Lectures on Lie Groups and Differential Equations

Peter J. Olver†
School of Mathematics
University of Minnesota
Minneapolis, MN 55455
olver@math.umn.edu
http://www.math.umn.edu/∼olver

Chapter 1

Geometric Foundations

The study of symmetry groups and equivalence problems requires a variety of tools and techniques, many of which have their origins in geometry. Even our study of differential equations and variational problems will be fundamentally geometric in nature, in contrast to the analytical methods of importance in existence and uniqueness theory. We therefore begin our exposition with a brief review of the basic prerequisites from differential geometry which will be essential to the proper development of our subject. These include the definition and fundamental properties of manifolds and submanifolds, of vector fields and flows, and of differential forms. Even though most of our concerns will be local, nevertheless it will be extremely useful to adopt the coordinate-free language provided by the geometric framework. The advantage of this approach is that it frees one from excessive reliance on complicated local coordinate formulas. On the other hand, when explicit computations need to be done in coordinates, one has the added advantage of being able to choose a particular coordinate system adapted to the problem at hand.

Manifolds

A manifold is an object which, locally, just looks like an open subset of Euclidean space, but whose global topology can be quite different. Although most of our manifolds are realized as subsets of Euclidean space, the general definition is worth reviewing. Although almost all the important examples and applications deal with analytic manifolds, many of the constructions are valid for smooth, meaning infinitely differentiable (C^∞), manifolds, and it is this context that we take as our primary domain of exposition, restricting to the analytic category only when necessary.

---

† *Supported in part by NSF Grant DMS* 11–08894.

Definition 1.1. An m-dimensional *manifold* M is a topological space which is covered by a collection of open subsets $W_\alpha \subset M$, called *coordinate charts*, and one-to-one maps $\chi_\alpha : W_\alpha \to V_\alpha$ onto connected open subsets $V_\alpha \subset \mathbb{R}^m$, which serve to define local coordinates on M. The manifold is *smooth* (respectively *analytic*) if the composite "overlap maps" $\chi_{\beta\alpha} = \chi_\beta \circ \chi_\alpha^{-1}$ are smooth (respectively analytic) where defined, i.e., from $\chi_\alpha[W_\alpha \cap W_\beta]$ to $\chi_\beta[W_\alpha \cap W_\beta]$.

The *topology* of a manifold is induced by that of $\mathbb{R}^m$. Thus, a subset $V \subset M$ is *open* if and only if its intersection $V \cap W_\alpha$ with every coordinate chart maps it to an open subset $\chi_\alpha[V \cap W_\alpha] \subset \mathbb{R}^m$, thereby making each χ_α a continuous invertible map. We will always assume that our manifolds are *separable*, meaning that there is a countable dense subset, and satisfy the Hausdorff topological separation axiom, meaning that any two distinct points $x \neq y$ in M can be separated by open subsets $x \in V$, $y \in W$, with empty intersection, $V \cap W = \varnothing$. A manifold is *connected* if it cannot be written as the disjoint union of two nonempty open subsets; many of our results will require connectivity of the manifolds in question. More generally, every manifold is the union of a countable number of disjoint connected components.

Besides the basic coordinate charts provided in the definition of a manifold, one can always adjoin many additional coordinate charts $\chi: M \to V \subset \mathbb{R}^m$ subject to the condition that, where defined, the corresponding overlap maps $\chi_\alpha \circ \chi^{-1}$ satisfy the same smoothness or analyticity requirements as M itself. For instance, composing a given coordinate map with any local diffeomorphism, meaning a smooth, one-to-one map defined on an open subset of $\mathbb{R}^m$, will give a new set of local coordinates. Often one expands the collection of coordinate charts to include all possible compatible charts, the resulting maximal collection defining an *atlas* on the manifold M. In practice, it is convenient to omit explicit reference to the coordinate maps χ_α and identify a point of M with its image in $\mathbb{R}^m$. Thus, the points in the coordinate chart W_α are identified with their local coordinate expressions $x = (x^1, \ldots, x^m) \in V_\alpha$. The changes of coordinates provided by the overlap maps are then given by local diffeomorphisms $y = \eta(x)$ defined on the overlap of the two coordinate charts.

Objects defined on manifolds must be defined intrinsically, independent of any choice of local coordinate. Consequently, manifolds provide us with the proper category in which most efficaciously to develop a coordinate-free approach to the study of their intrinsic geometry. Of course, the explicit formulae for the object may change when one goes from one set of coordinates to another. Thus, in one sense, any equivalence problem can be viewed as the problem of determining whether two different local coordinate expressions define the same intrinsic object on the manifold. In this language, the problem of determining canonical forms is that of finding local coordinates in which the object assumes a particularly simple form. Explicit examples of this general, underlying philosophy will appear throughout the book.

Example 1.2. The basic example of a manifold is, of course, $\mathbb{R}^m$ itself, or any open subset thereof, which is covered by a single coordinate chart. Another simple example is provided by the unit sphere $S^{m-1} = \{ |x| = 1 \} \subset \mathbb{R}^m$, which is an analytic manifold of dimension $m-1$. It can be covered by two coordinate charts, obtained by omitting the north and south poles respectively; the local coordinates are provided by stereographic projection to $\mathbb{R}^{m-1}$. Alternatively, one can use local spherical coordinates on S^{m-1}, which are valid away from the poles.

Example 1.3. Another important example is provided by the projective space $\mathbb{RP}^{m-1}$ which is defined as the space of lines through the origin in $\mathbb{R}^m$. Two nonzero points $x, y \in \mathbb{R}^m$ determine the same point $p \in \mathbb{RP}^{m-1}$ if and only if they are scalar multiples of each other: $x = \lambda y$, $\lambda \neq 0$. (Alternatively, we can realize $\mathbb{RP}^{m-1}$ by identifying antipodal points on the sphere S^{m-1}.) Coordinate charts on $\mathbb{RP}^{m-1}$ are constructed by considering all lines with a given component, say x^i, nonzero; the coordinates are then provided by ratios $p^k = x^k / x^i$, $k \neq i$, which amounts to the choice of canonical representative of such a line given by normalizing its i^{th} component to unity. In particular, the one-dimensional projective space $\mathbb{RP}^1$ can be identified with a circle, since each line through the origin in $\mathbb{R}^2$ is uniquely specified by the angle $0 \leq \theta < \pi$ it makes with the horizontal. The coordinate chart consisting of nonvertical lines, i.e., with $y \neq 0$, has local coordinate $p = x/y$, and covers all but one point on the projective line; the horizontal x-axis is traditionally identified with the "point at infinity". In this way one regards $\mathbb{RP}^1 = \mathbb{R} \cup \{\infty\}$ as the completion of the real line $\mathbb{R}^1$. In a similar fashion $\mathbb{RP}^{m-1}$ is viewed as the completion of $\mathbb{R}^{m-1}$ obtained by adjoining all "directions at infinity". Similar considerations apply to the complex projective space $\mathbb{CP}^{m-1}$, which is defined as the space of all complex lines through the origin in $\mathbb{C}^m$. The complex projective line $\mathbb{CP}^1$ can be identified with the Riemann sphere S^2, which is obtained by adjoining a single point at infinity to the complex plane $\mathbb{C}$.

Exercise 1.4. Prove that if M and N are manifolds of respective dimensions m and n, then their Cartesian product $M \times N$ is an $(m+n)$-dimensional manifold.

Example 1.5. Let M be a manifold. A *vector bundle* over M is a manifold E whose local coordinate charts are of the form $W_\alpha \times \mathbb{R}^q$, where $W_\alpha \subset M$ is a local coordinate chart on M, so that the local coordinates have the form (x, u), where $x \in \mathbb{R}^m$ are referred to as the base coordinates and $u \in \mathbb{R}^q$ the fiber coordinates. The overlap functions are restricted to be linear in the fiber coordinates, $(y, v) = (\eta(x), \mu(x) u)$, where $\mu(x)$ is an invertible $q \times q$ matrix of functions defined on the overlap $W_\alpha \cap W_\beta$ of the two coordinate charts. Thus, locally a vector bundle looks like the Cartesian product $M \times \mathbb{R}^q$, although its global topology might be quite different. Every vector bundle comes with a natural projection $\pi: E \to M$ to its base manifold, defined by $\pi(x, u) = x$ in local coordinates. Two simple examples are provided by a cylinder $S^1 \times \mathbb{R}$, and a Möbius band, both of which are vector bundles with one-dimensional fiber (i.e., line bundles) over the circle S^1.

Although it is useful to have the full vocabulary of manifold theory at our disposal, many of our results will be local in nature. By the phrase "locally" we shall generally mean in a neighborhood U, usually a coordinate neighborhood, of a point $x_0 \in M$. Many of our maps $F: M \to N$ will only be defined locally, i.e., not on the entire manifold M but rather on an open subset $U \subset M$; nevertheless, it is convenient to retain the notation $F: M \to N$ even when the domain of F is a proper subset of M. In such cases, when we write $F(x)$ we are always implicitly assuming that x lies in the domain of F.

