1.7), (1.8) guarantees that the flow generated by $\mathbf{v}$ coincides with the given local action of $\mathbb{R}$ on M on the common domain of definition. Thus there is a one-to-one correspondence between local one-parameter groups of transformations and their infinitesimal generators.

The computation of the flow or one-parameter group generated by a given vector field $\mathbf{v}$ (in other words, solving the system of ordinary differential equations) is often referred to as *exponentiation* of the vector field. The suggestive notation

$$\exp(\varepsilon\mathbf{v})x \equiv \Psi(\varepsilon, x)$$

for this flow will be adopted in this book. In terms of this exponential notation, the above three properties can be restated as

$$\exp[(\delta + \varepsilon)\mathbf{v}]x = \exp(\delta\mathbf{v}) \exp(\varepsilon\mathbf{v})x \tag{1.13}$$

whenever defined,

$$\exp(0\mathbf{v})x = x, \tag{1.14}$$

and

$$\frac{d}{d\varepsilon}[\exp(\varepsilon\mathbf{v})x] = \mathbf{v}|_{\exp(\varepsilon\mathbf{v})x}. \tag{1.15}$$

for all $x \in M$. (In particular, $\mathbf{v}|_x$ is obtained by evaluating (1.15) at $\varepsilon = 0$.) These properties mirror the properties of the usual exponential function, justifying the notation.

Example 1.28. *Examples of Vector Fields and Flows.*

(a) Let $M = \mathbb{R}$ with coordinate x, and consider the vector field $\mathbf{v} = \partial/\partial x \equiv \partial_x$. (In the sequel, we will often use the notation ∂_x for $\partial/\partial x$ to save space.) Then

$$\exp(\varepsilon\mathbf{v})x = \exp(\varepsilon\partial_x)x = x + \varepsilon,$$

which is globally defined. For the vector field $x\partial_x$ we recover the usual exponential

$$\exp(\varepsilon x\partial_x)x = e^\varepsilon x,$$

since it must be the solution to the ordinary differential equation $\dot{x} = x$ with initial value x at $\varepsilon = 0$.

(b) In the case of $\mathbb{R}^m$, a constant vector field $\mathbf{v}_a = \sum a^i\partial/\partial x^i$, $a = (a^1, \ldots, a^m)$ exponentiates to the group of translations

$$\exp(\varepsilon\mathbf{v}_a)x = x + \varepsilon a, \qquad x \in \mathbb{R}^m,$$

in the direction a. Similarly, a linear vector field

$$\mathbf{v}_A = \sum_{i=1}^{m} \left(\sum_{j=1}^{m} a_{ij} x^j \right) \frac{\partial}{\partial x^i},$$

where $A = (a_{ij})$ is an $m \times m$ matrix of constants, has flow

$$\exp(\varepsilon \mathbf{v}_A) x = e^{\varepsilon A} x,$$

where $e^{\varepsilon A} = \mathrm{I} + \varepsilon A + \frac{1}{2}\varepsilon^2 A^2 + \cdots$ is the usual matrix exponential.

(c) Consider the group of rotations in the plane

$$\Psi(\varepsilon, (x, y)) = (x \cos \varepsilon - y \sin \varepsilon, x \sin \varepsilon + y \cos \varepsilon).$$

Its infinitesimal generator is a vector field $\mathbf{v} = \xi(x, y)\partial_x + \eta(x, y)\partial_y$, where, according to (1.12),

$$\xi(x, y) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} (x \cos \varepsilon - y \sin \varepsilon) = -y,$$

$$\eta(x, y) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} (x \sin \varepsilon + y \cos \varepsilon) = x.$$

Thus $\mathbf{v} = -y\partial_x + x\partial_y$ is the infinitesimal generator, and indeed, the above group transformations agree with the solutions to the system of ordinary differential equations

$$dx/d\varepsilon = -y, \qquad dy/d\varepsilon = x.$$

(d) Finally, consider the local group action

$$\Psi(\varepsilon, (x, y)) = \left(\frac{x}{1 - \varepsilon x}, \frac{y}{1 - \varepsilon x} \right)$$

introduced in Example 1.27(c). Differentiating, as before, we find the infinitesimal generator to be $\mathbf{v} = x^2\partial_x + xy\partial_y$. This demonstrates that a smooth vector field may still generate only a local group action.