Functions

A map $F: M \to N$ between smooth manifolds is called smooth if it is smooth in local coordinates. In other words, given local coordinates $x = (x^1, \ldots, x^m)$ on M, and $y = (y^1, \ldots, y^n)$ on N, the map has the form $y = F(x)$, or, more explicitly, $y^i = F^i(x^1, \ldots, x^m)$, $i = 1, \ldots, n$, where $F = (F^1, \ldots, F^n)$ is a C^∞ map from an open subset of $\mathbb{R}^m$ to $\mathbb{R}^n$. The definition readily extends to analytic maps between analytic manifolds.

Definition 1.6. The *rank* of a map $F: M \to N$ at a point $x \in M$ is defined to be the rank of the $n \times m$ Jacobian matrix $(\partial F^i / \partial x^j)$ of any local coordinate expression for F at the point x. The map F is called *regular* if its rank is constant.

Standard transformation properties of the Jacobian matrix imply that the definition of rank is independent of the choice of local coordinates. In particular, the set of points where the rank of F is maximal is an open submanifold of the manifold M (which is dense if F is analytic), and the restriction of F to this subset is regular.

The first of the equivalence problems which we encounter, then, is to determine whether two different maps $y = F(x)$ and $\bar{y} = \overline{F}(\bar{x})$ between manifolds of the same dimension are locally the same, meaning that they can be transformed into each other by appropriate changes of coordinates $\bar{x} = \chi(x)$, $\bar{y} = \psi(y)$. In the regular case, the Implicit Function Theorem solves the (local) equivalence problem, and, in fact, provides a canonical form for regular maps.

Theorem 1.7. *Let $F: M \to N$ be a regular map of rank r. Then there exist local coordinates $x = (x^1, \ldots, x^m)$ on M and $y = (y^1, \ldots, y^n)$ on N such that F takes the canonical form*

$$y = F(x) = (x^1, \ldots, x^r, 0, \ldots, 0). \tag{1.1}$$

Thus, all maps of constant rank are locally equivalent and can be linearized by the introduction of appropriate local coordinates. The places where the rank of a map decreases are *singularities*. The canonical forms at singularities are much more complicated — this is the subject of singularity theory (a.k.a. catastrophe theory), cf. [**1**, **18**]. In order to keep the scope manageable, this book will be exclusively concerned with the regular cases — in this instance meaning regular maps — and will steer away from the more complicated (but perhaps even more interesting) investigation of singularities.

An important case occurs when M and N have the same dimension, and $F: M \to N$ is a regular map of rank $m = \dim M = \dim N$. The Inverse Function Theorem (which is a special case of Theorem 1.7) shows that F defines a local diffeomorphism between M and N, hence the inverse image $F^{-1}\{y\}$ of any point $y \in N$ is a discrete collection of points in M. If F is defined on all of $M =$ domain F, and, moreover, is onto $N =$ range F, then we shall call F a *covering map* and say that M *covers* N. For example, the map $F(t) = (\cos t, \sin t)$ provides a covering map from the real line $M = \mathbb{R}$ to the circle $N = S^1$. The reader should be warned that our definition of covering map is more general than the standard one, cf. [**54**; p. 98], in that N is not necessarily covered evenly by M — the cardinality of the inverse image $F^{-1}\{y\}$ of $y \in N$, can vary from point to point. For instance, according to our definition, the restriction of any covering map to any open subset $\widetilde{M} \subset M$ remains a covering map; thus the preceding example remains a covering map when restricted to any open interval of length at least 2π.

The notion of rank has a natural generalization to (infinite) families of smooth functions. First recall that the *differential* of a smooth function $f: M \to \mathbb{R}$ is given by the expression

$$df = \sum_{i=1}^{m} \frac{\partial f}{\partial x^i} \, dx^i. \tag{1.2}$$

Definition 1.8. Let $\mathcal{F}$ be a family of smooth real-valued functions $f: M \to \mathbb{R}$. The *rank* of $\mathcal{F}$ at a point $x \in M$ is the dimension of the space spanned by their differentials. The family is *regular* if its rank is constant on M.

Definition 1.9. A set $\{f_1, \ldots, f_k\}$ of smooth real-valued functions on a manifold M having a common domain of definition is called *functionally dependent* if, for each $x_0 \in M$, there is neighborhood U and a smooth function $H(z_1, \ldots, z_k)$, not identically zero on any subset of $\mathbb{R}^k$, such that $H(f_1(x), \ldots, f_k(x)) = 0$ for all $x \in U$. The functions are called *functionally independent* if they are not functionally dependent when restricted to any open subset of M.

For example $f_1(x,y) = x/y$ and $f_2(x,y) = xy/(x^2 + y^2)$ are functionally dependent on the upper half plane $\{y > 0\}$ since the second can be written as a function of the first: $f_2 = f_1/(1+f_1^2)$. For a regular family of functions, the rank tells us how many functionally independent functions it contains.

Theorem 1.10. *If a family of functions $\mathcal{F}$ is regular of rank r, then, in a neighborhood of any point, there exist r functionally independent functions $f_1, \ldots, f_r \in \mathcal{F}$ with the property that any other function $f \in \mathcal{F}$ can be expressed as a function thereof: $f = H(f_1, \ldots, f_r)$.*

Proof: Given $x_0 \in M$, choose $f_1, \ldots, f_r \in \mathcal{F}$ such that their differentials $df_1, \ldots, df_r$ are linearly independent at x_0, and hence, by continuity, in a neighborhood of x_0. According to Theorem 1.7, we can locally choose coordinates (y,z) near x_0 such that $f_i(y,z) = y^i$, $i = 1, \ldots, r$. If $f(y,z)$ is any other function in $\mathcal{F}$, then, since the rank is r, its differential must be a linear combination of the differentials df_i, so that in the new coordinates $df = \sum_{i=1}^{r} h_i(y,z) dy^i$. We now invoke a simple lemma.

Lemma 1.11. *Let $U \subset \mathbb{R}^m$ be a convex open set. A function $f: U \to \mathbb{R}$ has differential $df = \sum_{i=1}^{r} h_i(x)\, dx^i$ given as a linear combination of the first r coordinate differentials if and only if $f = f(x^1, \ldots, x^r)$ is a function of the first r coordinates. In particular, $df = 0$ if and only if f is constant.*

Thus, by shrinking the neighborhood of x_0 if necessary so that it is convex in the (y,z)-coordinates, Lemma 1.11 implies that $f(y^1, \ldots, y^r)$ is a function of y alone. Reverting to the original x coordinates completes the proof of the Theorem. *Q.E.D.*

Consequently, if $f_1, \ldots, f_r$ is a set of functions whose $m \times r$ Jacobian matrix $(\partial f_i / \partial x^k)$ has maximal rank r at x_0, then, by continuity, they also have rank r in a neighborhood of x_0, and hence are functionally independent near x_0. Note also that Theorem 1.10 implies that, locally, there are at most m functionally independent functions on any m-dimensional manifold M.

Submanifolds

Naïvely, a submanifold of a manifold M is a subset which is a manifold in its own right. To be more precise, we need to parametrize the subset by a suitable map.

Definition 1.12. A smooth (analytic) n-dimensional *immersed submanifold* of a manifold M is a subset $N \subset M$ parametrized by a smooth (analytic), one-to-one map $F: \widetilde{N} \to N \subset M$, whose domain $\widetilde{N}$, the *parameter space*, is a smooth (analytic) n-dimensional manifold, and such that F is everywhere regular, of maximal rank n.

A submanifold is *regular* if, in addition to the regularity of the parametrizing map, at each point $x \in N$ there exist arbitrarily small neighborhoods $x \in U \subset M$ such that $F^{-1}[U \cap N]$ is a connected open subset of $\widetilde{N}$. The Implicit Function Theorem 1.7 provides an immediate canonical form for regular submanifolds.

Theorem 1.13. *An n-dimensional submanifold $N \subset M$ of an m-dimensional manifold M is regular if and only if for each $x_0 \in N$ there exist local coordinates $x = (x^1, \ldots, x^m)$ defined on a neighborhood U of x_0 such that*

$$U \cap N = \{ \, x \mid x^1 = \cdots = x^{m-n} = 0 \, \} .$$

Thus, every regular n-dimensional submanifold of an m-dimensional manifold locally looks like an n-dimensional subspace of $\mathbb{R}^m$. We conclude that all regular n-dimensional submanifolds are locally equivalent. The topology on M induces the manifold topology on N, and so, in the regular case, we can effectively dispense with reference to the parameter space $\widetilde{N}$. Irregular submanifolds are trickier, since the same subset can be parametrized as a submanifold in several inequivalent ways — see Example 1.14 below.

It will occasionally be useful to enlarge our repertoire of submanifolds yet further to include submanifolds which have self-intersections. Consider an arbitrary regular map $F: \widetilde{N} \to N \subset M$ of maximal rank n, which is the dimension of the parameter space $\widetilde{N}$. We will call the image $N = F(\widetilde{N})$ a *submanifold with self-intersections* of M. According to Theorem 1.7, the image N of such a map F is, locally, a regular submanifold, but if F is not one-to-one, the submanifold N will intersect itself.

A *curve* in a manifold M is defined by a smooth map $\phi: I \to M$, where $I \subset \mathbb{R}$ is an open subinterval. The curve $C = \phi(I)$ defines a one-dimensional submanifold, with self-intersections, provided the regularity condition $\phi'(t) \neq 0$ holds. (Points at which the derivative of ϕ vanishes will, in general, be singularities of the curve.) If ϕ is one-to-one, then C is an ordinary one-dimensional submanifold, and is regular provided $\lim_{i\to\infty} \phi(t_i) = \phi(t)$ if and only if $\lim_{i\to\infty} t_i = t$.

Example 1.14. The curve $\phi_0(t) = \big(\sin t, 2\sin(2t)\big)$ describes a figure eight in the plane, which is a one-dimensional manifold with self-intersections. The map $\phi_1(t) = \big(\sin(2\arctan t), 2\sin(4\arctan t)\big)$ parametrizes the same figure eight, but is now one-to-one, and so describes an immersed submanifold, which, however, is not regular since

$\lim_{t\to\infty}\phi(t) = 0 = \phi(0)$. Note further that the alternative one-to-one map $\phi_2(t) = \bigl(-\sin(2\arctan t), 2\sin(4\arctan t)\bigr)$ defines an inequivalent means of parametrizing the figure eight. Thus, in the irregular case, a given subset can be parametrized as a submanifold in fundamentally different ways.

Exercise 1.15. Prove the map $\phi(t) = (e^{-t}\cos t, e^{-t}\sin t)$ describes a regular one-dimensional submanifold $N \subset \mathbb{R}^2$. Draw a picture of N.

Example 1.16. A *torus* is defined as the Cartesian product of circles. Consider the two-dimensional torus $T^2 = S^1 \times S^1$, with angular coordinates (θ, φ) with $0 \le \theta, \varphi < 2\pi$. The curve $\phi(t) = (t, \kappa\, t) \bmod 2\pi$ is closed if κ/π is a rational number, and hence defines a regular submanifold of T^2, with parameter space S^1. On the other hand, if κ/π is irrational, then the curve forms a dense subset of T^2 and hence is not a regular submanifold.

Exercise 1.17. Let M and N be smooth manifolds of respective dimensions m and n. Prove that the *graph* $\Gamma = \{(x, F(x))\}$ of a smooth map $F: M \to N$ (which might only be defined on an open subset of M) forms a regular m-dimensional submanifold of the Cartesian product manifold $M \times N$.

Example 1.18. The preceding construction is generalized by the concept of a section of a vector bundle $\pi: E \to M$. A map $F: M \to E$ is called a *section* of E if $\pi \circ F = \mathbb{1}$ is the identity map on M, and, in the local coordinates $(x, u) \in W_\alpha \times \mathbb{R}^q$ of Example 1.5, it has the form $u = F(x)$ where $F: W_\alpha \to \mathbb{R}^q$ is a smooth function. The image $F(M)$ of a section is a smooth m-dimensional submanifold of E that intersects each fiber $E|_x = \pi^{-1}\{x\}$ in only one point. (*Exercise*: Is every submanifold having the latter property a section?)

Since manifolds are modeled on Euclidean space, it is not hard to see that every connected manifold is *pathwise connected* meaning that any two points can be joined by a smooth curve. A manifold is *simply connected* if every smooth curve can be smoothly contracted to a point. For example, the circle S^1 and the punctured plane $\mathbb{R}^2 \setminus \{0\}$ are not simply connected, whereas the sphere S^2 is simply connected.

An alternative to the parametric approach to submanifolds is to define them *implicitly* as a common level set of a collection of functions. In general, the *variety* $\mathcal{S}_\mathcal{F}$ determined by a family of real-valued functions $\mathcal{F}$ is defined to be the subset where they simultaneously vanish:

$$\mathcal{S}_\mathcal{F} = \{\, x \mid\ f(x) = 0 \text{ for all } f \in \mathcal{F}\,\}.$$

In particular, given $F: M \to \mathbb{R}^r$, the variety $\mathcal{S}_F = \{F(x) = 0\}$ is just the set of solutions to the simultaneous system of equations $F^1(x) = \cdots = F^r(x) = 0$ defined by the components of F. We will call the variety (or system of equations) *regular* if $\mathcal{S}_\mathcal{F}$ is not empty, and the rank of $\mathcal{F}$ is constant in a neighborhood of $\mathcal{S}_\mathcal{F}$; the latter condition clearly holds if $\mathcal{F}$ itself is a regular family. In particular, regularity holds if the variety is defined by the vanishing of a map $F: M \to \mathbb{R}^r$ which has maximal rank r at each point $x \in \mathcal{S}_F$, i.e., at each solution x to the system of equations $F(x) = 0$. The Implicit Function Theorem 1.10, coupled with Theorem 1.13, shows that a regular variety is a regular submanifold.

Theorem 1.19. *Suppose $\mathcal{F}$ is a family of functions defined on an m-dimensional manifold M. If the associated variety $\mathcal{S}_{\mathcal{F}} \subset M$ is regular, then it defines a regular submanifold of dimension $m - r$.*

For example, the function $F(x, y, z) = x^2 + y^2 + z^2 - 1$ has rank one everywhere except at the origin, and hence its variety — the unit sphere — is a regular two-dimensional submanifold of $\mathbb{R}^3$. On the other hand, the function $F(x, y, z) = x^2 + y^2$ has rank zero on its variety — the z-axis — but rank one elsewhere; thus Theorem 1.19 does not apply (even though its variety is a submanifold, albeit of the "wrong" dimension). Finally, the function $F(x, y, z) = xyz$ is also not regular, and, in this case, its variety — the three coordinate planes — is not a submanifold.

Vector Fields

A *tangent vector* to a manifold M at a point $x \in M$ is geometrically defined by the tangent to a (smooth) curve passing through x. In local coordinates, the tangent vector $\mathbf{v}|_x$ to the parametrized curve $x = \phi(t)$ is determined by the derivative: $\mathbf{v}|_x = \phi'(t)$. The collection of all such tangent vectors forms the *tangent space* to M at x. Each tangent space $TM|_x$ is a vector space of the same dimension as M; they are "sewn" together in the obvious manner to form the *tangent bundle* $TM = \bigcup_{x \in M} TM|_x$, which is a $2m$-dimensional manifold, and forms a vector bundle over the m-dimensional manifold M.

Exercise 1.20. Prove that the tangent bundle $TS^1 \simeq S^1 \times \mathbb{R}$ to a circle is a cylinder. On the other hand, the tangent bundle to a sphere is not a trivial Cartesian product: $TS^2 \neq S^2 \times \mathbb{R}^2$.

A *vector field* $\mathbf{v}$ is a smoothly (or analytically) varying assignment of tangent vectors $\mathbf{v}|_x \in TM|_x$, i.e., a section of the tangent bundle TM. We write the local coordinate formula for a vector field in the form

$$\mathbf{v} = \sum_{i=1}^{m} \xi^i(x) \frac{\partial}{\partial x^i}\,, \tag{1.3}$$

where the coefficients $\xi^i(x)$ are smooth (analytic) functions. The motivation for adopting the standard derivational notation in (1.3) will appear presently. In particular, the tangent vectors to the coordinate axes are denoted by $\partial/\partial x^i = \partial_{x^i}$, and form a basis for the tangent space $TM|_x$ at each point in the coordinate chart. If $y = \eta(x)$ is any change of coordinates, then the vector field (1.3) is, in the y coordinates, re-expressed using the basic change of variables formula

$$\mathbf{v} = \sum_{j=1}^{m} \left(\sum_{i=1}^{m} \xi^i(x) \frac{\partial \eta^j}{\partial x^i} \right) \frac{\partial}{\partial y^j}\,, \tag{1.4}$$

where the coefficients are evaluated at $x = \eta^{-1}(y)$. Equation (1.4) is a direct consequence of the chain rule applied to the original definition of tangent vectors to curves.