The effect of a change of coordinates $y = \psi(x)$ on a vector field $\mathbf{v}$ is determined by its effect on each individual tangent vector $\mathbf{v}|_x$, $x \in M$, as given by (1.6). Thus if $\mathbf{v}$ is a vector field whose expression in the x-coordinates is

$$\mathbf{v} = \sum_{i=1}^{m} \xi^i(x) \frac{\partial}{\partial x^i},$$

and $y = \psi(x)$ is a change of coordinates, then $\mathbf{v}$ has the formula

$$\mathbf{v} = \sum_{j=1}^{m} \sum_{i=1}^{m} \xi^i(\psi^{-1}(y)) \frac{\partial \psi^j}{\partial x^i}(\psi^{-1}(y)) \frac{\partial}{\partial y^j} \tag{1.16}$$

in the y-coordinates.

The next result illustrates our earlier remarks that by suitably choosing local coordinates, we can often simplify the expressions for objects on manifolds, in this case vector fields.

Proposition 1.29. *Suppose* $\mathbf{v}$ *is a vector field not vanishing at a point* $x_0 \in M$: $\mathbf{v}|_{x_0} \neq 0$. *Then there is a local coordinate chart* $y = (y^1, \ldots, y^m)$ *at* x_0 *such that in terms of these coordinates,* $\mathbf{v} = \partial/\partial y^1$.

Proof. First linearly change coordinates so that $x_0 = 0$ and $\mathbf{v}|_{x_0} = \partial/\partial x^1$. By continuity the coefficient $\xi^1(x)$ of $\partial/\partial x^1$ is positive in a neighbourhood of x_0. Since $\xi^1(x) > 0$, the integral curves of $\mathbf{v}$ cross the hyperplane $\{(0, x^2, \ldots, x^m)\}$ transversally, and hence in a neighbourhood of $x_0 = 0$, each point $x = (x^1, \ldots, x^m)$ can be defined uniquely as the flow of some point $(0, y^2, \ldots, y^m)$ on this hyperplane. Consequently

$$x = \exp(y^1\mathbf{v})(0, y^2, \ldots, y^m),$$

for y^1 near 0, gives a diffeomorphism from $(x^1, \ldots, x^m)$ to $(y^1, \ldots, y^m)$ which defines the y-coordinates. (Geometrically, we have "straightened out" the integral curves passing through the hyperplane perpendicular to the x^1-axis.) In terms of the y-coordinates, we have by (1.13), for small ε,

$$\exp(\varepsilon\mathbf{v})(y^1, \ldots, y^m) = (y^1 + \varepsilon, y^2, \ldots, y^m),$$

so the flow is just translation in the y^1-direction. Thus every nonvanishing vector field is locally equivalent to the infinitesimal generator of a group of translations. (Of course, the global picture can be very complicated, as the irrational flow on the torus makes clear.) □

Action on Functions

Let $\mathbf{v}$ be a vector field on M and $f: M \to \mathbb{R}$ a smooth function. We are interested in seeing how f changes under the flow generated by $\mathbf{v}$, meaning we look at $f(\exp(\varepsilon\mathbf{v})x)$ as ε varies. In local coordinates, if $\mathbf{v} = \sum \xi^i(x)\partial/\partial x^i$, then using the chain rule and (1.15) we find

$$\begin{aligned}\frac{d}{d\varepsilon} f(\exp(\varepsilon\mathbf{v})x) &= \sum_{i=1}^{m} \xi^i(\exp(\varepsilon\mathbf{v})x)\frac{\partial f}{\partial x^i}(\exp(\varepsilon\mathbf{v})x) \\ &\equiv \mathbf{v}(f)[\exp(\varepsilon\mathbf{v})x].\end{aligned} \tag{1.17}$$

In particular, at $\varepsilon = 0$,

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} f(\exp(\varepsilon\mathbf{v})x) = \sum_{i=1}^{m} \xi^i(x)\frac{\partial f}{\partial x^i}(x) = \mathbf{v}(f)(x).$$

Now the reason underlying our notation for vector fields becomes apparent: the vector field $\mathbf{v}$ acts as a first order partial differential operator on real-valued functions $f(x)$ on M. Furthermore, by Taylor's theorem,

$$f(\exp(\varepsilon\mathbf{v})x) = f(x) + \varepsilon\mathbf{v}(f)(x) + \mathrm{O}(\varepsilon^2),$$

so $\mathbf{v}(f)$ gives the *infinitesimal change* in the function f under the flow generated by $\mathbf{v}$. We can continue the process of differentiation and substitution into the Taylor series, obtaining

$$f(\exp(\varepsilon\mathbf{v})x) = f(x) + \varepsilon\mathbf{v}(f)(x) + \frac{\varepsilon^2}{2}\mathbf{v}^2(f)(x) + \cdots + \frac{\varepsilon^k}{k!}\mathbf{v}^k(f)(x) + \mathrm{O}(\varepsilon^{k+1}),$$