A parametrized curve $\phi: \mathbb{R} \to M$ is called an *integral curve* of the vector field $\mathbf{v}$ if its tangent vector coincides with the vector field $\mathbf{v}$ at each point; this requires that its parametrization $x = \phi(t)$ satisfy the first order system of ordinary differential equations

$$\frac{dx^i}{dt} = \xi^i(x), \qquad i = 1, \ldots, m. \tag{1.5}$$

Standard existence and uniqueness theorems for systems of ordinary differential equations imply that, through each point $x \in M$, there passes a unique, maximal integral curve. We shall employ the suggestive notation $\phi(t) = \exp(t\mathbf{v})x$ to denote the maximal integral curve passing through $x = \exp(0\mathbf{v})x$ at $t = 0$; the curve $\exp(t\mathbf{v})x$ may or may not be defined for all t.

Motivated by fluid mechanics, where $\mathbf{v}$ is the fluid velocity vector field, the family of (locally defined) maps $\exp(t\mathbf{v})$ is known as the *flow* generated by the vector field $\mathbf{v}$, which is classically referred to as the *infinitesimal generator* of the flow. The flow obeys standard exponential rules:

$$\begin{aligned} &\exp(t\mathbf{v})\exp(s\mathbf{v})x = \exp\big[(t+s)\mathbf{v}\big]x, \qquad && \exp(0\mathbf{v})x = x, \\ &\exp(t\mathbf{v})^{-1}x = \exp(-t\mathbf{v})x, && \frac{d}{dt}\exp(t\mathbf{v})x = \mathbf{v}\big|_{\exp(t\mathbf{v})x}, \end{aligned} \tag{1.6}$$

the first and third equations holding where defined. Conversely, given a flow obeying the first two equations in (1.6), we can reconstruct its generating vector field by differentiation:

$$\mathbf{v}|_x = \frac{d}{dt}\exp(t\mathbf{v})x\,\bigg|_{t=0}, \qquad x \in M. \tag{1.7}$$

In other words, identifying tangent vectors in Euclidean space with ordinary vectors, we have the local coordinate expansion

$$\exp(t\mathbf{v})x = x + t\mathbf{v}|_x + \mathrm{O}(t^2), \tag{1.8}$$

for the flow. Thus, the theory of first order systems of autonomous ordinary differential equations (1.5) is the same as the theory of flows of vector fields.

A point x where the vector field vanishes, $\mathbf{v}|_x = 0$, determines a *singularity* or *equilibrium point*. In this case, x is a fixed point under the induced flow: $\exp(t\mathbf{v})x = x$, for all t. Points at which $\mathbf{v}$ is not zero are called *regular*. The existence of flows implies that, away from singularities, all vector fields look essentially the same. Indeed, as with regular maps and submanifolds, there is a simple canonical form for a vector field in a neighborhood of any regular point.

Theorem 1.21. *Let $\mathbf{v}$ be a vector field defined on M. If x_0 is not a singularity of $\mathbf{v}$, so $\mathbf{v}|_{x_0} \neq 0$, then there exist local rectifying coordinates $y = (y^1, \ldots, y^m)$ near x_0 such that $\mathbf{v} = \partial/\partial y^1$ generates the translational flow $\exp(t\mathbf{v})y = (y^1 + t, y^2, \ldots, y^m)$.*

Theorem 1.21 provides a solution to the basic equivalence problem for systems of autonomous first order systems of ordinary differential equations. This result states that, away from equilibrium points, all such systems are locally equivalent, since they can all be mapped to the elementary system $dy^1/dt = 1$, $dy^2/dt = \cdots = dy^m/dt = 0$ by a suitable change of variables. Indeed, if $\xi^1(x_0) \neq 0$, then the rectifying y coordinates are defined so that $x_0 = 0$, and each $x = \exp(t y^1)(0, y^2, \ldots, y^m)$ lies on a unique integral curve emanating from the hyperplane $\{y^1 = 0\}$. However, Theorem 1.21 is of little practical use for actually solving the system of ordinary differential equations governing the flow, since finding the change of variables required to place the system in canonical form is essentially the same problem as solving it in the first place.

Example 1.22. Consider the following three vector fields on $M = \mathbb{R}$. First, the vector field ∂_x generates the translation flow $\exp(t\partial_x)x = x + t$. The vector field $x\partial_x$ generates the scaling flow $\exp(tx\partial_x)x = e^t x$. In this case, away from the singularity at $x = 0$, the rectifying coordinate is given by $y = \log|x|$, in terms of which the vector field takes the form $\mathbf{v} = \partial_y$. Finally, the vector field $x^2\partial_x$ generates the "inversional" flow $\exp(tx^2\partial_x)x = x/(1 - tx)$. Note that this vector field only generates a local flow on $\mathbb{R}$. In this case, the rectifying coordinate is $y = 1 - 1/x$.

Exercise 1.23. Prove that the three vector fields in Proposition 2.48 can be extended to the projective line $\mathbb{RP}^1$ and, in fact, define global flows there. (See [**37**, **47**] for general results on globalizing flows and group actions.)

The problem of equivalence of vector fields at singularities is much more delicate. There is a large body of literature on the determination of normal forms for vector fields near equilibrium points; see, for instance, [**17**], [**19**]. The global equivalence problem is also of interest, but again more delicate since topological data come into play. For example, unless the Euler characteristic of M is trivial, every smooth vector field must have at least one singularity; see [**51**; Chapter 11].

Applying a vector field $\mathbf{v}$ to a function $f\colon M \to \mathbb{R}$ determines the infinitesimal change in f under the induced flow:

$$\mathbf{v}(f(x)) = \sum_{i=1}^{n} \xi^i(x) \frac{\partial f}{\partial x^i} = \frac{d}{dt} f(\exp(t\mathbf{v})x) \bigg|_{t=0}.$$

In this way, vector fields act as *derivations* on the smooth functions, meaning that they are linear and satisfy a Leibniz Rule:

$$\mathbf{v}(f + g) = \mathbf{v}(f) + \mathbf{v}(g), \qquad \mathbf{v}(fg) = f\mathbf{v}(g) + g\mathbf{v}(f). \tag{1.9}$$

In particular, vector fields annihilate constant functions: $\mathbf{v}(c) = 0$. Indeed, an alternative definition of the tangent space $TM|_x$ is as the space of derivations on the smooth functions defined in a neighborhood of x, cf. [**51**; Chapter 3]. The action of the flow on a function can be reconstructed from its generating vector field using the *Lie series* expansion

$$f(\exp(t\mathbf{v})x) = f(x) + t\mathbf{v}(f(x)) + \tfrac{1}{2}t^2\mathbf{v}(\mathbf{v}(f(x))) + \cdots, \tag{1.10}$$

which, in the analytic case, converges for t near 0.

Lie Brackets

The most important operation on vector fields is the Lie bracket or commutator.

Definition 1.24. Given vector fields $\mathbf{v}$ and $\mathbf{w}$ on a manifold M, their *Lie bracket* is the vector field $[\mathbf{v},\mathbf{w}]$ which satisfies $[\mathbf{v},\mathbf{w}]f = \mathbf{v}(\mathbf{w}(f)) - \mathbf{w}(\mathbf{v}(f))$ for any smooth function $f\colon M \to \mathbb{R}$.

The fact that the Lie bracket is a well-defined vector field rests on the readily verified fact that the commutator of two derivations is itself a derivation. In local coordinates, if

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

then

$$[\mathbf{v},\mathbf{w}] = \sum_{i=1}^{m} \bigl\{ \mathbf{v}(\eta^i) - \mathbf{w}(\xi^i) \bigr\} \frac{\partial}{\partial x^i} = \sum_{i=1}^{m}\sum_{j=1}^{m} \left\{ \xi^j\,\frac{\partial \eta^i}{\partial x^j} - \eta^j\,\frac{\partial \xi^i}{\partial x^j} \right\} \frac{\partial}{\partial x^i}\,. \tag{1.11}$$

The Lie bracket is anti-symmetric, $[\mathbf{v},\mathbf{w}] = -[\mathbf{w},\mathbf{v}]$, and bilinear; moreover it satisfies the important *Jacobi identity*

$$[\mathbf{u},[\mathbf{v},\mathbf{w}]] + [\mathbf{v},[\mathbf{w},\mathbf{u}]] + [\mathbf{w},[\mathbf{u},\mathbf{v}]] = 0, \tag{1.12}$$

for any triple of vector fields $\mathbf{u},\mathbf{v},\mathbf{w}$.

The Lie bracket between two vector fields can be identified as the infinitesimal generator of the commutator of the two associated flows. This interpretation is based on the local coordinate expansion (see (1.8))

$$\exp(-t\mathbf{w})\exp(-t\mathbf{v})\exp(t\mathbf{w})\exp(t\mathbf{v})x = x + t^2[\mathbf{v},\mathbf{w}]\big|_x + \mathrm{O}(t^3), \tag{1.13}$$

which follows directly from (1.10). For example, the Lie bracket of the scaling and translation vector fields on $\mathbb{R}$ is $[\partial_x, x\partial_x] = \partial_x$, reflecting the non-commutativity of the operations of translation and scaling. In particular, if the two flows commute, then the Lie bracket of their infinitesimal generators is necessarily zero. This statement admits an important converse, which is a consequence of the basic existence and uniqueness theorems for ordinary differential equations — see [**43**; Theorem 1.34].