where $\mathbf{v}^2(f) = \mathbf{v}(\mathbf{v}(f))$, $\mathbf{v}^3(f) = \mathbf{v}(\mathbf{v}^2(f))$, etc. If we assume convergence of the entire Taylor series in ε, then we obtain the *Lie series*

$$f(\exp(\varepsilon\mathbf{v})x) = \sum_{k=0}^{\infty} \frac{\varepsilon^k}{k!}\mathbf{v}^k(f)(x) \tag{1.18}$$

for the action of the flow on f. The same result holds for vector-valued functions $F\colon M \to \mathbb{R}^n$, $F(x) = (F^1(x), \ldots, F^n(x))$, where we let $\mathbf{v}$ act component-wise on F: $\mathbf{v}(F) = (\mathbf{v}(F^1), \ldots, \mathbf{v}(F^n))$. In particular, if we let F be the coordinate functions x, we obtain (again under assumptions of convergence) a Lie series for the flow itself, given by

$$\exp(\varepsilon\mathbf{v})x = x + \varepsilon\xi(x) + \frac{\varepsilon^2}{2}\mathbf{v}(\xi)(x) + \cdots = \sum_{k=0}^{\infty} \frac{\varepsilon^k}{k!}\mathbf{v}^k(x), \tag{1.19}$$

where $\xi = (\xi^1, \ldots, \xi^m)$, $\mathbf{v}(\xi) = (\mathbf{v}(\xi^1), \ldots, \mathbf{v}(\xi^m))$, etc., providing even further justification for our exponential notation.

According to our new interpretation of the symbols $\partial/\partial x^i$, each tangent vector $\mathbf{v}|_x$ at a point x defines a *derivation* on the space of smooth real-valued functions f defined near x in M. This means that $\mathbf{v}|_x$, when applied to a smooth function f, gives a real number $\mathbf{v}(f) = \mathbf{v}(f)(x)$, and, moreover, this operation determined by $\mathbf{v}$ has the basic derivational properties of

(a) *Linearity*

$$\mathbf{v}(f + g) = \mathbf{v}(f) + \mathbf{v}(g), \tag{1.20}$$

(b) *Leibniz' Rule*

$$\mathbf{v}(f \cdot g) = \mathbf{v}(f) \cdot g + f \cdot \mathbf{v}(g). \tag{1.21}$$

(Here both sides of (1.20), (1.21) are evaluated at the point x.) Conversely, it is not hard to show that every derivation on the space of smooth functions at x is a tangent vector, and in particular is given in local coordinates by $\sum \xi^i \partial/\partial x^i$. (See Exercise 1.12.) This approach is often used to *define* tangent vectors and the tangent bundle in an abstract, coordinate-free manner. Further, if $\mathbf{v}$ is a vector field on M, then $\mathbf{v}(f)$ is a smooth function for any $f\colon M \to \mathbb{R}$. Thus we can also define vector fields as derivations, i.e. maps satisfying (1.20), (1.21), on the space of smooth functions on M. This point of view is especially useful for defining various operations on vector fields in a coordinate-free manner. (See Warner, [1; Chap. 1] for more details of this construction and the correspondence between tangent vectors and derivations.)

Differentials

Let M and N be smooth manifolds and $F: M \to N$ a smooth map between them. Each parametrized curve $C = \{\phi(\varepsilon): \varepsilon \in I\}$ on M is mapped by F to a parametrized curve $\tilde{C} = F(C) = \{\tilde{\phi}(\varepsilon) = F(\phi(\varepsilon)): \varepsilon \in I\}$ on N. Thus F induces a map from the tangent vector $d\phi/d\varepsilon$ to C at $x = \phi(\varepsilon)$ to the corresponding tangent vector $d\tilde{\phi}/d\varepsilon$ to $\tilde{C}$ at the image point $F(x) = F(\phi(\varepsilon)) = \tilde{\phi}(\varepsilon)$. This induced map is called the *differential* of F, and denoted by

$$dF(\dot{\phi}(\varepsilon)) = \frac{d}{d\varepsilon}\{F(\phi(\varepsilon))\}. \tag{1.22}$$

As every tangent vector $\mathbf{v}|_x \in TM|_x$ is tangent to some curve passing through x, the differential maps the tangent space to M at x to the tangent space to N at $F(x)$:

$$dF: TM|_x \to TN|_{F(x)}.$$

The local coordinate formula for the differential is found using the chain rule in the same manner as the change of variables formula (1.6). If

$$\mathbf{v}|_x = \sum_{i=1}^{m} \xi^i \frac{\partial}{\partial x^i}$$

is a tangent vector at $x \in M$, then

$$dF(\mathbf{v}|_x) = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} \xi^i \frac{\partial F^j}{\partial x^i}(x) \right) \frac{\partial}{\partial y^j} = \sum_{j=1}^{n} \mathbf{v}(F^j(x)) \frac{\partial}{\partial y^j}. \tag{1.23}$$

Note that the differential $dF|_x$ is a linear map from $TM|_x$ to $TN|_{F(x)}$, whose matrix expression in local coordinates is just the Jacobian matrix of F at x.