Theorem 1.25. *Let* $\mathbf{v}$, $\mathbf{w}$ *be vector fields on a manifold* M*. Then* $[\mathbf{v},\mathbf{w}] = 0$ *if and only if* $\exp(t\mathbf{v})\exp(s\mathbf{w})x = \exp(s\mathbf{w})\exp(t\mathbf{v})x$ *for all* $x \in M$ *all* $t,s \in V$ *where* $V \subset \mathbb{R}^2$ *is a connected open subset of the* (t,s) *plane containing* $(0,0)$ *and such that both sides of the preceding equation are defined at all* $(t,s) \in V$*.*

The Differential

A smooth map $F\colon M \to N$ between manifolds will map smooth curves on M to smooth curves on N, and thus induce a map between their tangent vectors. The result is a linear map $dF\colon TM|_x \to TN|_{F(x)}$ between the tangent spaces of the two manifolds, called the *differential* of F. More specifically, if the parametrized curve $\phi(t)$ has tangent vector $\mathbf{v}|_x = \phi'(t)$ at $x = \phi(t)$, then the image curve $\psi(t) = F[\phi(t)]$ will have tangent vector

$\mathbf{w}|_y = dF(\mathbf{v}|_x) = \psi'(t)$ at the image point $y = F(x)$. Alternatively, if we regard tangent vectors as derivations, then we can define the differential by the chain rule formula

$$dF\,(\mathbf{v}|_x)\,[h(y)] = \mathbf{v}[h \circ F(x)] \qquad \text{for any} \qquad h\colon N \to \mathbb{R}. \tag{1.14}$$

In terms of local coordinates,

$$dF\,(\mathbf{v}|_x) = dF\left(\sum_{i=1}^{m} \xi^i \frac{\partial}{\partial x^i}\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} \xi^i \frac{\partial F^j}{\partial x^i}\right)\frac{\partial}{\partial y^j}\,. \tag{1.15}$$

(Note that the change of variables formula (1.4) is a special case of (1.15).) Consequently, the differential dF defines a linear map from $TM|_x$ to $TN|_{F(x)}$, whose local coordinate matrix expression is just the $n \times m$ Jacobian matrix $\left(\partial F^j/\partial x^i\right)$ of F at x. In particular, the *rank* of the map F can now be defined intrinsically as the rank of the linear map determined by its differential dF. Note that the differential of the composition of two maps $F\colon M \to N$, and $H\colon N \to P$, is just the linear composition of the two differentials: $d(H \circ F) = dH \circ dF$; this fact is merely the coordinate-free version of the usual chain rule for Jacobian matrices.

An important remark is that, in general, unless F is one-to-one, its differential dF does *not* map vector fields to vector fields. Indeed if $\mathbf{v}$ is a vector field on M and x and $\tilde{x}$ are two points in M with the same image $F(x) = F(\tilde{x})$ in N, there is no reason why $dF(\mathbf{v}|_x)$ should necessarily agree with $dF(\mathbf{v}|_{\tilde{x}})$. However, if $\mathbf{v}$ *is* mapped to a well-defined vector field $dF(\mathbf{v})$ on N, then the two flows match up, meaning

$$F[\exp(t\mathbf{v})x] = \exp(t\,dF(\mathbf{v}))\,F(x), \tag{1.16}$$

where defined. Moreover, the differential dF respects the Lie bracket operation:

$$dF\big([\mathbf{v},\mathbf{w}]\big) = [dF(\mathbf{v}), dF(\mathbf{w})], \tag{1.17}$$

whenever $dF(\mathbf{v})$ and $dF(\mathbf{w})$ are well-defined vector fields on N. This is a consequence of the commutator formulation; it can also be verified directly from (1.11), (1.15). A particular consequence is that the Lie bracket preserves tangent vector fields to submanifolds.

Proposition 1.26. *Suppose* $\mathbf{v}$ *and* $\mathbf{w}$ *are vector fields which are tangent to a submanifold* $N \subset M$*. Then their Lie bracket* $[\mathbf{v},\mathbf{w}]$ *is also tangent to* N*.*

Example 1.27. The vector fields $\mathbf{v} = x\partial_y - y\partial_x$, $\mathbf{w} = y\partial_z - z\partial_y$, generate the rotational flows around the z- and the x-axis respectively. They are both tangent to the unit sphere $S^2 \subset \mathbb{R}^3$, hence their Lie bracket $[\mathbf{v},\mathbf{w}] = x\partial_z - z\partial_x$, which generates the rotational flow around the y-axis, is also tangent to S^2.

Vector Field Systems

In general, by a *vector field system* we mean a set $\mathcal{V}$ of vector fields on a manifold M which forms a linear space under the operations of addition and multiplication by smooth functions. Therefore, we require that if $\mathbf{v},\mathbf{w} \in \mathcal{V}$, and $f, h \in C^\infty(M)$, then $f\mathbf{v} + h\mathbf{w} \in \mathcal{V}$. In most applications, $\mathcal{V}$ is finitely generated by vector fields $\mathbf{v}_1, \ldots, \mathbf{v}_r$, and so consists of all linear combinations $\sum_i f_i \mathbf{v}_i$ with arbitrary smooth functions f_i for coefficients. We define $\mathcal{V}|_x$ to be the subspace of $TM|_x$ spanned by the tangent vectors $\mathbf{v}|_x$ for all $\mathbf{v} \in \mathcal{V}$; in the finitely generated case $\mathcal{V}|_x = \operatorname{Span}\{\mathbf{v}_1|_x, \ldots, \mathbf{v}_r|_x\}$.

Definition 1.28. A submanifold $N \subset M$ is called an *integral submanifold* of the vector field system $\mathcal{V}$ if and only if its tangent space $TN|_x$ is contained in the subspace $\mathcal{V}|_x$ for each $x \in N$.

The *rank* of the vector field system at a point $x \in M$ is, by definition, the dimension of the subspace $\mathcal{V}|_x$. The dimension of any integral submanifold passing through x, then, is bounded by the rank of the system at x. In the standard approach, one looks exclusively for integral submanifolds of maximal dimension, meaning ones whose dimension equals the rank of a vector field system at each of its points. However, we have chosen to keep the more general definition so as to correspond more closely to the differential form case. Often the rank of the vector field system is assumed to be constant, and so all the maximal integral submanifolds have the same dimension, but the general vector field version of Frobenius' Theorem does not require this extra hypothesis.

Definition 1.29. A system of vector fields $\mathcal{V}$ is called *integrable* if through every point $x \in M$ there passes an integral submanifold of dimension $n = \dim \mathcal{V}|_x$.

Note that if a vector field system $\mathcal{V}$ is integrable, and $\mathbf{v}$ is a vector field having the property that it is tangent to every n-dimensional integral submanifold, then $\mathbf{v}$ necessarily belongs to $\mathcal{V}$. Consequently, an immediate necessary condition for the integrability of a vector field system is provided by the fact that the Lie bracket of any two vector fields tangent to a submanifold is also tangent to the submanifold, cf. Proposition 1.26. Thus, if $\mathbf{v}, \mathbf{w}$ are any two vector fields in the system, their Lie bracket $[\mathbf{v}, \mathbf{w}]$ must be tangent to each integral submanifold, and hence belong to the system. With this in mind, we make the following definition.

Definition 1.30. A system of vector fields $\mathcal{V}$ is *involutive* if, whenever $\mathbf{v}, \mathbf{w} \in \mathcal{V}$ are any two vector fields in $\mathcal{V}$, their Lie bracket $[\mathbf{v}, \mathbf{w}]$ also belongs to $\mathcal{V}$.

In the case that $\mathcal{V}$ is finitely generated, the basic properties (1.9) of the Lie bracket imply that we need only check the involutivity condition on a basis for the system. Therefore, a vector field system generated by $\mathbf{v}_1, \ldots, \mathbf{v}_r$ is involutive if and only if there exist smooth functions $a_{ij}^k(x)$, $i, j, k = 1, \ldots r$, such that

$$[\mathbf{v}_i, \mathbf{v}_j] = \sum_{k=1}^{r} a_{ij}^k \, \mathbf{v}_k, \qquad i, j = 1, \ldots, r. \tag{1.18}$$

Note that the vector fields $\mathbf{v}_i$ generating the system need *not* span a Lie algebra of vector fields since we are not necessarily requiring that the coefficients a_{ij}^k in (1.18) be constant.

Frobenius' Theorem

Frobenius' Theorem for vector field systems states that a regular system is integrable if and only if it is involutive. The precise statement and proof will be given at the beginning of the following chapter. A simple example, though, is provided by a vector field system generated by a single vector field. Such a system is automatically involutive. On the other hand, the basic existence theorem for systems of ordinary differential equations provides the required integral curves (and equilibrium points) of the vector field. Therefore the system is also automatically integrable. Thus, for a vector field system generated by a single vector field, Frobenius' Theorem reduces to the usual existence theorem for systems of ordinary differential equations.

Let M be a smooth manifold of dimension m and consider a regular vector field system $\mathcal{V}$ of constant rank n. In this case, any integral submanifold N of the system can have dimension at most n. Frobenius' Theorem says that, for constant rank systems in involution, this dimension is actually attained, and the vector field system is integrable.

Theorem 1.31. *Let $\mathcal{V}$ be a system of* smooth *vector fields on a manifold M of constant rank n. Then $\mathcal{V}$ is* integrable *if and only if it is* involutive.

Remark: If the vector field system is finitely generated, or consists of analytic vector fields, then the theorem remains true even if the rank varies: If $\mathcal{V}$ has rank $n = n(x)$ at x then there exists an n-dimensional integral submanifold N passing through x, and, moreover, at every point of N the rank of the vector field system is the same, namely n; see [**23**].

Proof: Let $x_0 \in M$. We introduce local coordinates

$$x = (y, z) = (y^1, \ldots, y^n, z^1, \ldots, z^{m-n})$$

by first translating x_0 to the origin, and then applying a linear transformation so that the subspace $\mathcal{V}|_0 \subset TM|_0$ corresponding to the vector field system at $x_0 = 0$ is spanned by the first n coordinate tangent vectors $\partial/\partial y^1, \ldots, \partial/\partial y^n$. By continuity, for $x = (y, z)$ in a neighborhood of $x_0 = 0$, the subspace $\mathcal{V}|_x \subset TM|_x$ is spanned by vector fields of the form

$$\widehat{\mathbf{v}}_i = \frac{\partial}{\partial y^i} + \sum_{l=1}^{m-n} \xi_i^l(y, z) \frac{\partial}{\partial z^l}, \qquad i = 1, \ldots, n, \tag{1.19}$$

where, at $x_0 = (0, 0)$, the coefficients satisfy $\xi_i^l(0, 0) = 0$ for all i, l. The Lie bracket of any two of these vector fields has the form

$$[\, \widehat{\mathbf{v}}_i, \widehat{\mathbf{v}}_j \,] = \sum_{l=1}^{m-n} \eta_{ij}^l(y, z) \frac{\partial}{\partial z^l}, \qquad i, j = 1, \ldots, n. \tag{1.20}$$

In order that the system be involutive, each vector field (1.20) must be a linear combination of the vector fields (1.19). However, owing to the form of the vector fields $\widehat{\mathbf{v}}_i$, this can only happen if all the coefficients are 0, so $\eta_{ij}^l(y, z) \equiv 0$ for all i, j, l. Therefore the vector fields pairwise commute: $[\, \widehat{\mathbf{v}}_i, \widehat{\mathbf{v}}_j] = 0$.

Let $\exp(t\, \widehat{\mathbf{v}}_i)$ denote the flow of the i^{th} vector field. According to Theorem 1.25, commutativity of the vector fields implies that the flows commute, so that, where defined,

$$\exp(t\, \widehat{\mathbf{v}}_i) \circ \exp(s\, \widehat{\mathbf{v}}_j) = \exp(s\, \widehat{\mathbf{v}}_j) \circ \exp(t\, \widehat{\mathbf{v}}_i). \tag{1.21}$$

Our desired integral submanifolds will be found by starting at a given point and successively flowing in all directions prescribed by the vector fields (1.19). More explicitly, for $(t,s) = (t_1, \ldots, t_n, s_1, \ldots, s_{m-n})$ in a neighborhood U of the origin in $\mathbb{R}^m$, we define the map $\Phi: U \to M$ by

$$x = \Phi(t,s) = \exp(t_1 \widehat{\mathbf{v}}_1) \circ \exp(t_2 \widehat{\mathbf{v}}_2) \circ \cdots \circ \exp(t_n \widehat{\mathbf{v}}_n)(0,s), \tag{1.22}$$

i.e., we start at the point $x = (0,s)$ and flow an amount t_j by the vector field $\widehat{\mathbf{v}}_j$. Note that, by commutativity, the order in which we perform the flows is immaterial. The integral submanifolds will be the images, under the map Φ, of the slices $H_s = \{\, (y,s) \mid y \in \mathbb{R}^n \,\} \subset \mathbb{R}^m$, for $s \in \mathbb{R}^{m-n}$ sufficiently near 0, which implies that (t,s) are the desired rectifying coordinates. More specifically, I claim that *a*) the map $\Phi: V \to M$ is a diffeomorphism in a neighborhood $0 \in V \subset U$, and *b*) for each $s \in \mathbb{R}^{m-n}$, the submanifold $N_s = \Phi[V \cap H_s]$ is an n-dimensional integral submanifold of $\mathcal{V}$. The two claims will suffice to prove the Theorem of Frobenius.

According to the Inverse Function Theorem, the first claim will follow if we show that the differential (Jacobian matrix) of the map Φ is nonsingular at the origin. To compute the t-derivatives of (1.22), we use the commutativity of the flows:

$$\begin{aligned} d\Phi \left(\left. \frac{\partial}{\partial t_i} \right|_{(t,s)} \right) &= \frac{\partial}{\partial t_i} \Big[\exp(t_1 \widehat{\mathbf{v}}_1) \circ \exp(t_2 \widehat{\mathbf{v}}_2) \circ \cdots \circ \exp(t_n \widehat{\mathbf{v}}_n)(0,s) \Big] \\ &= \frac{\partial}{\partial t_i} \exp(t_i \widehat{\mathbf{v}}_i) \Big[\exp(t_1 \widehat{\mathbf{v}}_1) \circ \cdots \circ \exp(t_{i-1} \widehat{\mathbf{v}}_{i-1}) \circ \\ &\qquad\qquad \circ \exp(t_{i+1} \widehat{\mathbf{v}}_{i+1}) \circ \cdots \circ \exp(t_n \widehat{\mathbf{v}}_n)(0,s) \Big] \\ &= \widehat{\mathbf{v}}_i \big|_{\Phi(t,s)}, \end{aligned}$$

the latter equality following from equation (1.6). Therefore,

$$d\Phi \left(\frac{\partial}{\partial t_i} \right) = \widehat{\mathbf{v}}_i, \qquad i = 1, \ldots, n. \tag{1.23}$$

On the other hand, $\Phi(0,s) = (0,s)$ by definition, so

$$d\Phi \left(\left. \frac{\partial}{\partial s_j} \right|_{(0,s)} \right) = \left. \frac{\partial}{\partial s_j} \right|_{(0,s)}, \qquad j = 1, \ldots, m-n. \tag{1.24}$$

In particular, at the origin, $\widehat{\mathbf{v}}_i = \partial / \partial y^i$, and hence equations (1.23, 24) imply that the differential $d\Phi(0,0)$ is the identity matrix, which proves the first claim. To prove the second claim, we note that (1.23) implies that the tangent space to the submanifold is spanned by the vector fields $\widehat{\mathbf{v}}_i$, so $TN_s = d\Phi(TH_s) = \mathcal{V}$. This suffices to prove that N_s is an integral submanifold (of maximal dimension). *Q.E.D.*

Theorem 1.31 and its proof demonstrate the existence of local integral submanifolds for an involutive vector field system. Moreover, just as the integral curves of a single vector field can be extended to maximal integral curves, so can we extend the integral submanifolds to be maximal. In the sequel, the term "integral submanifold" will always refer to the maximal, connected integral submanifolds of the given differential system. The integral submanifolds of a constant rank vector field system $\mathcal{V}$ provide a foliation of the manifold M by n-dimensional submanifolds — see [**39**; §2.11], [**54**] for the details.

Example 1.32. Consider the three vector fields $\mathbf{v}_1 = \partial_x$, $\mathbf{v}_2 = \partial_y$, $\mathbf{v}_3 = x\partial_x + zy\partial_y$, which act on $M = \mathbb{R}^3$. Since $[\mathbf{v}_1, \mathbf{v}_2] = 0$, $[\mathbf{v}_1, \mathbf{v}_3] = \mathbf{v}_1$, $[\mathbf{v}_2, \mathbf{v}_3] = z\mathbf{v}_2$, the system spanned by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is involutive. Indeed, the integral submanifolds are just the planes $N_c = \{z = c\}$ for c constant. Restricted to the integral submanifolds, the vector fields generate nonisomorphic three-parameter planar group actions, $(x, y) \mapsto (\lambda x + \delta, \lambda^z y + \varepsilon)$, corresponding to Case 1.7 with $k = 1$, $z = \alpha$ in our tables. Therefore, $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ do not generate a three-parameter group of transformations on $\mathbb{R}^3$. Moreover, one cannot include these vector fields in a finite-dimensional Lie algebra, since $[\mathbf{v}_2, \mathbf{v}_3] = \mathbf{v}_4 = z\partial_y$, $[\mathbf{v}_4, \mathbf{v}_3] = \mathbf{v}_5 = z^2\partial_y$, and so on, hence the successive commutators span an infinite-dimensional Lie algebra of vector fields.