If we prefer to think of tangent vectors as derivations on the space of smooth functions defined near a point x, then the differential dF has the alternative definition

$$dF(\mathbf{v}|_x)f(y) = \mathbf{v}(f \circ F)(x), \qquad y = F(x), \tag{1.24}$$

for all $\mathbf{v}|_x \in TM|_x$ and all smooth $f: N \to \mathbb{R}$. The equivalence of (1.22) and (1.24) is easily verified using local coordinates.

Example 1.30. Let $M = \mathbb{R}^2$, with coordinates (x, y), and $N = \mathbb{R}$, with coordinate s, and let $F: \mathbb{R}^2 \to \mathbb{R}$ be any map $s = F(x, y)$. Given

$$\mathbf{v}|_{(x,y)} = a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y},$$

then, by (1.23),

$$dF(\mathbf{v}|_{(x,y)}) = \left\{ a\frac{\partial F}{\partial x}(x, y) + b\frac{\partial F}{\partial y}(x, y) \right\} \frac{d}{ds}\bigg|_{F(x,y)}.$$

For example, if $F(x, y) = \alpha x + \beta y$ is a linear projection, then

$$dF(\mathbf{v}|_{(x,y)}) = (a\alpha + b\beta)\frac{d}{ds}\bigg|_{s=\alpha x+\beta y}.$$

Lemma 1.31. *If $F: M \to N$ and $H: N \to P$ are smooth maps between manifolds, then*

$$d(H \circ F) = dH \circ dF, \tag{1.25}$$

where $dF: TM|_x \to TN|_{y=F(x)}$, $dH: TN|_y \to TP|_{z=H(y)}$, and $d(H \circ F): TM|_x \to TP|_{z=H(F(x))}$.

The proof is immediate from either of the two definitions. In local coordinates, (1.25) just says that the Jacobian matrix of the composition of two functions is the product of their respective Jacobian matrices.

It is important to note that if $\mathbf{v}$ is a vector field on M, then in general $dF(\mathbf{v})$ will *not* be a well-defined vector field on N. For one thing, $dF(\mathbf{v})$ may not be defined on all of N; for another, if two points x and $\tilde{x}$ in M are mapped to the same point $y = F(x) = F(\tilde{x})$ in N, there is no guarantee that $dF(\mathbf{v}|_x)$ and $dF(\mathbf{v}|_{\tilde{x}})$ (both of which are in $TN|_y$) are the same. For instance, if $\mathbf{v} = y\partial_x + \partial_y$ and $s = F(x, y) = \alpha x + \beta y$ is the projection of Example 1.30, then

$$dF(\mathbf{v}|_{(x,y)}) = (\alpha y + \beta)\frac{d}{ds}\bigg|_{s=\alpha x+\beta y},$$

which is not a well-defined vector field on $\mathbb{R}$ unless $\alpha = 0$. However, if F is a diffeomorphism onto N, then $dF(\mathbf{v})$ is always a vector field on N. More generally, two vector fields $\mathbf{v}$ on M and $\mathbf{w}$ on N are said to be *F-related* if $dF(\mathbf{v}|_x) = \mathbf{w}|_{F(x)}$ for all $x \in M$. If $\mathbf{v}$ and $\mathbf{w} = dF(\mathbf{v})$ are F-related, then F maps integral curves of $\mathbf{v}$ to integral curves of $\mathbf{w}$, with

$$F(\exp(\varepsilon\mathbf{v})x) = \exp(\varepsilon\, dF(\mathbf{v}))F(x). \tag{1.26}$$

Lie Brackets

The most important operation on vector fields is their Lie bracket or commutator. This is most easily defined in terms of their actions as derivations on functions. Specifically, if $\mathbf{v}$ and $\mathbf{w}$ are vector fields on M, then their *Lie bracket* $[\mathbf{v}, \mathbf{w}]$ is the unique vector field satisfying

$$[\mathbf{v}, \mathbf{w}](f) = \mathbf{v}(\mathbf{w}(f)) - \mathbf{w}(\mathbf{v}(f)) \tag{1.27}$$