Exercise 1.33. Suppose $\mathcal{V}$ is a commutative vector field system, so $[\mathbf{v}, \mathbf{w}] = 0$ for all $\mathbf{v}, \mathbf{w} \in \mathcal{V}$. Prove that if $\mathcal{V}$ has constant rank n, then there exist local coordinates $(t_1, \ldots, t_n, s_1, \ldots, s_{m-n})$ such that every vector field $\mathbf{v} \in \mathcal{V}$ has the form $\mathbf{v} = \sum_{i=1}^n \eta^i(s)\, \partial_{t_i}$.

Exercise 1.34. Let $M = \mathbb{R}^2$ and let $\mathcal{V}$ denote the vector field system spanned by the horizontal vector field ∂_x and all vertical vector fields of the form $f(x)\partial_y$ where f is any smooth scalar function such that it and all of its derivatives vanish at $x = 0$, so $f^{(n)}(0) = 0$, $n = 0, 1, 2, \ldots$. Prove that $\mathcal{V}$ has rank one on the y-axis and rank two elsewhere. Prove that any point on the y-axis is contained in a (nonunique) integral curve C transverse to the y-axis. Therefore the rank of $\mathcal{V}$ at any other point in C is strictly greater than one. This example shows that, for infinitely generated vector field systems of variable rank, one must allow integral submanifolds of nonmaximal dimension. See [**38**] for more details.

Differential Forms

The dual objects to vector fields are differential forms. Given a point $x \in M$, a real-valued linear function $\omega: TM|_x \to \mathbb{R}$ on the tangent space is said to define a *one-form* at x. The evaluation of ω on a tangent vector $\mathbf{v}$ will be indicated by the bilinear pairing $\langle\, \omega\,; \mathbf{v}\,\rangle$. The space of one-forms is the dual vector space to the tangent space $TM|_x$, and is called the *cotangent space*, denoted $T^*M|_x$. The cotangent spaces are sewn together to form the *cotangent bundle* $T^*M = \bigcup_{x\in M} T^*M|_x$, which, like the tangent bundle, forms an m-dimensional vector bundle over the m-dimensional manifold M. A *differential one-form* or *Pfaffian form*, then, is just a (smooth or analytic) section of T^*M, i.e., a smoothly varying assignment of linear maps on the tangent spaces $TM|_x$. Given a smooth real-valued function $f: M \to \mathbb{R}$, its differential df, as given in (1.2), determines a one-form whose evaluation on any tangent vector (vector field) $\mathbf{v}$ is defined by $\langle\, df\,; \mathbf{v}\,\rangle = \mathbf{v}(f)$. In local coordinates $x = (x^1, \ldots, x^m)$, the differentials dx^i of the coordinate functions provide a basis of the cotangent space at each point of the coordinate chart, which forms the dual to the coordinate basis ∂_{x^j} of the tangent space; thus $\langle\, dx^i\,; \partial_{x^j}\,\rangle = \delta^i_j$, where δ^i_j denotes the *Kronecker delta*, which is 1 if $i = j$ and 0 otherwise. In terms of this basis, a general one-form takes the local coordinate form

$$\omega = \sum_{i=1}^{m} h_i(x) dx^i, \qquad \text{so that} \qquad \langle\, \omega\,; \mathbf{v} \,\rangle = \sum_{i=1}^{m} h_i(x) \xi^i(x) \tag{1.25}$$

defines its evaluation on the vector field (1.3).

Exercise 1.35. Let $\mathcal{S}_F = \{F_1(x) = \cdots = F_k(x) = 0\} \subset M$ be a regular variety. Prove that its tangent space is the common kernel of the differentials of its defining functions:

$$T\mathcal{S}_F|_x = \{\, \mathbf{v} \in TM|_x \mid \langle\, dF_\nu(x)\,; \mathbf{v} \,\rangle = 0,\ \nu = 1, \ldots, k \,\}\,.$$

Differential forms of higher degree are defined as alternating multi-linear maps on the tangent space. Thus a *differential k-form* Ω at a point $x \in M$ is a k-linear map

$$\Omega\colon\ \overbrace{TM|_x \times \cdots \times TM|_x}^{k \text{ times}} \longrightarrow \mathbb{R},$$

which is anti-symmetric in its arguments, meaning that

$$\langle\, \Omega\,; \mathbf{v}_{\pi 1}, \ldots, \mathbf{v}_{\pi k} \,\rangle = (\operatorname{sign} \pi)\, \langle\, \Omega\,; \mathbf{v}_1, \ldots, \mathbf{v}_k \,\rangle,$$

for any permutation π of the indices $\{1, \ldots, k\}$. We refer to k as the *degree* of the differential form Ω. A real-valued function is considered as a form of degree 0. The space of all k-forms at x is the k-fold *exterior power* of the cotangent space at x, denoted by $\bigwedge^k T^*M|_x$, and forms a vector space of dimension $\binom{m}{k}$. In particular, the only differential form whose degree is greater than the dimension of the underlying manifold is the trivial one $\Omega = 0$. These spaces are sewn together to form the k^{th} exterior tangent bundle $\bigwedge^k T^*M$.

If $\omega^1, \ldots, \omega^k$ are one-forms at x, their *wedge product* defines a *decomposable* k-form $\omega^1 \wedge \cdots \wedge \omega^k$, which is defined by the determinantal formula

$$\langle\, \omega^1 \wedge \cdots \wedge \omega^k\,; \mathbf{v}_1, \ldots, \mathbf{v}_k \,\rangle = \det\Big(\langle\, \omega^i\,; \mathbf{v}_j \,\rangle \Big).$$

In particular, the one-forms $\omega^1, \ldots, \omega^k$ are linearly dependent if and only if their wedge product vanishes: $\omega^1 \wedge \cdots \wedge \omega^k = 0$. Not every k-form is decomposable, although they can all be written as linear combinations of decomposable forms. In local coordinates $x = (x^1, \ldots, x^m)$, the $\binom{m}{k}$ coordinate k-forms $dx^I = dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ corresponding to all strictly increasing multi-indices $1 \leq i_1 < i_2 < \cdots < i_k \leq m$ form a basis for the exterior tangent space $\bigwedge^k T^*M|_x$. A general differential k-form, then, is a section of $\bigwedge^k T^*M$, and, in local coordinates, can be written as a sum

$$\Omega = \sum_I h_I(x)\, dx^I, \tag{1.26}$$

over the basis k-forms with smoothly varying coefficients $h_I(x)$. For example, every m-form on the m-dimensional manifold M is a multiple $h(x)\, dx$ of the standard volume form $dx = dx^1 \wedge \cdots \wedge dx^m$ on $\mathbb{R}^m$.

If $f_1, \ldots, f_k$ are smooth, real-valued functions on M, then the wedge product $df_1 \wedge \cdots \wedge df_k$ of their differentials forms a decomposable k-form on M, which vanishes at a point x if and only if the differentials are linearly dependent there. If we expand this form in terms of the coordinate differentials, as in (1.26), then the coefficient of the basis k-form dx^I is the Jacobian determinant

$$\frac{\partial(f_1, \ldots, f_k)}{\partial(x^{i_1}, \ldots, x^{i_k})} = \det \left(\frac{\partial F_j}{\partial x^{i_l}} \right).$$

Thus, Theorem 1.10 implies the following simple test for functional independence.†

Proposition 1.36. *If $f_1, \ldots, f_k$ satisfy $df_1 \wedge \cdots \wedge df_k \neq 0$, then they are* functionally independent. *On the other hand, if $df_1 \wedge \cdots \wedge df_k \equiv 0$ for all $x \in M$, then $f_1, \ldots, f_k$ are* functionally dependent.

If $\Omega = \omega^1 \wedge \cdots \wedge \omega^k$ and $\Theta = \theta^1 \wedge \cdots \wedge \theta^l$ are decomposable k- and l-forms, we define their *wedge product* to be the decomposable $(k+l)$-form $\Omega \wedge \Theta = \omega^1 \wedge \cdots \wedge \omega^k \wedge \theta^1 \wedge \cdots \wedge \theta^l$; this definition extends by linearity to arbitrary differential forms. The resulting wedge product between differential forms is bilinear and "super-symmetric": $\Omega \wedge \Theta = (-1)^{kl} \Theta \wedge \Omega$. In particular, the wedge product of a 0-form f, otherwise known as a smooth function, and a k-form Ω is the k-form $f\,\Omega$ obtained by multiplying Ω by f.

Remark: We shall *always* assume that our differential forms are of homogeneous degree. This means that, although we are allowed to take the wedge product of any pair of differential forms, we are only allowed to sum differential forms of the *same* degree.

Exercise 1.37. Let $\omega^1, \ldots, \omega^k$ be a set of linearly independent one-forms, so $\omega^1 \wedge \cdots \wedge \omega^k \neq 0$. Prove *Cartan's Lemma*, which states that the one-forms $\theta^1, \ldots, \theta^k$ satisfy $\sum_i \theta^i \wedge \omega^i = 0$ if and only if $\theta^i = \sum_j A^i_j \, \omega^j$ for some symmetric matrix of functions: $A^i_j = A^j_i$.

If $F \colon M \to N$ is a smooth map, then there is an induced map on differential forms, called the *pull-back* of F and denoted F^*, which maps a differential form on N *back* to a differential form on M. (*Warning*: The direction of the pull-back is reversed from that of F and its differential dF.) If $\theta \in T^*N|_y$ is a one-form at $y = F(x)$, then $\omega = F^*\theta \in T^*M|_x$ is a one-form on M defined so that $\langle\, \omega \,;\mathbf{v}\,\rangle = \langle\, F^*\theta \,;\mathbf{v}\,\rangle = \langle\, \theta \,; dF(\mathbf{v})\,\rangle$ for any tangent vector $\mathbf{v} \in TM|_x$. In local coordinates, the pull-back of a one-form θ on N is the one-form

$$F^*\theta = F^*\left(\sum_{j=1}^{n} h_j(y)\, dy^j \right) = \sum_{j=1}^{n} h_j(F(x))\, dF^j(x) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} h_j(F(x))\, \frac{\partial F^j}{\partial x^i} \right) dx^i. \tag{1.27}$$

† Actually, Theorem 1.10 only applies if the rank of $f_1, \ldots, f_k$ is constant. The proof of the second statement in Proposition 1.36 in the more general case when the rank is $< k$, but not necessarily constant, can be found in [**39**; Theorem 1.4.14].

Thus, in local coordinates the pull-back is represented by the transpose of the Jacobian matrix of F. The pull-back extends to arbitrary differential forms by requiring that it commute with addition and the wedge product:

$$F^*(\Omega + \Theta) = F^*(\Omega) + F^*(\Theta), \qquad F^*(\Omega \wedge \Theta) = F^*(\Omega) \wedge F^*(\Theta). \tag{1.28}$$

In particular, the pull-back of a smooth function $h: N \to \mathbb{R}$ (0-form) is given by composition, $F^*(h) = h \circ F$. In contrast to the behavior of vector fields under the differential, the pull-back of a differential k-form on N is *always* a well-defined smooth k-form on M. Note that the pull-back *reverses* the order of composition of maps: If $F: M \to N$ and $H: N \to P$, then $(H \circ F)^* = F^* \circ H^*$.

Exercise 1.38. Prove that the pull-back action of a smooth map $F: \mathbb{R}^m \to \mathbb{R}^m$ on the volume form $dx = dx^1 \wedge \cdots \wedge dx^m$ is given by $F^*(dx) = (\det J)\, dx$ where $J = \left(\partial F^i / \partial x^j\right)$ is the Jacobian matrix of F. Conclude that F is volume-preserving if and only if its Jacobian determinant has value ± 1.

The differential (1.2) of a smooth function has an important and natural extension to differential forms of higher degree. In general, the *differential* or *exterior derivative* of a k-form is the $(k+1)$-form defined in local coordinates by

$$d\Omega = \sum_I dh_I(x) \wedge dx^I = \sum_{I,j} \frac{\partial h_I}{\partial x^j}\; dx^j \wedge dx^I, \tag{1.29}$$

for Ω given by (1.26). For example, if $\omega = \sum h_i\, dx^i$ is a one-form, its differential is the two-form

$$d\omega = \sum_{i=1}^m dh_i \wedge dx^i = \sum_{i<j} \left(\frac{\partial h_j}{\partial x^i} - \frac{\partial h_i}{\partial x^j} \right) \; dx^i \wedge dx^j. \tag{1.30}$$

The differential satisfies the generalized ("super") derivational property

$$d(\Omega \wedge \Theta) = d\Omega \wedge \Theta + (-1)^k\, \Omega \wedge d\Theta, \tag{1.31}$$

whenever Ω is a k-form.

The fact that the differential is coordinate free is perhaps not so clear from the local coordinate definition (1.29). An alternative, intrinsic definition in the case of one-form is via the useful formula

$$\langle\, d\omega\, ; \mathbf{v}, \mathbf{w} \,\rangle = \mathbf{v} \langle\, \omega\, ; \mathbf{w} \,\rangle - \mathbf{w} \langle\, \omega\, ; \mathbf{v} \,\rangle - \langle\, \omega\, ; [\mathbf{v}, \mathbf{w}] \,\rangle, \tag{1.32}$$

valid for any pair of vector fields $\mathbf{v}, \mathbf{w}$. In (1.32), the first term on the right hand side is the result of applying the vector field $\mathbf{v}$ to the function defined by the evaluation of the one-form ω on the vector field $\mathbf{w}$, cf. (1.25). Formula (1.32) has the implication that, in a certain sense, the Lie bracket and the differential are dual operations.

Exercise 1.39. Prove that formula (1.32) agrees with the local coordinate version (1.30). Find a generalization of (1.32) valid for arbitrary k-forms, cf. [**54**].

A *crucial* property of the differential is its invariance under smooth maps. The proof is straightforward, using either local coordinates or the intrinsic approach.

Theorem 1.40. *Let $F: M \to N$ be a smooth map with* pull-back F^*. *If Ω is any differential form on N, then*

$$d[F^*\Omega] = F^*[d\Omega]. \tag{1.33}$$

Suppose $F: \widetilde{N} \to M$ parametrizes a submanifold $N \subset M$. We will call the pull-back $F^*\Omega$ of a differential form Ω on M the *restriction* of Ω to the submanifold N, and denote it by $\Omega \mid N$. Actually, we should call this the restriction of Ω to the parameter space $\widetilde{N}$, but the notation should not cause any confusion; besides, if N is a regular submanifold, we can, as remarked earlier, unambiguously identify N with $\widetilde{N}$.

Corollary 1.41. *Let $N \subset M$ be a submanifold. If Ω is any differential form which vanishes when restricted to N, so $\Omega \mid N = 0$, then so does its differential, $d(\Omega \mid N) = 0$.*

Finally, the anti-symmetry of the wedge product, coupled with the equality of mixed partial derivatives, implies that applying the differential twice in a row always produces zero; see, for instance, (1.30).

Theorem 1.42. *If Ω is any differential form, then*

$$d(d\Omega) = 0. \tag{1.34}$$

This final property provides the foundation of a remarkable, deep connection between the structure of the space of differential forms and the global topology of the underlying manifold. A differential form Ω is said to be *closed* if it has zero differential, $d\Omega = 0$. A k-form Ω is said to be *exact* if it is the differential of a $(k-1)$-form: $\Omega = d\Theta$. Theorem 1.42 implies that every exact form is closed, but the converse may not hold. Indeed, (1.34) implies that the differential defines a "complex", the *deRham complex* of the manifold M, whose cohomology, meaning, very roughly, the extent to which closed forms fail to be exact, depends, remarkably, only on the global topological character of M. For example, the *Poincaré Lemma* states that convex subdomains of Euclidean space have trivial cohomology.

Theorem 1.43. *Let $k > 0$. If $M \subset \mathbb{R}^m$ is a convex open subset, and Ω is any closed k-form defined on all of M, then Ω is exact, so that there exists a $(k-1)$ form Θ on M such that $\Omega = d\Theta$. A closed 0-form, i.e., a function satisfying $df = 0$, is constant.*

Thus, on a general manifold M, every closed form is locally, but perhaps not globally, exact. As our subsequent considerations are primarily local, we will not pursue this fascinating aspect of the theory of differential forms any further, but refer the interested reader to [**10**].

Example 1.44. Let $M = \mathbb{R}^2 \setminus \{0\}$. An easy calculation shows that the one-form $\omega = (x^2 + y^2)^{-1}\,(y\,dx - x\,dy)$ is closed, so $d\omega = 0$. However, there is *no* globally defined smooth function $f(x,y)$ satisfying $df = \omega$, so that ω is not exact. Indeed, locally $\omega = d\theta$ is the differential of the polar angle $\theta = \tan^{-1}(y/x)$, but, of course, θ is not a globally defined single-valued function on M. It is not hard to show that ω is essentially the unique closed one-form with this property; any closed one-form η can be written as $\eta = c\,\omega + df$ for some constant c and some smooth, globally defined function f. This result reflects the fact that M is a topological surface with a single hole.

